# 3D object identification with color and curvature signatures

Adnan A.Y. Mustafa[a],*, Linda G. Shapiro[b], Mark A. Ganter[c]

[a] *Department of Mechanical and Industrial Engineering, Kuwait University, PO Box 5969, Safat-13060, Kuwait*
[b] *Department of Computer Science and Engineering, University of Washington, Seattle, WA 98195, U.S.A.*
[c] *Department of Mechanical Engineering University of Washington Seattle, WA 98195, U.S.A.*

## Abstract

In this paper we describe a model-based object identification system. Given a set of 3D objects and a scene containing one or more of these objects, the system identifies which objects appear in the scene by matching surface signatures. Surface signatures are feature vectors that reflect the probability of occurrence of the features for a given surface. Two types of surface signatures are employed; curvature signatures and spectral (i.e. color) signatures. Furthermore, the system employs an inexpensive acquisition setup consisting of a single CCD camera and two light sources. The system has been tested on 95 observed surfaces and 77 objects with varying degrees of curvature and color with good results. © 1999 Pattern Recognition Society. Published by Elsevier Science Ltd. All rights reserved.

*Keywords*: Object identification; Object recognition; Color photometric stereo; Surface signatures; Surface matching; Color

## 1. Introduction

At the core of any complete vision system is the recognition subsystem that generates hypotheses about objects appearing in a given scene. A hypotheses is usually produced by comparing object data extracted from the scene with a database of model data stored on-line. Any reliable system should produce reliable identification results even if complicated scenes – scenes with object occlusion and shadowing effects – are present. Such scenes should not affect the outcome of the system and satisfactory interpretations should still prevail. Most vision identification systems today are laser-based systems. While such systems have a higher degree of accuracy than other similar systems, they can be quite expensive.

In this paper we describe an inexpensive system that employs a single CCD camera and two light sources for identifying 3D objects appearing in cluttered scenes. A block diagram of the object identification system appears in Fig. 1. The system input consists of two color images of the scene illuminated by two different white light sources. Color photometric stereo is employed to produce a surface normal map of the scene which is then integrated with the spectral data (i.e. color data) to produce a description of the scene. Finally, the system compares the extracted object data (both curvature and spectral data) to model data accessed from the model database (constructed through training) to produce hypotheses about objects appearing in the scene.

Object to model matching is accomplished by finding the best correspondence between an object's surfaces and a model's surfaces. The main features employed for matching are surface characteristics that are invariant to change in pose, are unaffected by shadows from adjacent or occluding objects, and are not sensitive to changes in lighting. In this work, we define one such class of invariant features called *surface signatures*. A surface signature
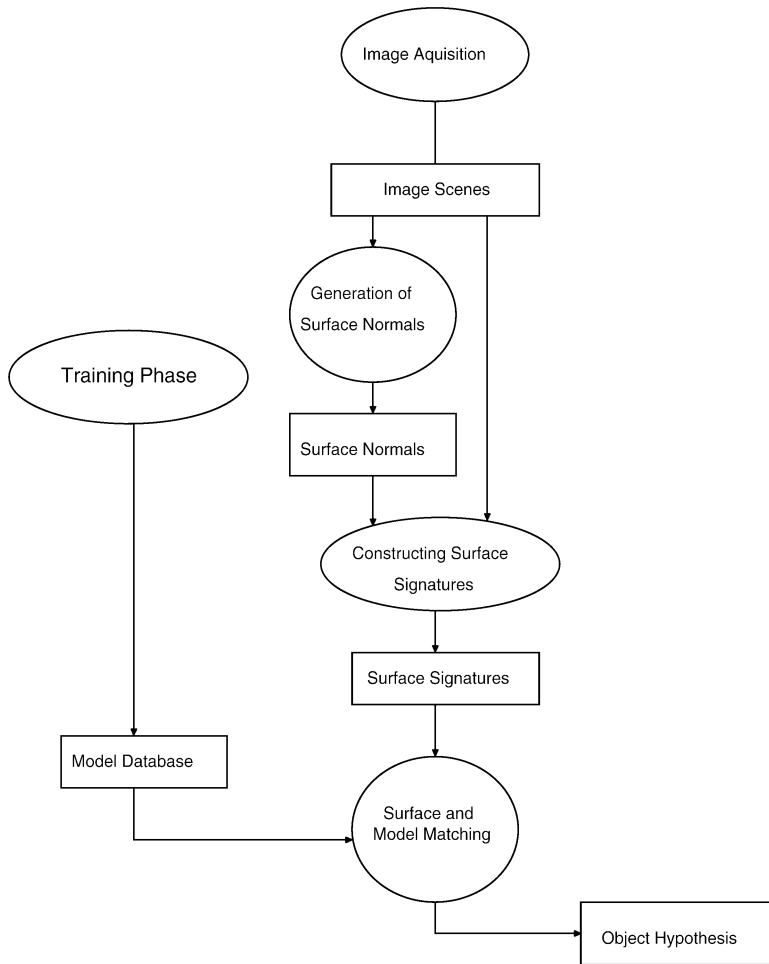
---

[†] Corresponding author.

Fig. 1. Diagram illustrating the object identification system.

is a feature vectors that reflects the probability of occurrence of a feature for a given surface. Currently, two types of surface signatures are employed; curvature signatures and spectral (i.e. color) signatures.

## 2. Related literature

While only recently has the use of color in computer vision started to play an important role in image understanding. The use of color and other spectral bands has been used in the image processing arena for some time, where it is more commonly known as multi-spectral image analysis [1–3]. As the field of computer vision matured and the price of color cameras decreased, the use of color in image analysis has seen a rapid growth during the last decade. As a result, the literature on the use of color in image analysis has expanded, though it is scattered under different subjects [4].

Surface curvature is an intrinsic feature that can play an important role in surface identification. Many strategies for classifying surfaces have been proposed. Besl and Jain [5] employed the mean and Gaussian curvature to classify surfaces extracted from range images to characterize surfaces into eight types. This classification is different from the topographic primal sketch (TPS) proposed by Haralick et al. [6], which labels each pixel of an intensity image surface with one of ten possible topographic labels based on gradients, Hessians and directional derivatives. Newman et al. [7] identified quadric surfaces in segmented range images by employing surface position and surface normal estimates with the known parameters of surfaces in the database.

Photometric stereo is a method that uses two (or more) grayscale images taken under different illumination to provide additional constraints to recover surface orientation. Woodham originally introduced the technique of Photometric stereo [8, 9] by building on the shape from

shading theory developed earlier by Horn [10]. A great deal of research has been done to extend the use of photometric stereo [11–13]. Christensen and Shapiro [14] extended photometric stereo to color images. In color photometric stereo, color is employed to provide additional surface constraints and increase accuracy [15].

Little research has been done that combine both color and curvature for object recognition. Higuchi et al. [16] modified the extended Gaussian image (EGI) to represent not only object curvature but also object color. They called their new structure the *spherical attribute image* (SAI). The SAI maps the values of curvature and color at every node of a mesh (approximating the object surface) onto a spherical image. A matching algorithm matches an observed surface to an object model by finding the rotation that maximizes the sum of two similarity measures (curvature and color).

Grewe and Kak [17] developed a system for the interactive learning of multiple hash tables for fast 3D object recognition. The system uses qualitative attributes of objects such as the shape of the object and quantitative attributes such as color. The system has a training phase where the system is trained on the model identities. A local feature set (LFS) approach to hypothesis generation is employed; a LFS is defined as a vertex of an object and three rotationally adjacent surfaces.

## 3. Surface signatures

*Surface signatures* are employed to characterize object surfaces. A surface signature is a feature vector that reflects the probability of occurrence of a feature for a given surface. In general, if **S** denotes a surface signature of size $N$, and $S_i$ denotes the $i$th component of **S**, then by definition,

- $0 \leqslant S_i \leqslant 1$   where $0 \leqslant i < N - 1$,

- $\sum_{i=0}^{N-1} S_i = 1$.

From this definition it can be seen that surface signatures are normalized histogram distributions of a feature on the surface. The motivation for using signatures is that they are invariant to change in pose and are not affected by partial occlusion or shadows. This is true because signatures are statistical features that are uniform for a uniform surface. This stems from the fact that surface curvature and hue[1] do not change under different environments (see Section 5).

Three curvature signatures are employed to characterize the curvature characteristics of the surface. The *angle curvature signature* ($\mathbf{S}_{AC}$) of a surface is defined as the normalized histogram distribution of the average crease angle (defined below) for the surface. Similarly, the *Gaussian curvature signature* ($\mathbf{S}_{GC}$) and the *mean curvature signature* ($\mathbf{S}_{MC}$) are defined as the normalized histogram distributions of the average Gaussian curvature and the average mean curvature on the surface, respectively. Three spectral signatures are employed to characterize the spectral characteristics of the surface. The *red spectral signature* ($\mathbf{S}_{RS}$) of a surface is defined as the normalized histogram distribution of the red color band. Similarly, the *green spectral signature* ($\mathbf{S}_{GS}$) and the *blue spectral signature* ($\mathbf{S}_{BS}$) are defined as the normalized histogram distributions of the green and blue color bands, respectively. Although only three spectral bands are employed here, the number of spectral signatures can be expanded to include other parts of the electromagnetic spectrum (e.g. infra-red).

### 3.1. Surface curvature signatures

Three curvature attributes are employed in our work to characterize surfaces; the *crease angle*, the *Gaussian curvature* and the *mean curvature*. The crease angle ($\gamma$) between two surface normals, $n_1$ and $n_2$, is defined as the angle between the normals and is calculated through the dot product of the normals,[2]

$$\gamma = \cos^{-1}(n_1 \cdot n_2). \tag{1}$$

The crease angle represents the amount of directional discrepancy between two neighboring normals.[3] Two normals that are parallel, such as those on a flat plate, will have a crease angle equal to zero. Two normals on surfaces that are perpendicular to one another will have a crease angle equal to $90°$.

To reduce the effect of noise and abrupt changes in normals as a result of inaccuracies, the crease angle is averaged in its 8-neighborhood (i.e. values are averaged in $3 \times 3$ window). The average crease angle, $\bar{\gamma}$, at location $(r,c)$ is calculated by

$$\bar{\gamma}_{r,c} = \frac{1}{N_\gamma} \sum_{i=-1}^{1} \sum_{\substack{j=-1 \\ j \neq 0 \text{ if } i=0}}^{1} \gamma_{r+1, c+1}, \tag{2}$$

where $N_\gamma$ is the number of normals defined in the neighborhood. Once the values of $\bar{\gamma}$ have been calculated for

---

[1] Surface specularity does affect object color in the specular region. However, if a specular region is large compared to the rest of the surface then surface spectral signatures are affected significantly.

[2] Crease angles are calculated only at locations where a normal exists.

[3] Note that values are dependent on the resolution (magnification) in which the object appears.

a given surface, a histogram count of quantized $\bar{\gamma}$ values is performed to construct the crease angle signature.

The Gaussian curvature ($K$) is calculated by computing the determinant of the curvature matrix (**C**) defined by

$$K = |\mathbf{C}|, \tag{3}$$

where

$$\mathbf{C} = (p^2 + q^2 + 1)^{-3/2} \begin{bmatrix} q^2 + 1 & -pq \\ -pq & p^2 + 1 \end{bmatrix} \mathbf{H}, \tag{4}$$

$$H = \begin{bmatrix} p_x & p_y \\ q_x & q_y \end{bmatrix}. \tag{5}$$

For our work we have

$$p = \frac{n_x}{n_z} \quad \text{and} \quad q = \frac{n_y}{n_z} \tag{6}$$

and $p_x$, $p_y$, $q_x$ and $q_y$ are by definition,

$$p_x = \frac{\partial p}{\partial x}, \, p_y = \frac{\partial p}{\partial y}, \, q_x = \frac{\partial q}{\partial x} \quad \text{and} \quad q_x = \frac{\partial q}{\partial x} \tag{7}$$

and $\mathbf{n} = (n_x \, n_y \, n_z)'$ is the surface normal at a given point. The mean curvature ($H$) is the average value of the principal curvatures and is calculated by

$$H = \tfrac{1}{2} trace(\mathbf{C}). \tag{8}$$

In our work, we use a $3 \times 3$ neighborhood around the normal to calculate the average value of Gaussian curvature in the neighborhood; we call this the *average Gaussian curvature* ($\bar{K}$) about the normal. Similarly, we use a $3 \times 3$ neighborhood around the normal to calculate the average value of mean curvature in the neighborhood. We call it the *average mean curvature* ($\bar{H}$) about the normal. $\bar{K}$ and $\bar{H}$ are calculated using neighborhood means of $p$, $q$, $p_x$, $p_y$, $q_x$ and $q_y$ in the neighborhood,

$$\bar{p} = \frac{1}{N_p} \sum_{i=-1}^{1} \sum_{j=-1}^{1} p_{i,j}, \tag{9}$$

$$\bar{q} = \frac{1}{N_p} \sum_{i=-1}^{1} \sum_{j=-1}^{1} q_{i,j}, \tag{10}$$

$$\bar{p}_x = \frac{1}{N_x} \sum_{i=-1}^{0} \sum_{j=-1}^{1} (p_{i,j+1} - p_{i,j}), \tag{11}$$

$$\bar{q}_x = \frac{1}{N_x} \sum_{i=-1}^{0} \sum_{j=-1}^{1} (q_{i,j+1} - q_{i,j}), \tag{12}$$

$$\bar{p}_y = \frac{1}{N_y} \sum_{i=-1}^{0} \sum_{j=-1}^{1} (p_{j+1,i} - p_{j,i}), \tag{13}$$

$$\bar{q}_y = \frac{1}{N_y} \sum_{i=-1}^{0} \sum_{j=-1}^{1} (q_{j+1,i} - q_{j,i}), \tag{14}$$

where $\bar{p}, \bar{q}, \bar{p}_x, \bar{p}_y, \bar{q}_x$ and $\bar{q}_y$ are the mean values of $p$, $q$, $p_x$, $p_y$, $q_x$ and $q_y$ in the $3 \times 3$ neighborhood, respectively. $N_p$ is the number of normals existing in the neighborhood. $N_x$ and $N_y$ are the number of defined normal pairs in the $x$ and $y$ direction, respectively. The above equations are employed to obtain $\bar{K}$ and $\bar{H}$,

$$\bar{K} = |\bar{X}|, \tag{15}$$

where

$$\bar{C} = (\bar{p}^2 + \bar{q}^2 + 1)^{-3/2} \begin{bmatrix} \bar{q}^2 + 1 & -\bar{p}\bar{q} \\ -\bar{p}\bar{q} & \bar{p}^2 + 1 \end{bmatrix} \bar{H}, \tag{16}$$

$$\bar{H} = \begin{bmatrix} \bar{p}_x & \bar{p}_y \\ \bar{q}_x & \bar{q}_y \end{bmatrix} \tag{17}$$

and

$$\bar{H} = \tfrac{1}{2} trace(\bar{\mathbf{C}}). \tag{18}$$

We calculate the values of $\bar{K}$ and $\bar{H}$ for all surface normals appearing in the normal map. For a given surface, a histogram count of quantized values of $\bar{K}$ and $\bar{H}$ is performed to construct the Gaussian curvature signature and the mean curvature signature, respectively.

### 3.2. Surface spectral signatures

A spectral signature for a given surface reflects the surface's spectral characteristics for a particular region of the electromagnetic spectrum. Hence, a red spectral signature represents the surface's characteristics with respect to the red region of the electro-magnetic spectrum. Similarly, an infra-red spectral signature reflects the surface's characteristics with respect to the infra-red region of a spectrum, and so on.

## 4. Object models and image scenes

All scenes employed (training and test) consist of two images, a left image and a right image. The left and a right images are taken from the same viewpoint but with the light source to the left and to the right of the scene, respectively. The left and right images of a scene are necessary so that photometric stereo can be employed to generate surface normal maps.

### 4.1. Object models

A model is represented by a data structure containing a description of the model's surface (see Table 1). The structure is constructed during the training process and consists of information, such as the number of surfaces that constitutes the model, surface adjacency data and surface signatures.

Table 1
Object model structure

| Model name |  |  |
| Number of |  |  |
| Surfaces |  |  |
| --- | --- | --- |
| Surface #1 | Surface name |  |
|  | Number of adjacent surfaces |  |
|  | List of adjacent surfaces |  |
|  | Curvature signatures | $\{\mathbf{S}_{AC}, \mathbf{S}_{GC}, \mathbf{S}_{MC}\}$ |
|  | Spectral signatures | $\{\mathbf{S}_{RS}, \mathbf{S}_{GS}, \mathbf{S}_{BS}\}$ |
| Surface #n | Surface name |  |
|  | Number of adjacent surfaces |  |
|  | List of adjacent surfaces |  |
|  | Curvature signatures | $\{\mathbf{S}_{AC}, \mathbf{S}_{GC}, \mathbf{S}_{MC}\}$ |
|  | Spectral signatures | $\{\mathbf{S}_{RS}, \mathbf{S}_{GS}, \mathbf{S}_{BS}\}$ |

### 4.2. Description of the training phase

The objective of the training phase is to train the system on the set of objects appearing in the scenes. Each model surface is analyzed to construct its six signatures. This data is stored in a database which is accessed during the matching stage. We will refer to object data stored in the database as *models* to distinguish them from the unknown scene objects that are to be identified, which we will simply refer to as *objects*.

Surface signatures are generated by analyzing the image scenes and their surface normal map from which the signatures are extracted. Initially the training images are converted into two new images; one containing an approximation to diffuse colors, called the diffuse image and the other containing an approximation to specular colors, called the specular image. The diffuse image is obtained by taking the minimum intensity values of the left and right images. By taking the minimum of the two images, object specularity is removed. The specular image is obtained by taking the maximum intensity values of the left and right images. Taking the maximum intensity values of the two images produces an image equivalent to having both light sources on at the same time, but with shadows removed. The resulting image also contains all specular regions appearing in the original images.

Clustering is performed on a quantized image color space to extract color values appearing in the scene. A variant of the iterative ISODATA clustering method [18] is employed to cluster the data in color space. Clusters are restricted to small variation in color; colors with large variation are split among two or more clusters. The final cluster values produced through clustering are filtered and the remaining clusters represent the RGB

values of colors appearing in the image. Since training scenes consist of single objects, clusters obtained through clustering reflect the colors of the object appearing in the scene. Hence, clustering the colors of the diffuse image produces RGB cluster values which approximate the diffuse reflection constants of the object appearing in the image, while clustering the colors of the specular image produces RGB cluster values which approximate the diffuse reflection constants of the object.

Surface normal maps are generated using a color photometric stereo program [14] developed at the Computer Science and Engineering Department at the University of Washington. The program input is the right and left image of a scene and accepts diffuse and specular reflection constants to generate a surface normal map of the scene. These parameters are employed to determine if a surface normal at an image point can be conceived using the shading information of the left and right images. If the shading information agrees with the parameters specified, then a surface normal is constructed at the point. Normals are filtered to remove inaccuracies and a median filter is applied to smooth the normal map. Surface edges are assumed where the crease angles exceed a threshold value (e.g. $90°$). These edges are then employed to segment the normal map into surfaces from which the surface signatures are constructed [19].

### 4.3. Analyzing scenes

Test scenes differ from training scenes in that any number of objects may appear in a given scene. Objects can be partially occluded or in shadow. Since the objects appearing in the scene are unknown and may be of several different colors, a surface normal map for each color appearing in the scene is constructed. The cluster colors extracted from the specular image are taken as the reflection parameters needed for color photometric stereo. The objects are segmented into surfaces by examining the complete surface map of the scene and identifying surface boundaries where the crease angle exceeds a pre-defined threshold value. Surface signatures are generated for each surface in the same way as that of the training phase producing three curvature signatures and three spectral signatures. In some instances surfaces may become over-segmented. This is generally not a problem as long as the oversegmented surfaces are large enough that sufficient surface characteristics are available to construct adequate representative signatures. Small surfaces are simply discarded by the system. Results of scenes tested are discussed in Section 6.

## 5. Surface matching

Objects appearing in scenes are matched to models by matching their surface signatures. Several matching

metrics and matching errors are employed to arrive at the best match.

## 5.1. Matching metrics

Surface matching consists of two parts; surface curvature matching and surface spectral matching. Surface curvature matching consists of matching its three curvature signatures, while surface spectral matching consists of matching its three spectral signatures.

### 5.1.1. Signature matching

Four signature error metrics are employed to measure the discrepancy between two signatures. These metrics are the *signature distance error*, the *signature shape error*, the *signature spread error*, and the *shape correlation error*.

*5.1.1.1. The signature distance error.* The *signature distance error* ($\varepsilon_\mu$) measures the distance between two signature distributions in space. Given two signature distributions, $\mathbf{S}_1$ and $\mathbf{S}_2$, $\varepsilon_\mu$ is calculated by

$$\varepsilon_\mu(\mathbf{S}_1, \mathbf{S}_2) = \frac{1}{c_n}|\mu_1 - \mu_2|, \tag{19}$$

where $c_n$ is a normalization constant and $\mu_i$ is the mean value of distribution $i$. Values of $\varepsilon_\mu$ lie between zero and unity. Small error values close to zero indicate that the distributions have the same mean value and are distributed around the same frequency points. Error values of unity indicate that the two distributions are accumulated around different frequency points that are far apart in space.

*5.1.1.2. The signature shape error.* The *signature shape error* ($\varepsilon_\sigma$) measures the variance difference between two signatures,

$$\varepsilon_\sigma(\mathbf{S}_1, \mathbf{S}_2) = \begin{cases} 0 & \text{if } \sigma_1 = \sigma_2 = 0, \\ \dfrac{|\sigma_1 - \sigma_2|}{\sigma_1 + \sigma_2} & \text{else,} \end{cases} \tag{20}$$

where $\sigma_i$ is the standard deviation of distribution $i$. Values of $\varepsilon_\sigma$ lie between zero and unity. Error values close to zero indicate that the general shape of the two signature distributions is similar, whereas values close to unity indicate that the general shape of the two distributions are different.

*5.1.1.3. The signature spread error.* The *signature spread error* ($\varepsilon_\lambda$) measures the distribution spread difference between two signatures around their mean value,

$$\varepsilon_\lambda(\mathbf{S}_1, \mathbf{S}_2) = \frac{|\lambda(\mathbf{S}_1, \bar{\sigma}) - \lambda(\mathbf{S}_2, \bar{\sigma})|}{\lambda(\mathbf{S}_1, \bar{\sigma}) + \lambda(\mathbf{S}_2, \bar{\sigma})}, \tag{21}$$

where

$$\lambda(\mathbf{S}, \bar{\sigma}) = \sum_{k=\mu-\bar{\sigma}}^{\mu+\bar{\sigma}} S_k \quad \text{and} \quad \bar{\sigma} = \tfrac{1}{2}(\sigma_1 + \sigma_2). \tag{22}$$

Values of $\varepsilon_\lambda$ lie between zero and unity. Values close to zero indicate that the distribution spread of two signatures around their mean value are similar, whereas error values in the neighborhood of unity indicate large differences.

*5.1.1.4. The signature correlation error.* The *signature correlation error* ($\varepsilon_c$) measures the amount of non-correlation between two signatures,

$$\varepsilon_c(\mathbf{S}_1, \mathbf{S}_2) = 1 - \left| \frac{\sum(\mathbf{S}_1 - \mu_1)(\mathbf{S}_2 - \mu_2)}{\sqrt{\sum(\mathbf{S}_1 - \mu_1)^2}\sqrt{\sum(\mathbf{S}_2 - \mu_2)^2}} \right|. \tag{23}$$

Values of $\varepsilon_c$ lie between zero and unity. Values close to zero indicate strong correlation between two signatures and hence great similarity, whereas error values close to unity indicate poor correlation.

### 5.1.2. Signature match error

The *signature match error* ($E$) is a measure of the discrepancy between two signatures based on the metrics defined above. It is calculated by summing the signature error metrics with appropriate weight factors,

$$E(\mathbf{S}_1, \mathbf{S}_2) = \boldsymbol{\varepsilon}\mathbf{W}^{\mathrm{T}} \tag{24}$$

where $\boldsymbol{\varepsilon} = (\varepsilon_\mu\ \varepsilon_\sigma\ \varepsilon_\lambda\ \varepsilon_c)$ and $\mathbf{W} = (w_\mu\ w_\sigma\ w_\lambda\ w_c)$. $w_\mu$, $w_\sigma$, $w_\lambda$ and $w_c$ are the signature distance, shape, spread and correlation weight factors, respectively, which add up to unity. In Section 7 we will present a method to find the optimum values of these weights.

Error values of $E$ lie between zero and unity. Values close to zero indicate a strong match between two signatures, whereas values close to unity indicate a poor match between two signatures. Note that there are six different signature match errors corresponding to the three curvature signatures and the three spectral signatures.

### 5.1.3. Surface match error

Matching surfaces using curvature signatures gives an indication of how well surfaces match with respect to their geometric attributes. Matching surfaces using spectral signatures gives an indication of similarity based on their spectral attributes. Combining matching results of both curvature signatures and spectral signatures produces a more complete measure of similarity.

### 5.1.4. Surface curvature match error.

The *surface curvature match error* ($E_C$) between two surfaces, $s_1$ and $s_2$, is the sum of the three curvature signature errors between the two surfaces: the angle

curvature signature match error, the Gaussian curvature signature match error, and the mean curvature signature match error.

$E_C$ is calculated by summing the three curvature match errors with appropriate weight factors,

$$E_C(s_1, s_2) = \mathbf{E}_C \mathbf{W}_C^T, \qquad (25)$$

where $\mathbf{E}_C = (E_{AC}\ E_{GC}\ E_{MC},$ and $\mathbf{W}_C = (w_{AC}\ w_{GC}\ w_{MC})$. $E_{AC}$, $E_{GC}$ and $E_{MC}$ are the angle, the Gaussian and the mean signature match errors, respectively. $w_{AC}$, $w_{GC}$ and $w_{MC}$ are the angle, Gaussian and mean curvature signature weight factors, respectively. Error values of $E_C$ lie between zero and unity. Values close to zero indicate strong similarity between surfaces based on curvature features. Large error values close to unity indicate that little similarity based on these features exist between the surfaces.

Specifying values for the weights $w_{AC}$, $w_{GC}$ and $w_{MC}$ depend on the object set being analyzed. For example, if the object set consists of flat surfaces and objects with cylindrical surfaces of different sizes, then the Gaussian curvature signature weight factor $w_{GC}$ should be assigned a value much smaller than the other weight factors since all objects will have Gaussian curvature signature distributions close to zero.[4] In our work, we have found that $w_{AC}$ is more accurate than $w_{GC}$ and $w_{MC}$ because of the errors introduced in the calculations of second order derivatives employed to calculate $w_{GC}$ and $w_{MC}$. In general, for object sets with varying degrees of curvature (flat surfaces, spheres, cones, cylinders) equal weights should be employed. In Section 7 we will present a method to find the optimum values of these weights.

### 5.1.5. Surface spectral match error

The *surface spectral match error* ($E_S$) between two surfaces, $s_1$ and $s_2$, is the sum of the *spectral ratio error* and its three spectral signature errors between the two surfaces: the *red spectral signature match error*, the *green spectral signature match error*, and the *blue spectral signature match error*. The spectral ratio error ($E_{rS}$) is a measure of the RGB color ratios between two surfaces. $E_{rS}$ is defined as,

$$E_{rS}(S_1, S_2) = \left| \frac{\mu_{GS_1}}{\mu_{RS_1}} - \frac{\mu_{GS_2}}{\mu_{RS_2}} \right| + \left| \frac{\mu_{BS_1}}{\mu_{RS_1}} - \frac{\mu_{BS_2}}{\mu_{RS_2}} \right|, \qquad (26)$$

where $\mu_{RS}$, $\mu_{GS}$ and $\mu_{BS}$ are the mean value of the red, green and blue signatures.[5] Error values of $E_{rS}$ lie between zero and unity. Values close to zero indicate strong

similarity between surfaces based on color. Large error values close to unity indicate little similarity based on color. This measure is an important indicator of error, because it measures the hue of the object. When surfaces are slightly darker or lighter due to illumination, the ratio of color bands remain – to a certain degree – constant.

The surface spectral match error ($E_S$) between two surfaces, $s_1$ and $s_2$, is the sum of the three spectral signature errors plus the spectral ratio error between the two surfaces appropriately weighted,

$$E_S(s_1, s_2) = E_s W_s^T, \qquad (27)$$

where $\mathbf{E}_S = (E_{RS}\ E_{GS}\ E_{BS}\ E_{rS})$ and $\mathbf{W}_S = (W_{RS}\ W_{GS}\ W_{BS}\ W_{rS})$. $E_{RS}$, $E_{GS}$ and $E_{BS}$ are the red, green and blue spectral signature match errors, respectively. $w_{RS}$, $w_{GS}$ and $w_{BS}$ are the red, green and blue spectral signature weight factors, respectively. $w_{rS}$ is the spectral ratio weight factor.

$E_S$ error values lie between zero and unity. Values close to zero indicate strong similarity between surfaces based on surface spectral attributes. Large error values close to unity indicate that little similarity exists between surfaces based on surface spectral attributes.

As was the case for weight assignment of the curvature weights for the surface curvature match error, specifying values for the specular weights $w_{RS}$, $w_{GS}$, $w_{BS}$ and $w_{rS}$ also depends on the object set being analyzed. For example, if the object set consists of objects that have surface colors that are red and green and no surfaces with a blue color, then the blue spectral signature weight factor $w_{BS}$ can be assigned a value of zero without any loss of matching accuracy. Generally, weight factors of equal values should be employed. We have found that $E_{rS}$ has consistently given good results in our work, and thus $w_{rS}$ is usually given a higher value than the other spectral weights.

### 5.1.6. Surface match error

The *surface match error* ($\bar{E}$) between two surfaces, $s_1$ and $s_2$, is the sum of the curvature signature match error and the spectral signature match error,

$$\bar{E}(s_1, s_2) = E\mathbf{W}_C^T, \qquad (28)$$

where $\mathbf{E} = (E_C\ E_S)$ and $\mathbf{W} = (w_C\ w_S)$. $w_C$ and $w_S$ are the surface curvature and surface spectral weight factors, respectively, which sum up to unity. $\bar{E}$ error values lie between zero and unity. Values close to zero indicate strong similarity between surfaces based on their surface curvature and spectral attributes. Large error values close to unity indicate poor similarity between the surfaces based on their surface curvature and spectral attributes.

Specifying values for the surface weight factors $w_C$ and $w_S$ depends, once again, on the object set being analyzed.

---

[4] In theory the Gaussian curvature should be exactly zero, but in practice Gaussian curvature values are not exactly zero at all points on the surface due to noise.

[5] If $E_{rS} > 1$, then $E_{rS} = 1$.

If the object set consists of objects that are more easily distinguishable based on object color than object curvature and shape, then the surface curvature weight factor ($w_C$) is more important than the surface spectral weight factor ($w_S$), and hence $w_C$ should have a smaller value than $w_S$. On the other hand, if the objects are similar in color and are more easily distinguishable based on curvature attributes, then $w_C$ should have a larger value than $w_S$. If no clear evidence supports dominance of either curvature or spectral attributes, then the weights should be assigned equal values, $w_C = w_S$. In Section 7 we investigate the effect of varying these weights, as well as the previous weights, on surface matching performance.

## 5.2. Matching and identification strategy

Matching *observed-surfaces* (surfaces extracted from image scenes) to *model-surfaces* (surfaces defined in the model database) is a two-stage process. Initially, the surface match error is calculated for each observed-surface/model-surface pair. The model-surface with the least match error for each observed-surface is marked as the best match for the observed-surface. In the second stage, the surface adjacency graph of the scene is exploited to further facilitate object hypothesis. At this stage it is assumed that all surfaces determined to be adjacent by the surface adjacency procedure[20] do indeed belong to the same object. We assume this to be true, as we have found that occluding surfaces usually cast shadows that prevent observed-surfaces of different objects to be determined as adjacent. This results from the light sources being placed to the right and to the left of the objects appearing in the scene. Based on the surface adjacency graph of an observed-surface, object hypothesis is determined as follows:

- If an observed-surface has no adjacent surfaces, then the object which constitutes the observed-surface as its only surface is identified as the model-surface-match. If the observed-surface has more than one model-match, than the object has more than one model-surface match.
- If the observed-surface has adjacent surfaces to it, then a search is made to find the model which best matches to these surfaces. For any two adjacent observed-surfaces two possibilities exist:

  1. Both observed-surfaces are part of the same object, or
  2. Each observed-surface belongs to a different object (the objects happen to be in close proximity to each other or one is occluding the other).

The first case is easily identified; if adjacent observed-surfaces have the same model-match and the surface adjacency graph is in agreement with the surface adjacency of the observed-surfaces, then they are part of the same object. However, if adjacent observed-surfaces do not have the same model-match, then the model database is examined to see if the adjacency graph can be rejected. For instance, if one of the observed-surfaces has a green color and the other adjacent observed-surface has a red color and there is no model with such a surface structure, then the adjacent graph is determined to be invalid and each observed-surface is assumed to belong to a different object. If the adjacent graph cannot be invalidated, then in order for the two adjacent surfaces to belong to the same model, each observed-surface must produce a surface error match less than a pre-defined error threshold, $\varepsilon_m$. If each pair of adjacent observed-surfaces do not produce a surface match error less than $\varepsilon_m$, then each observed-surface is assumed to belong to a different object, and the model-match of each observed-surface is the model hypotheses of the object of the observed-surfaces.

## 5.3. Performance evaluation

To evaluate the matching performance, the observed-surfaces are matched against their corresponding (correct) model-surfaces and several relative matching measures are evaluated. In this section we define several metrics that are employed to evaluate the surface matching and object identification performances of the analyzed scenes.

### 5.3.1. Relative performance index

The *relative performance index* (*RPI*) of a model-surface to an observed-surface is the relative surface match error of the model-surface to the observed-surface with respect to all model-surface errors in the model set. Hence, the *RPI* of a model-surface indicates how well a model-surface matched to the observed-surface with respect to other model-surfaces in the set. The *RPI* of an observed-surface, $s_1$, and a model-surface, $s_2$, is calculated by

$$RPI(s_1, s_2) = \frac{\bar{E}(s_1, s_2) - \bar{E}_{\min}(s_1)}{\bar{E}_{\max}(s_1) - \bar{E}_{\min}(s_1)} \tag{29}$$

where, $\bar{E}(s_1, s_2)$ is the surface match error between $s_1$ and $s_2$, $\bar{E}_{\min}(s_1)$ is the minimum surface error for $s_1$ in the model set, and $\bar{E}_{\max}(s_1)$ is the maximum surface error for $s_1$ in the model set.

Values of *RPI* lie in the range between zero and unity. A *RPI* value close to zero indicates that the model-surface matched very well to the observed-surface in comparison with the other model surfaces. A *RPI* value close to unity indicates that the model-surface matched poorly to the observed-surface in comparison with the other model surfaces, even though $\bar{E}(s_1, s_2)$ might be small.

### 5.3.2. Model hypothesis error

To evaluate incorrect object hypotheses, we define the *model hypothesis error* ($\Theta$) as a measure of how well the correct model matches to an object appearing in a scene. $\Theta$ for a model, *m*, is defined as the average relative matching error for all its model-surfaces that are visible in the scene,

$$\Theta(m) = \frac{1}{n} \sum_{i=1}^{n} RPI(\Psi^{-1}(s_i), s_i), \qquad (30)$$

where $s_i$ is the surface *i* of model *m* in the scene, $\Psi^{-1}(s)$ is the the correct observed-surface match to the model-surface *s*, and *n* is the number of visible model-surfaces of model *m* in the scene.

Values of $\Theta$ lie in the range between zero and unity. A $\Theta$ value of zero implies an object is correctly hypothesized. A $\Theta$ value of unity indicates that the correct model is the worst hypothesis among all models in the set.

### 5.3.3. Surface error performance index of a scene

The *surface error performance index* (*SEPI*) of a scene is a measure of how well all observed-surfaces match to their correct model-surfaces for a given scene. *SEPI* for a scene is calculated as the average surface match error of all the correct model-surfaces,

$$SEPI = \frac{1}{n} \sum_{i=1}^{n} \bar{E}(s_i, \Psi(s_j)), \qquad (31)$$

where $\Psi(s)$ is the correct surface match to observed-surface *s* and *n* is the number of observed-surfaces in the scene.

Values of *SEPI* lie in the range between zero and unity. A *SEPI* value close to zero indicates that most of the observed-surfaces matched very well to their correct model surfaces, whereas a *SEPI* value close to unity indicates that the observed-surfaces matched poorly to their correct surfaces.

### 5.3.4. Surface matching performance index of a scene

The *surface matching performance index* (*SMPI*) of a scene is a measure of how well all observed-surfaces match to their correct model-surfaces, with respect to all model-surfaces in the set

$$SMPI = (1 - ARPI) \times 100\%, \qquad (32)$$

where

$$ARPI = \frac{1}{n} \sum_{i=1}^{n} RPI(s_i, \Psi(s_j)) \qquad (33)$$

with $\Psi(s)$ being the correct model-surface match to observed-surface *s*, and *n* being the number of observed-surfaces in the scene.

*ARPI* is the average relative performance index of the scene. Values of *SMPI* lie in the range between [0%, 100%]. A value close to 100% indicates that most of the observed-surfaces match very well to their correct model-surfaces, with respect to all model surfaces in the set. A *SMPI* value of 100% indicates that all surfaces match correctly. A low *SMPI* value indicates that the observed-surfaces match poorly to their correct surfaces, and that most observed-surfaces matched to incorrect surfaces. A value of 0% indicates that the correct model-surfaces have the largest surface errors in the model set – the worst possible case.

### 5.3.5. Model hypothesis performance index of a scene

The *model hypothesis performance index* (*MHPI*) of a scene is a measure of how well *correct* models are identified for a given scene. *MHPI* is calculated as the average model hypotheses error of a scene for all correct models in the scene,

$$MHPI = \left(1 - \frac{1}{N} \sum_{i=1}^{N} \Theta(\Omega(m))\right) \times 100\%, \qquad (34)$$

where, $\Omega(m)$ is the correct *model-match* to object *m*, and *N* is the number of objects in the scene.

Values of *MHPI* lie in the range between [0%, 100%]. A *MHPI* value of 100% indicates that all objects are correctly identified for the scene. On the other hand, a value of 0% indicates that all correct models produce the largest hypothesis error in the model set – the worst possible case.

## 6. Experimental results

Testing of the system was conducted on three different model sets with objects of varying degrees of curvature and color. Synthetic images as well as real images were employed. The synthetic images were generated using a raytracing program to produce scenes with realistic shadowing effects. The real image scenes consisted of real 3D objects that were obtained with a grayscale CCD camera and three spectral filters (red, green and blue). Testing was conducted on three different model sets with varying degrees of curvature and color. The sets in increasing complexity are:

- A *Geometry* set consisting of seven simulated geometric objects (*cylinders*, *spheres*, *cone* and *block*). See Fig. 2.
- A *Toy* set consisting of seven real 3D geometric objects (*cylinder*, *elliptical cylinders*, *triangular prism* and *hexagonal prism*). See Fig. 3.
- A *Fruit* set consisting of seven plastic fruit models (an *apple*, a *banana*, *corn*, *grape*, a *lemon*, *pepper*, *strawberry*, and a *tomato*). See Fig. 4.
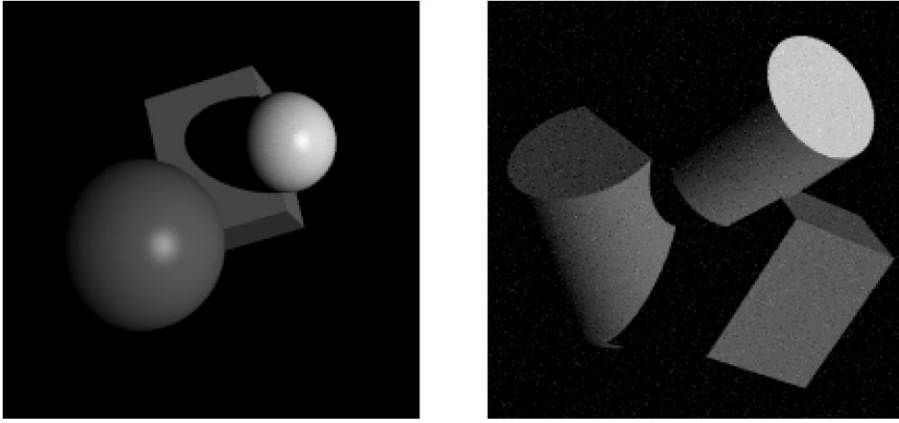
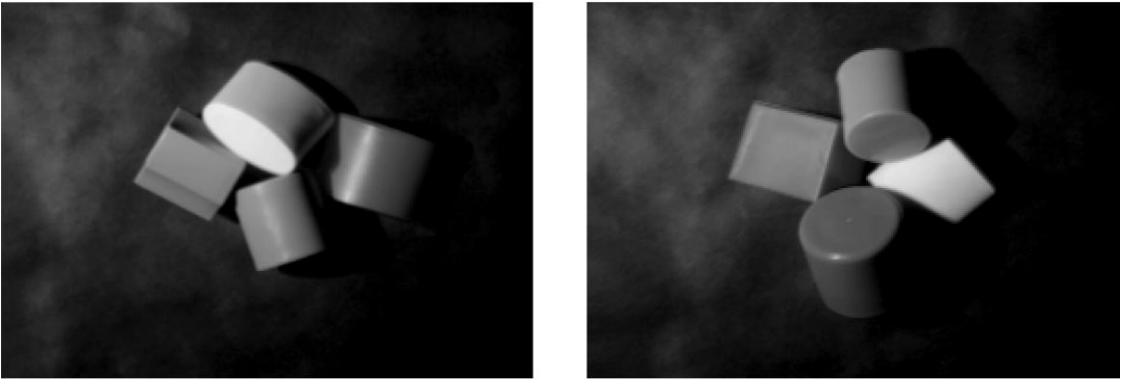Fig. 2. *Scene* #1 and *Scene* #2 of the *Geometry* set.



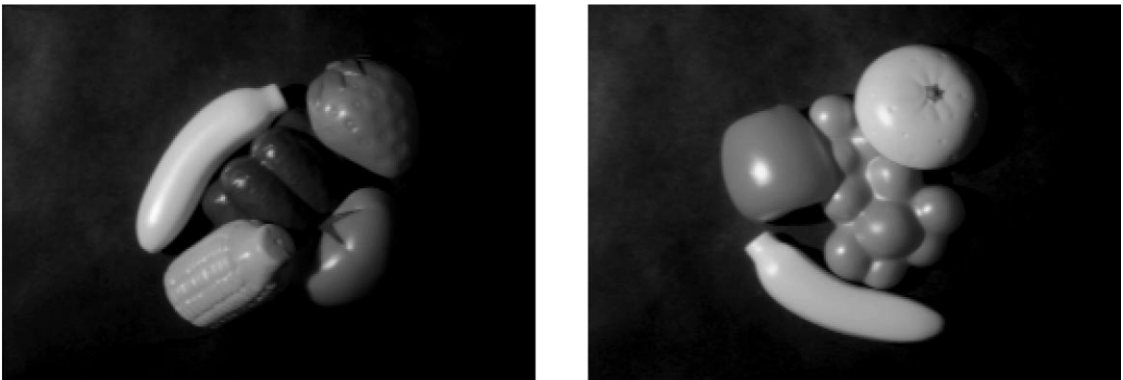Fig. 3. *Scene* #1 and *Scene* #2 of the *Toy* set.



Fig. 4. *Scene* #1 and *Scene* #2 of the *Fruit* set.

Each set consists of several objects with identical colors but different types of surfaces as well as objects with similar surfaces but different colors (e.g. a yellow cone, a yellow cylinder, a blue cylinder). This ensures that both types of signatures (curvature and spectral) will be essential employed to arrive at a correct hypothesis of the object. Fig. 5 shows the surface adjacency graph for two objects of the *Toy* set.
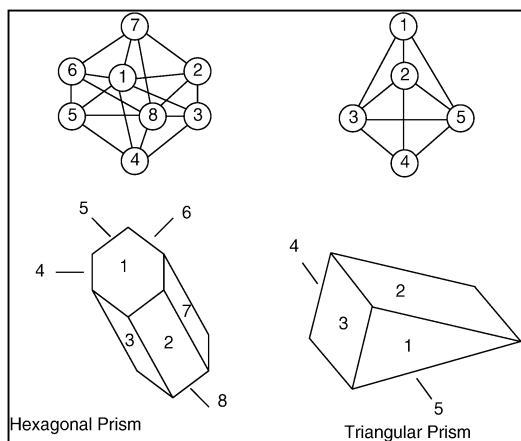
Fig. 5. Surface labeling and surface adjacency graphs for two primitives of the *Toy* set.

### 6.1. A sample test: Fruit Scene #1

A sample test scene from the *Fruit* set, *Fruit Scene* #1, was shown in Fig. 4. The scene consists of five objects; a *banana*, *corn*, *pepper*, *strawberry*, and a *tomato*. Fig. 6 shows the normals found using color photometric stereo and the resulting segmented image with surfaces labeled. From the surface segmented image we see that shadowing effects are evident in the scene; a noticeable effect is that much of the *corn* surface as well as that of the *pepper* surface is in shadow. Figs. 7 and 8 show surface signatures for the banana model and the corn model, respectively. An apparent difference between the signatures of the two models is the staggered distributions of the corn model, as a result of the nature of the faceted surface of the corn object.

Since all models of the *Fruit* set consist of models with a single surface, the model producing the smallest surface match error to the observed-surface is hypothesized as the model match. Table 2 shows the results of matching

*observed-surface*$_1$ to the model-surfaces. We see that matching the curvature signatures of *observed-surface*$_1$ to model curvature signatures resulted in *corn* with the least curvature signature match error (0.03). This is due to the distinct curvature signatures of the corn model as compared to other model signatures, as previously discussed. With the use of spectral signatures, we see that the system was more decisive in eliminating most of the incorrect model-surfaces as these produced large spectral signature match errors ($> 0.12$), while the *corn* model-surface produced the smallest spectral error (0.04). Although there were other yellow colored models (e.g. *banana*, *lemon*), the system was able to correctly distinguish the spectral signatures of the observed-surface as being that of the *corn* model-surface. From the matching result, the system correctly hypothesizes that the object of *observed-surface*$_1$ is the *corn* model.

Table 3 shows the results of matching all observed-surfaces appearing in the scene to their correct model-surfaces. Both the surface match error and the relative performance index for all observed-surfaces appearing in the scene were relatively small (less than 0.1). For this scene the surface error performance index (*SEPI*) was a low 0.06 and the surface matching performance index (*SMPI*) was an excellent 98%.

### 6.2. Summary of results for the geometry set

Seven synthetic scenes with several geometric primitives were simulated using a raytracing program. The attributes of these objects are shown in Table 4. The synthesized objects consist of three cylinders, two spheres, a block and a cone of different sizes and color. Four different types of surfaces were present in these scenes: flat surfaces, cylindrical surfaces, spherical surfaces and conic surfaces. The colors of the objects were: cyan, blue, green, yellow, red and violet. The scenes consisted of 26 observed-surfaces and 21 scene objects. Scenes #1 and #2 were shown in Fig. 2.
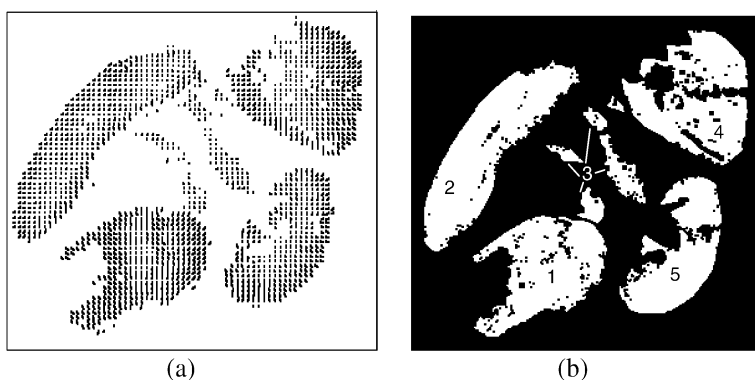


(a)                                        (b)

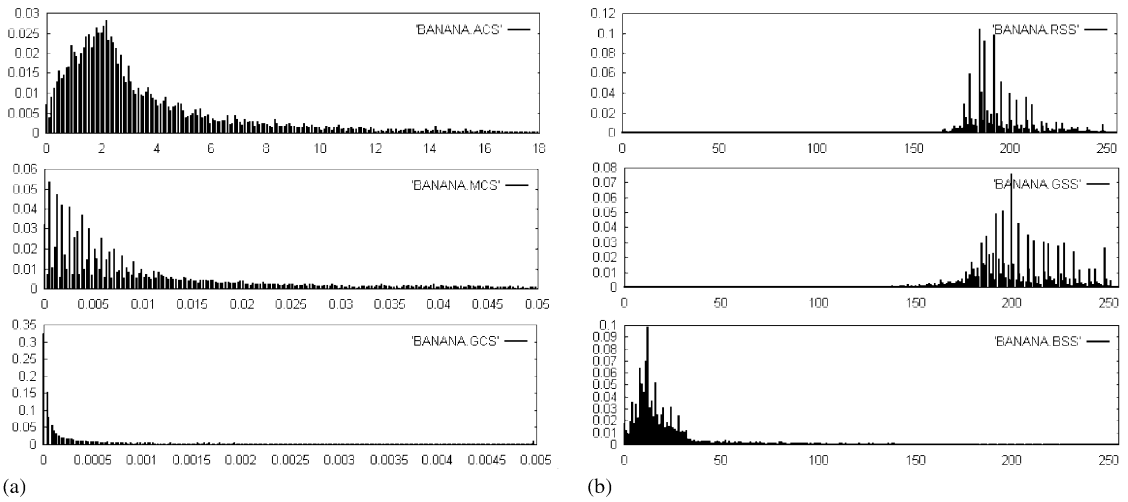Fig. 6. (a) Normal map for *Fruit Scene* #1 (b) segmented normal map with labeling.

Fig. 7. *Banana* signatures. (a) Curvature signatures: from top to bottom; $\mathbf{S}_{AC}$, $\mathbf{S}_{MC}$ and $\mathbf{S}_{GC}$. (b) Spectral signatures: from top to bottom; $\mathbf{S}_{RS}$, $\mathbf{S}_{GS}$, and $\mathbf{S}_{BS}$.
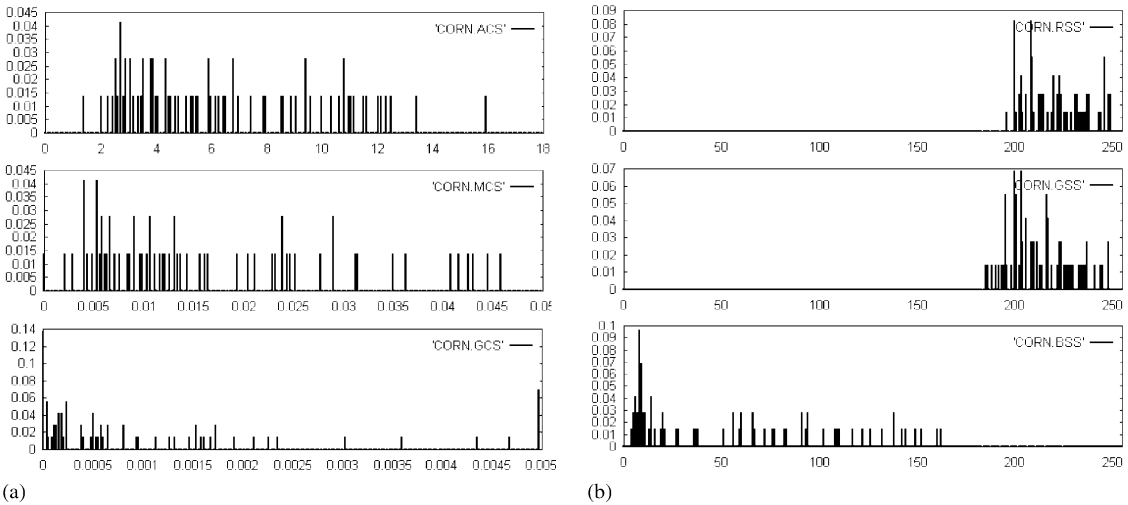


Fig. 8. *Corn* signatures: (a) Curvature signatures: from top to bottom; $\mathbf{S}_{AC}$, $\mathbf{S}_{MC}$ and $\mathbf{S}_{GC}$. (b) Spectral signatures: from top to bottom; $\mathbf{S}_{RS}$, $\mathbf{S}_{GS}$, and $\mathbf{S}_{BS}$.

Table 2
Surface match errors for *observed-surface$_1$* of *Fruit Scene* #1

| Error | Apple | Banana | Corn | Grape | Lemon | Pepper | Strawberry | Tomato |
|---|---|---|---|---|---|---|---|---|
| Curvature | 0.07 | 0.13 | **0.03** | 0.03 | 0.08 | 0.06 | 0.07 | 0.09 |
| Spectral | 0.23 | 0.14 | **0.04** | 0.34 | 0.12 | 0.42 | 0.25 | 0.18 |
| Surface | 0.15 | 0.14 | **0.03** | 0.19 | 0.10 | 0.24 | 0.16 | 0.14 |
| RPI/Θ | 0.44 | 0.38 | **0.00** | 0.57 | 0.25 | 0.74 | 0.47 | 0.39 |

Table 3
Surface matching performance for *Fruit Scene* #1

| Observed-surface | Model–Match | Surface-Error | RPI |
|---|---|---|---|
| 1 | *Banana* | 0.08 | 0.05 |
| 2 | *Corn* | 0.03 | 0.00 |
| 3 | *Pepper* | 0.05 | 0.00 |
| 4 | *Strawberry* | 0.07 | 0.04 |
| 5 | *Tomato* | 0.05 | 0.00 |

A summary of the matching and identification results for the *Geometry* set are shown in Table 5. Matching and identification results for this set were as follows:

- All scenes – except scene #2 – had a small average scene surface matching error, where values of *SEPI* fluctuated between 0.08 and 0.11, producing an overall average *SEPI* value of 0.11. *Scene* #2 is the same as scene #5 but with white noise added ($\mu = 0$, $\sigma = 7.8$). As a result of noise, *SEPI* increased 13%.
- All observed-surfaces appearing in all the scenes – except *scene* #2 – matched to their correct model-surfaces with the lowest surface errors, producing an overall *SMPI* value of 100%. For *Scene* #2: 3 of 5 observed-surfaces matched to their correct model-surfaces with the lowest surface errors, as a result *SMPI* decreased from 100% to 86%. However, the system was able to correctly match the models by exploiting the surface adjacency graph of the scene and comparing it to the models.
- All objects appearing in all the scenes were correctly hypothesized to be the correct model. The overall *MHPI* value of all the scenes was 100%.

These results indicate that the matching and identification system worked flawlessly for these scenes as all observed-surfaces and objects were correctly identified.

### 6.3. Summary of results for the toy scenes

The second test set, the *Toy* set, consisted of a set of seven real scenes obtained through a CCD camera. Scenes #1 and #2 were shown in Fig. 3. The objects employed in this set consisted of seven toy objects: two cylinders, two elliptic cylinders, an elliptic cylinder half sectioned, a right hexagonal prism and a right triangular prism. Three different types of surfaces were present in these scenes: flat surfaces, cylindrical surfaces and elliptic surfaces. Surface colors of the objects are blue, cyan, white and red. The test images contained 26 scene objects which once analyzed produced 39 surfaces. The identification and matching results are shown in Table 6. An examination of the table results in the following observations:

- Values of *SEPI* were in the range 0.09–0.25. This produced an overall *SEPI* value of 0.16 for the whole set. Scenes that had large *SEPI* values (*scenes* #5 and #7), had these errors due to the fact that the observed-surfaces in these scenes were mostly occluded or in shadow.
- Values of *SMPI* were in the range 83–100%. These values produced an overall *SMPI* value for the whole set of 93.4%.
- Even with large scene errors as measured by *SEPI*, the overall model identification performance for the whole set, as measured by *MHPI*, was an impressive 99.8%. In fact, all scenes – except one scene (*Real Scene* #6) – had a *MHPI* value of 100%. In other words, out of the 26 objects appearing in the scenes, objects were correctly hypothesized 25 out of 26 times. For the single case when the correct model was not correctly hypothesized, the correct model produced the second best model-match with a small hypothesis error of 0.06. Such excellent identification was possible, even though matching performance was not as excellent, due to the integration of surface adjacency information into the identification process.

Table 4
*Geometry* set object attributes

| Object name | Type | Radius | Length | Diffuse color[a] | Specular color |
|---|---|---|---|---|---|
| *Cyan cylinder* | Cylinder | 2 | 5 | 0.22:0.45:0.64 | 0.35:0.35:0.35 |
| *Blue Cylinder* | Cylinder | 1 | 6 | 0.1:0.2:0.89 | 0.35:0.35:0.35 |
| *Yellow cylinder* | Cylinder | 1.5 | 4 | 0.89:0.84:0.34 | 0.35:0.35:0.35 |
| *Red sphere* | Sphere | 2 | – | 0.95:0:0 | 0.4:0.4:0.4 |
| *Green sphere* | Sphere | 1.25 | – | 0.6:0.98:0.6 | 0.4:0.4:0.4 |
| *Violet block* | Block | – | 2:3:4 | 0.6:0.13:0.94 | 0.4:0.4:0.4 |
| *Yellow cone* | Cone | 2 | 4 | 0.89:0.6:0.2 | 0.35:0.35:0.35 |

[a] Color values listed are normalized RGB values

Table 5
Summary of results for the *Geometric* Set

| Scene | Observed surfaces | SEPI | SMPI | Scene objects | *MHPI* |
|---|---|---|---|---|---|
| 1 | 3 | 0.09 | 100% | 3 | 100% |
| 2 | 3 | 0.21 | 86% | 3 | 100% |
| 3 | 3 | 0.09 | 100% | 2 | 100% |
| 4 | 5 | 0.11 | 100% | 3 | 100% |
| 5 | 4 | 0.11 | 100% | 3 | 100% |
| 6 | 5 | 0.08 | 100% | 5 | 100% |
| 7 | 3 | 0.09 | 100% | 2 | 100% |
| Overall | 26 | 0.11 | 98% | 21 | 100% |

Table 6
Summary of results for the *Toy* Set

| Scene | Observed surfaces | SEPI | SMPI | Scene objects | *MHPI* |
|---|---|---|---|---|---|
| 1 | 7 | 0.14 | 100% | 5 | 100% |
| 2 | 5 | 0.13 | 97% | 4 | 100% |
| 3 | 8 | 0.09 | 100% | 3 | 100% |
| 4 | 4 | 0.13 | 97% | 4 | 100% |
| 5 | 7 | 0.23 | 83% | 4 | 100% |
| 6 | 4 | 0.15 | 95% | 3 | 98.5% |
| 7 | 4 | 0.25 | 84% | 3 | 100% |
| Overall | 39 | 0.16 | 93.4% | 26 | 99.8% |

### 6.4. Summary of results for the Fruit Scenes

The third test set consisted of a set of seven real scenes, consisting of 30 observed-surfaces and 30 scene objects, obtained as before through a CCD camera. Scenes #1 and #2 were shown in Fig. 4. The objects of this set consisted of eight plastic fruit models: *apple*, *banana*, *corn*, *grape*, *lemon*, *pepper*, *strawberry* and *tomato*. These objects differed from the objects of the other sets by having complex shapes and surfaces. i.e. surfaces could not be simply described as spherical, cylindrical or any other type of uniform surface. It was assumed that each object consisted of one surface. The object colors of this set are red, yellow, purple, green. A summary of the matching and identification results for the *Fruit* scene set are shown in Table 7. From the results we observe that:

- Although objects of this set had complex shapes, the surface matching errors were lower than those

Table 7
Summary of results for *Fruit Scenes*

| Scene | Observed Surfaces | SEPI | SMPI | Scene objects | *MHPI* |
|---|---|---|---|---|---|
| 1 | 4 | 0.05 | 99.5% | 4 | 99.5% |
| 2 | 5 | 0.10 | 100% | 5 | 100% |
| 3 | 5 | 0.06 | 98% | 5 | 98% |
| 4 | 4 | 0.05 | 98% | 4 | 98% |
| 5 | 3 | 0.05 | 99% | 3 | 99% |
| 6 | 4 | 0.09 | 96% | 4 | 96% |
| 7 | 5 | 0.06 | 98% | 5 | 98% |
| Overall | 30 | 0.07 | 98.4% | 30 | 98.4% |

obtained for the *Toy* model set (overall *SEPI* for the whole set was 0.07), this is because surfaces in the set are of the same general type – curved surfaces, whereas the other set had two types of surfaces – curved and flat surfaces. Thus, the contribution of curvature errors to the overall surface error for this set was less than those for other sets.

- Values of *SMPI* for this set were in the range 96–100% which were higher than *SMPI* values for the *Toy* model set. The overall *SMPI* value for this set was 98.4%.
- Since the objects of this set consisted of single surfaces, values of *MHPI* for this set are equivalent to *SMPI* values. Hence, the *MHPI* values for this scene were slightly lower than the those obtained for the other set. The overall *MHPI* for the whole set was nevertheless an excellent 98.4%. The reason that this set had a slightly lower *MHPI* value than the other set, even though the *SMPI* value for this set was higher than the other set, was because model identification was based only on the results of surface matching, and surface adjacency information was not applicable.

## 7. Effect of weight factor variation on matching performance

In this section we study the effect of weight variation on matching and identification performance. We will study the effect of variation on performance by examining the following performance indices:

1. the surface match error ($\bar{E}$),
2. the scene error performance index (*SEPI*),

Table 8
Error weights employed for surface matching for the *Fruit Scenes*

| $w_\mu$ | $w_\sigma$ | $w_\lambda$ | $w_c$ | $w_{AC}$ | $w_{GC}$ | $w_{MC}$ | $w_{RS}$ | $w_{GS}$ | $w_{BS}$ | $w_{rS}$ | $w_C$ | $w_S$ |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| 0.45 | 0.22 | 0.22 | 0.11 | 0.5 | 0.15 | 0.35 | 0.22 | 0.22 | 0.22 | 0.34 | 0.5 | 0.5 |

3. the relative performance index (*RPI*),
4. the average relative performance index (*ARPI*).

We employ *Fruit Scene* #1 for our analysis. As earlier presented, *Fruit Scene* #1 consisted of the following objects: *banana*, *corn*, *pepper*, *strawberry*, and *tomato*. We will analyze the effect of weight variation on each of these objects and the overall effect of the variation on scene performance.

In order to adequately study the effect of weight variation for each of the 12 error weights employed to calculate $\bar{E}$ (i.e., $w_\mu$, $w_\sigma$, $w_\lambda$, $w_c$, $w_{AC}$, $w_{GC}$, $w_{MC}$, $w_{RS}$, $w_{GS}$, $w_{BS}$, $w_{rS}$, $w_C$ and $w_S$), a 12-dimensional system would be required to analyze the performance in detail. Since such an analysis is impractical and difficult to visualize, we will simplify our analysis by varying the value of each weight factor in their respective error groups, while keeping the weights in the other groups constant. The weight variation groups are:

- Variation of the signature match weights: $w_\mu$, $w_\sigma$, $w_\lambda$ and $w_c$.
- Variation of the curvature signature weights: $w_{AC}$, $w_{GC}$ and $w_{MC}$.
- Variation of the spectral signature weights: $w_{RS}$, $w_{GS}$, $w_{BS}$ and $w_{rS}$.
- Variation of the surface error weights: $w_C$ and $w_S$.

Since each group of weights are constrained to sum to unity, we will vary each weight factor from zero to unity while keeping the remaining weights in the group equal. Let $w^*$ represent the varied weight and let $w'$ represent any of the remaining group weights, then we have

$$\sum_i^n w_i = 1, \tag{35}$$

$$w' = \frac{1}{n-1}(1 - w^*), \tag{36}$$

where $n$ is the number of weight factors in the group. As an example, in studying the effect of varying the signature match weights which consist of the four weight factors: $w_\mu$, $w_\sigma$, $w_\lambda$ and $w_c$, if $w_\mu$ is to be studied, then $w_\mu$ is varied from zero to unity and

$$w_\sigma = w_\lambda = w_c = \tfrac{1}{3}(1 - w_\mu), \tag{37}$$

$$w_\mu + w_\sigma + w_\lambda + w_c = 1. \tag{38}$$

The default values taken for the weights are shown in Table 8 (these are the same weight values employed in the analysis of the *Fruit* scenes discussed in the previous section).

As an example, Fig. 9 shows the effect of varying $w_\mu$ on the matching and identification performance for *Fruit Scene* #1. We see that all values of $\bar{E}$ for all objects in the scene decrease as $w_\mu$ increases. As a result, the *SEPI* value for the scene decreases as $w_\mu$ increases. The *RPI* value for *corn* and *pepper* are constant at zero for all values of $w_\mu$, which implies that these two objects are correctly identified regardless of the values of $w_\mu$. This is explained by the fact that the *corn* curvature signatures have unique profiles that easily distinguish it from the curvature signatures of other objects. The *RPI* value for *pepper* remains constant because its spectral signatures easily distinguish it from others (it is the only object with a green color). The optimum *RPI* value is at $w_\mu \cong 0.8$, where *ARPI* attains a minimum and *RPI* curves of all objects – except *strawberry* – also attains a minimum.

Table 9 shows the optimum weight values found for *Fruit Scene* #1. With the application of these weights, the results now show that all observed-surfaces match correctly to their model surfaces ($SMPI = MHPI = 100\%$), whereas before only three of the five observed-surfaces matched correctly to their model-surfaces. The overall scene match error decreased 16% from a *SEPI* value of 0.06 to a *SEPI* value of 0.05. It should be emphasized that
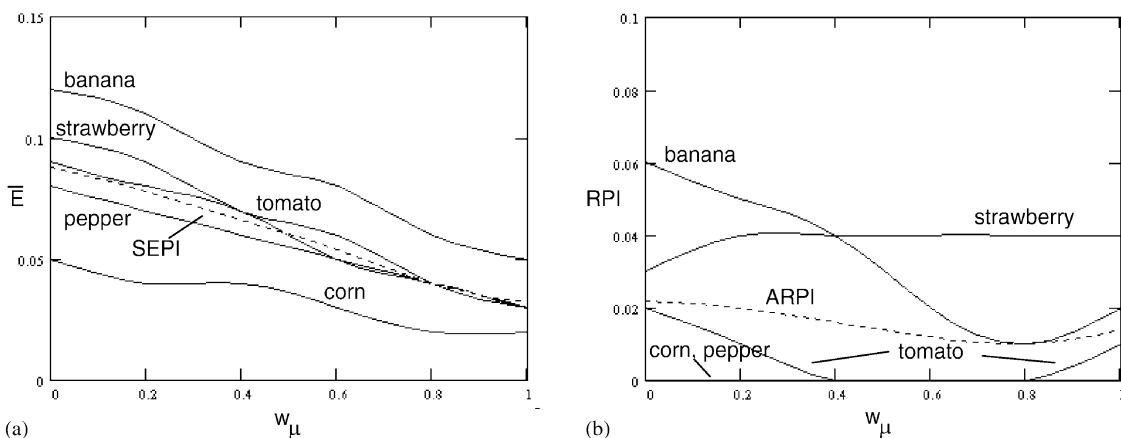


Fig. 9. Effect of varying $w_\mu$ on *Fruit Scene* #1. (a) $\bar{E}$, (b) *RPI*.

Table 9
Optimum error weights for surface matching for *Fruit Scene* #1

| $w_\mu$ | $w_\sigma$ | $w_\lambda$ | $w_c$ | $w_{AC}$ | $w_{GC}$ | $w_{MC}$ | $w_{RS}$ | $w_{GS}$ | $w_{BS}$ | $w_{rS}$ | $w_C$ | $w_S$ |
|------|------|------|------|------|------|------|------|------|------|------|------|------|
| 0.42 | 0.42 | 0.11 | 0.05 | 0.3 | 0.2 | 0.5 | 0.25 | 0.12 | 0.25 | 0.38 | 0.3 | 0.7 |

the optimum weight values obtained are valid only for these objects. To arrive at optimum weight values applicable to all objects in the set, an image containing all objects, or more than one image containing all objects, should be analyzed. As the knowledge base becomes large and new objects are added to the knowledge base, the weight factors would not have to be recalculated as their values would only change slightly.

## 8. Conclusion and future work

We have described a system for 3D object identification that integrates object curvature with object color. The system takes two color images of a scene taken under different lighting conditions and produces a hypothesis about objects appearing in the scene. The system is based on analyzing object surfaces from which surface signatures are constructed. Surface signatures are surface attributes that characterize a surface. We employ two types of signatures; curvature signatures and spectral signatures. Surface signatures extracted from objects appearing in the scene are matched to model surface signatures – obtained through training – to arrive at a hypothesis about the objects.

Tests were conducted on three different sets of 3D objects – one simulated and two real – consisting of 77 objects appearing in 21 scenes. Once analyzed these scenes produced 95 surfaces of different curvatures: flat, cylindrical, spherical, conic, elliptical and other complex types (e.g. banana surface). Different surface colors also appeared in the scenes: cyan, blue, green, yellow, red, violet, white, and purple.

Tests conducted on simulated images produced perfect surface matching and correct object hypothesis. In the presence of noise, simulated images did not produce perfect surface matching, but nevertheless produced correct object hypothesis. Using real images proved to be a much more difficult task than using simulated images. Noise inherent in images sometimes led to inaccurate construction of surface normals. With such inaccuracies, surface curvature signatures constructed though still representing the surface curvature to an acceptable degree in most cases had more inaccuracies and errors associated with them. The surface spectral signatures represented the object colors more accurately, though not as accurate as those constructed for the simulated images. The use of both surface curvature signatures and surface spectral

signatures, with more weight given to the spectral signatures, provided sufficient matching constraint to produce better matching and identification performance. Current work is directed at refining the system by:

1. Integrating the use of the *extended Gaussian image* (*EGI*) in the system. This can resolve some of the ambiguities found when matching objects to models.
2. Integrating procedures to analyze 2D shapes of objects into the system can increase the performance of the system. The shape of an observed-surface can be another item by which surfaces are compared. Our previous work in this area [21] has shown that such measures are robust and can be easily implemented. Furthermore, these measures are usually efficient and do not slowdown the system's performance.
3. Integration of texture analysis into the system.

## References

[1] M. Weathersby, Multi-spectral Image Processing and Enhancement, Orlando, FL, 6–8 April, SPIE Proc., vol. 933, 1988.
[2] M. Trivedi, C. Chen, D. Cress, Object detection by stepwise analysis of spectral, spatial, and topographic features, Computer Vision, Graphics Image Process, 51 (1990) 235–255.
[3] J. Kagel, J. Reeder, Prototype Neural Network Processor for Multispectral image fusion. Int. Joint Conf. on Neural Networks, Seattle, WA, vol. 1, July 1991, pp. 245–250.
[4] G. Healey, S. Shafer, L. Wolff (Eds.), Physics-Based Vision: Color, Jones and Bartlett Publishers, Boston, 1992.
[5] P. Besl, J. Ramesh, Three dimensional object recognition, Computing surveys 17(1) (1985) 75–145.
[6] R. Haralick, L. Watson, T. Laffey, The topographic primal sketch, Int. J. Robot Res. 2(1) (1983) 50–72.
[7] T. Newman, P. Flynn, A. Jain, Model-based classification of quadric surfaces, CVGIP: Image Understanding 58(2) (1983) 235–249.
[8] R. Woodham, Photometric stereo: a reflectance map technique for determining surface orientation from a single view, Proc. SPIE, vol. 155, 1978, pp. 136–143.
[9] R. Woodham, Photometric method for determining surface orientation from multiple images, Opt. Engng 19(1) (1980) 139–144.
[10] P. Horn, Understanding Image Intensities, Artificial Intelligence 8 (1977) 201–231.
[11] K. Ikeuchi, Determining surface orientations of specular surfaces by using the photometric stereo method, IEEE Trans. Pattern Anal. Machine Intell. 3(6) (1981) 661–669.

[12] K. Ikeuchi, Determining a depth map using a dual photometric stereo, IEEE Trans. Pattern Anal. Machine Intell. 3(6) (1987) 661–669.

[13] R. Woodham, Surface curvature from photometric stereo. Technical Rep. 90-29, Lab. of Computational Vision, Univ. of BC, 1990.

[14] P. Christensen, L.G. Shapiro, Determining the shape of multi-colored dichromatic surfaces using color photometric stereo, Conference on Computer Vision and Pattern Recognition, New York, NY, Jun, 1993.

[15] P. Christensen, L.G. Shapiro, Three dimensional shape from colored photometric stereo, Int. J. Comput. Vision 13(2) (1994) 213–227.

[16] K. Higuchi, H. Delingette, K. Ikeuchi, Merging multiple views using a spherical representation, Second CAD–Based Vision Workshop, Champion, PA, 8–11 February, 1994, pp. 17–26.

[17] L. Grewe, A. Kak, Interactive learning of multiple attribute Hash table for fast 3D object recognition, Second CAD-Based Vision Workshop, Champion, PA, 8–11 February, 1994, pp. 17–26.

[18] R. Gnanadesikan, Methods of Statistical Data Analysis of Multivariate Observations, Wiley, New York, 1977.

[19] A. Mustafa, L.G. Shapiro, M. Ganter, 3D object identification from color intensity images, 13th Int. Conf. on Pattern Recognition, Vienna, Austria, 25–30 August, 1996.

[20] A. Mustafa, Object identification with surface signatures using color photometric stereo. Ph.D. Dissertation, Department of Mechanical Engineering, University of Washington, March, 1995.

[21] A. Mustafa, Determining symbolic differences between two-dimensional models using structural matching. Master of Science Thesis, Department of Mechanical Engineering, University of Washington, August, 1988.

**About the Author**—ADNAN MUSTAFA received his B.S. in Mechanical Engineering from the California State University in 1983. He worked in the oil industry a few years before persuing his graduate studies at the University of Washington (Seattle). He received his M.S. and his Ph.D degrees in Mechanical Engineering in 1988 and 1995, respectively. He is currently teaching at Kuwait university. His research interests are in the field of Computer Vision.

**About the Author**—LINDA G. SHAPIRO was born in Chicago, Illnois, 1949. She received the B.S. degree in mathematics from the University of Illinois, Urbana, in 1970, and the M.S. and Ph.D. degrees in computer science from the University of Iowa, Iowa City, in 1972 and 1974, respectively.

She was an Assistant Professor of Computer Science at Kansas State University, Manhattan, from 1974 to 1978 and was an Assistant Professor of Computer Science from 1979 to 1981 and Associate Professor of Computer Science from 1981 to 1984 at Virginia Polytechnic Institute and State University, Blacksburg. She was Director of Intelligent Systems at Machine Vision International in Ann Arbor from 1984 to 1986. She is currently Professor of Computer Science and Engineering and of Electrical Engineering at the University of Washington. Her research interests include computer vision, image database systems, artificial intelligence, pattern recognition, and robotics. She has co-authored two textbooks, one on data structures and one on computer and robot vision.

Dr. SHAPIRO is a Fellow of the IEEE and a member of the Pattern Recognition Society and the American Association for Artificial Intelligence. She is a past Chair of the IEEE Computer Society Technical Committee on Pattern Analysis and Machine Intelligence and is currently an editorial board member of Pattern Recognition and of Computer Vision and Image Understanding. She has served as Editor of CVGIP: Image Understanding, co-Program Chairman of the IEEE Conference on Computer Vision and Pattern Recognition in 1994, General Chairman of the IEEE Workshop on Directions in Automated CAD-Based Vision in 1991, and General Chairman of the IEEE Conference on Computer Vision and Pattern Recognition in 1986. She has also served on the program committees of numerous vision and AI workshops and conferences.

**About the Author**—DR MARK GANTER is involved in applied computational geometry, solid modeling, computer graphics, design automation, rapid prototyping and rapid manufacturing. His main research activities focus on application of computational techniques to engineering design problems involving geometry. He received his Ph.D. in Mechanical Engineering from University of Wisconsin in 1985, and has been at the University of Washington since 1986.