Properties of color images of surfaces under multiple illuminants

A. P. Petrov

Atomic Energy Institute, Kurchatov Square, Moscow 123182, Russia

L. L. Kontsevich

Smith-Kettlewell Eye Research Institute, 2232 Webster Street, San Francisco, California 94115

Received June 2, 1993; revised manuscript received June 9, 1994; accepted June 10, 1994

We present a theoretical framework for object-based segmentation of color images that extracts the uniformly colored regions on an arbitrary surface all the points of which are exposed to the same set of light sources. This segmentation is a necessary stage before shape-from-shading processing. We introduce a consistent definition of rank and provide a complete classification of interactions among shape, surface color, and illumination. The proposed theory is an extension of known approaches.

1. INTRODUCTION

The shape-from-shading reconstruction, like any other inverse vision task, is based on assumptions that constrain possible solutions. For the shape-from-shading task the specific assumption, presumably exploited by the visual system, is that the variations of brightness within an image are caused by gradients in surface orientation rather than by variations of surface color or incident light.[1,2] This assumption is generally valid for local parts of a scene: an adequate segmentation process is required for locating such parts. Each region extracted in this process corresponds to an area on the surface that is uniformly colored and illuminated by a fixed set of light sources. All algorithms known to the authors that solve the shape-from-shading task apply only to regions of this type.[3–6]

The segmentation process is a complicated problem for monochrome images for which information about illuminants or surface color and shape is not available. For multispectral images, however, the segmentation can be elegantly performed without additional data, as recognized by Nikolaev.[7] He developed a theory that classified image fields corresponding to uniformly colored areas of matte surfaces illuminated by a fixed set of light sources. He also proposed several algorithms for segmentation. Unfortunately, Nikolaev's papers[7–9] were not available in English until recently. Most of his results were described by Brill,[10] who surveyed research of this kind of segmentation and demonstrated that Nikolaev's theory encompasses all others.[11–14] In addition, Brill improved, extended, and clarified Nikolaev's theory on some points.

Central to Nikolaev's theory is the concept of image-field rank. According to Brill's interpretation, the rank is equal to the number of light sources illuminating the corresponding surface area. (Nikolaev proposed a slightly different definition of rank, but here we shall follow Brill's definition.) Nikolaev proposed a method for the unique description of rank-1 and rank-2 regions. However, because of the trichromatic nature of the visual system, it is impossible to extend this method to regions of higher ranks. This shortcoming was caused by an insufficient definition of rank.

The remainder of this paper is concerned with providing a consistent definition of rank and unique descriptions of regions of all possible ranks. Furthermore, answers to mathematical questions that arise in relation to segmentation are provided. The definition is now possible because of the new language for the description of interaction among illumination, surface, and viewer proposed by Petrov.[15] Let us start with this description.

2. BASIC DEFINITIONS AND RELATIONS

First consider a simplified task. Suppose that a small patch of matte surface is situated in space and is illuminated by a single light source with a very small aperture (see Fig. 1). The orientation of the surface can be defined by unit normal $\hat{n}$, and its reflectance properties by spectral reflection function $R(\lambda)$. The incident light can be characterized by its direction $\hat{p}$ (unit vector) and spectral power density function $S(\lambda)$.

Suppose that this patch of matte surface is then imaged by a trichromatic eye with sensitivities $\nu_i(\lambda)$, where $i = 1, 2, 3$ might represent long-wavelength-sensitive, medium-wavelength-sensitive, and short-wavelength-sensitive cones in the eye. Response of ith sensor for the patch of surface is given by

$$r_i = \hat{n}^T \hat{p} \int_0^\infty S(\lambda)R(\lambda)\nu_i(\lambda)\mathrm{d}\lambda \qquad (1)$$

(here superscript T denotes matrix transposition). Note that the effect of the illumination geometry on the response is determined by the scalar product $\hat{n}^T\hat{p}$ and does not depend on the position of the viewer. By definition this is an intrinsic property of purely matte surfaces.

Equation (1) can be represented in the more convenient form

$$\mathbf{r} = \mathbf{M}\hat{n}, \qquad (2)$$

0740-3232/94/102745-05$06.00

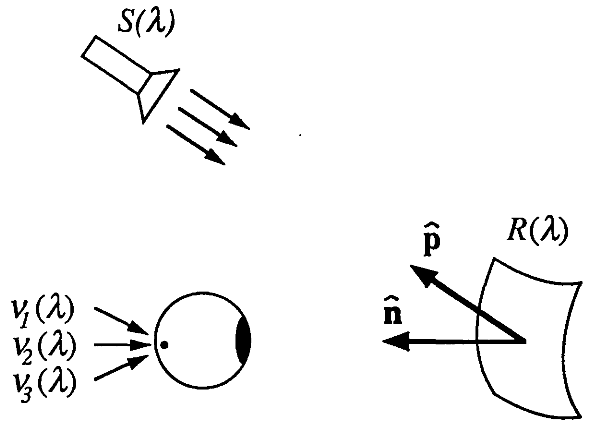


Fig. 1. Small patch of matte surface illuminated by a point light source with spectral power density function $S(\lambda)$; the direction at the light source is $\hat{p}$ (unit vector). The orientation of the surface is defined by unit normal $\hat{n}$, and its reflectance properties by spectral reflection function $R(\lambda)$. For every point of the receptive surface, three different receptors with spectral sensitivities $\nu_i(\lambda)$ are present.

where **r** is a three-dimensional response vector, $\hat{n}$ a normal unit vector, and **M** a 3×3 matrix with elements

$$m_{i,j} = \int_0^\infty S(\lambda)R(\lambda)\nu_i(\lambda)p_j \mathrm{d}\lambda \qquad (3)$$

(here p_j means jth coordinate of vector $\hat{p}$). Note that Eq. (2) is valid for all surface orientations when the surface is illuminated from the outside. This condition can be written explicitly as

$$\hat{n}^T \hat{p} \geq 0\,. \qquad (4)$$

Equation (2) states that the normals and the responses are linked by linear mapping when relation (4) holds.

In the case of complex illumination, when the surface is illuminated by a set of light sources its response equals the sum of responses to each individual light source (assuming that the responses of the color-sensitive detectors are linear and without saturation). Therefore Eq. (2) is also valid for complex illumination, but in this case matrix **M** is the sum of matrices corresponding to each light source. In the case in which some of the light sources have large apertures, integration of spatial angles is necessary.

Matrix **M** is invariant under all orientations of the normal when the set of light sources illuminating the surface patch does not change. These normal directions form a convex cone of admissible directions. This cone is defined by conditions similar to relation (4) that must be valid for all sources from the illuminant set corresponding to **M** and invalid for all other illuminants.

Now let us consider an extended area of smooth surface. Suppose that the distance between light sources and the surface is much greater than size of the surface, so that the vectors $\hat{p}$ pointing to any light source are equal for all points on the surface. Then any two points on the surface having the same color and illuminated by the same set of light sources have the same matrix **M**. Thus we can reformulate the segmentation problem described in Section 1 as a task of finding the regions consisting of all the points that have responses corresponding to the same matrix **M**.

If $L(\mathbf{r}, G)$ denotes the linear span for all response vectors **r** from region G, then $\dim[L(\mathbf{r}, G)]$ is the *rank of region G*.[16]

For classification purposes let us call $\dim L(\hat{n}, G)$ the *rank of shape for* region G and rank **M** the *rank of color*. For the most common case of rank-3 shape, the rank of shape is equal to the rank of color, and therefore

$$\dim L(\mathbf{r}, G) = \operatorname{rank} \mathbf{M}\,. \qquad (5)$$

The remaining rank-2 and rank-1 shapes are cylindrical surfaces and planes, respectively. These cases will be considered separately.

One difference between the definition of color rank proposed here and Brill's definition for image-field rank is that our definition does not allow for color rank of greater than 3 in a trichromatic visual system. The basis for this definition is that the distribution of light corresponding to any set of light sources can be emulated by three chromatic light sources with small apertures [at least formally, in cases in which $S(\lambda)$ can be negative and restrictions such as relation (4) are not applied]. This fact was missed by Nikolaev and Brill. Another difference is that the proposed definition decomposes region rank into color rank and shape rank. This decomposition provides a powerful tool for analysis of various cases in general terms.

In summary, Eq. (2) describes responses in arbitrary regions. This relation provides two possible methods for image analysis. The first considers the same surface in different illumination conditions. As a result, one can obtain descriptions of both illumination and shape that are invariant to color. This possibility was explored by Petrov.[17] The other method of analysis is to reconstruct the metric introduced by mapping **M** for a given area of the image. This approach was first proposed by Kontsevich[17] and is developed further in this paper.

The concept of rank is useful for the classification of regions, but because there are only four possible rank values (0, 1, 2, 3), rank does not specify regions uniquely. Matrix **M**, however, could be used as a unique descriptor for a region. In some cases it can be determined after completion of the shape-from-shading process.[6] The intrinsic contradiction of this method is that one must perform image segmentation beforehand, because the analysis of shape from shading depends on knowledge of image-segmentation results.

We consider simpler descriptors for regions that can be obtained from the image directly.

3. DESCRIPTORS OF IMAGE FIELDS

For further analysis it is important to remember that surface normals are of unit length. If all surface normals were translated to the origin, their end points would thus lie on the unit sphere. Each surface point has a corresponding point on this sphere. This is a well-known Gaussian mapping of the surface onto the unit (Gaussian) sphere. Obviously, the mapping of surface normals onto trichromatic response space can be considered the superposition of the mapping of surface normals onto a

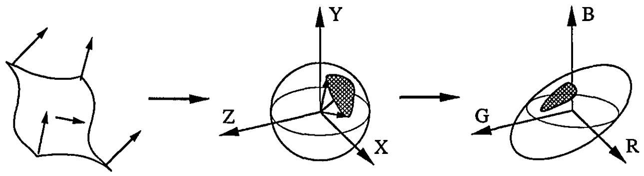


Fig. 2. Mapping of surface normals (left) onto trichromatic response space (right) can be considered the superposition of the mapping of surface normals onto a Gaussian sphere (middle) and the mapping of the Gaussian sphere onto response space (right). The images of the surface normals on the Gaussian sphere and its image (ellipsoid) in the trichromatic response space are depicted by the shaded regions.

Gaussian sphere and the mapping of the Gaussian sphere onto the response space (Fig. 2). Without loss of generality we can omit the first mapping for rank-3 shapes and consider the whole Gaussian sphere an original domain for the second mapping.

The results of the following analysis are collected in Table 1. In each cell, region rank and the locations of its corresponding set of response vectors are depicted graphically (we draw only the response vectors having positive coordinates).[18]

Rank-1 Region for Rank-3 Shape

This is the case in which matrix **M** corresponding to the region has rank 1. The image of the Gaussian sphere is a one-dimensional subspace in the response space, and all response vectors of the region lie on a line passing over the origin. This line can be determined by the unit vector collinear with it.

Thus the main property of the rank-1 region is colinearity of its response vectors. Such a region can be described by a single nonzero vector in response space.

Rank-2 Region for Rank-3 Shape

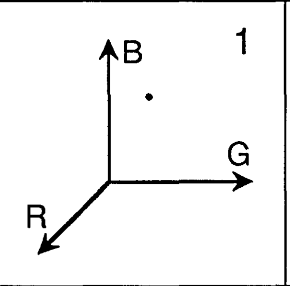
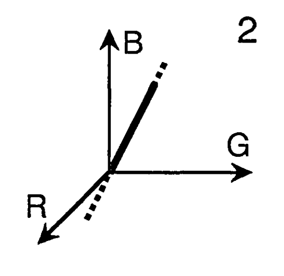
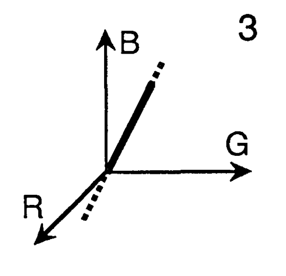
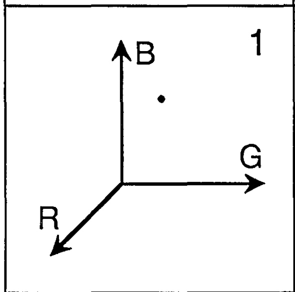
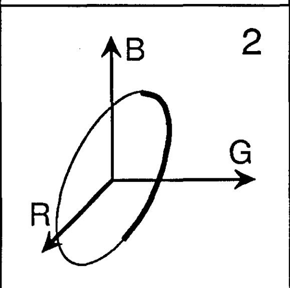
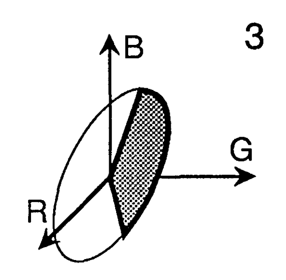
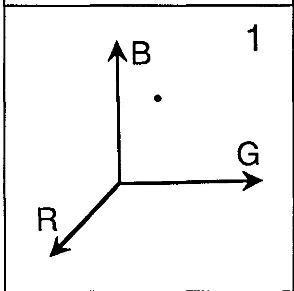
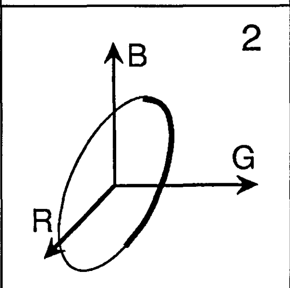
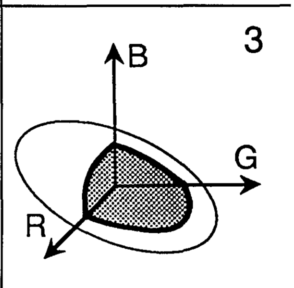
In this case matrix **M** has rank 2, which means that the image of the Gaussian sphere lies in an elliptical area on a two-dimensional subspace of response space. All response vectors of the field lie on a plane passing through the origin. Moreover, all response vectors lie in some bounded area on this plane, which is an image of the Gaussian sphere. The plane can be determined by two linearly independent vectors from this plane. If the response space possesses some metric, the plane can be determined by a vector that is orthogonal to the plane.

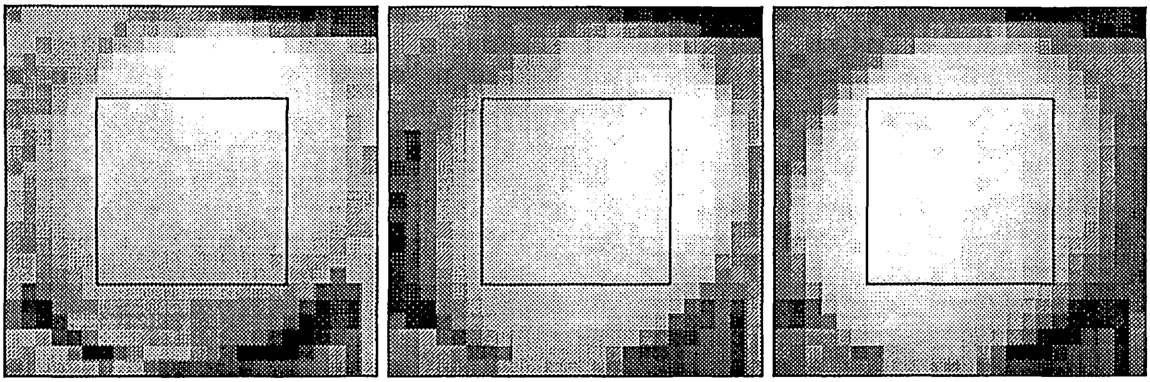
Thus the main property of the rank-2 region is the coplanarity of its response vectors. Such a region can be described by two independent response vectors in the nonmetric case or by a single vector in the metric one.

Rank-3 Region for Rank-3 Shape

In this case the matrix of the region has rank 3. Consequently, the image of the Gaussian sphere is an ellipsoid

Table 1. Rank-Guided Classification of the Image Field Loci in Relation to Shape and Color

	Rank of Shape		
Rank of Color	1	2	3
1			
2			
3			



Red

91	91	94	94	95	98	98	98	99	99	98	97
89	90	92	93	93	96	96	98	99	100	97	96
88	90	91	91	92	94	95	95	96	96	95	92
88	91	91	92	93	91	94	94	93	93	93	93
90	90	90	90	91	93	92	91	92	92	91	90
88	88	87	89	89	90	89	90	90	91	90	87
86	86	86	87	86	89	89	88	88	88	89	87
81	85	85	83	86	87	86	85	87	87	86	83
81	83	84	84	83	83	84	82	85	83	82	81
79	78	83	82	82	81	81	80	77	81	77	77
75	78	78	79	80	79	78	76	75	76	75	70
71	75	74	75	74	76	74	74	75	71	68	68

Green

77	80	83	88	90	92	95	95	96	98	98	98
76	82	86	88	93	92	97	97	97	100	100	99
80	84	88	90	93	95	96	98	97	101	101	100
81	85	89	90	92	92	96	97	99	101	101	100
81	87	89	90	93	93	96	96	99	98	98	97
83	86	88	89	94	93	95	95	99	99	99	96
85	86	89	90	92	94	96	95	96	97	97	97
84	85	89	89	91	92	95	95	96	98	96	95
83	85	87	89	91	92	94	95	94	94	96	95
81	84	87	88	91	92	92	91	92	94	94	93
80	83	87	88	89	89	91	92	92	92	92	90
79	81	84	87	90	89	90	90	91	91	90	88

Blue

93	95	95	97	97	97	98	96	96	96	94	92
95	95	96	95	98	100	97	97	97	95	94	91
96	97	96	99	97	100	99	96	98	96	94	92
97	97	97	98	96	99	100	96	96	95	94	92
96	96	98	96	99	100	101	96	97	95	91	91
96	97	98	96	97	98	98	95	96	95	92	91
96	95	96	95	96	97	97	96	93	92	91	90
94	96	96	95	97	96	95	93	94	93	90	86
95	94	97	96	95	95	94	95	91	90	88	86
95	96	95	96	95	95	92	92	91	90	88	85
95	95	95	95	94	94	93	92	90	90	87	84
94	95	95	96	93	94	93	91	90	88	82	78

Fig. 3. Top: Egg illuminated by three chromatic light sources (orange, green, and blue) as it is seen through three chromatic channels. The area within the rectangular frame is illuminated by all three lights. The responses from this area were taken for computation of the quadratic form **Q**. Bottom: The numerical values of the responses of the channels.

with its center at the origin of the response space. Parameters of the ellipsoid can be reconstructed from the image field.

This ellipsoid is a unit sphere in the response-space metric[19] induced by the mapping **M**. The scalar product of two response vectors in an induced metric is equal to the scalar product of those origins (normals). It can be demonstrated that the symmetrical matrix **Q** of the induced scalar product is related to **M** according to

$$\mathbf{Q} = (\mathbf{M}^{-1})^T \mathbf{M}^{-1} . \tag{6}$$

The ellipsoid is completely determined by the matrix **Q**.

The matrix **Q** can be easily obtained from the image field. When its normals have unit length, the following relation is valid:

$$\mathbf{r}^T \mathbf{Q} \mathbf{r} = 1 . \tag{7}$$

This relation is a linear equation for elements of matrix **Q**. Symmetrical matrix **Q** is defined by six independent parameters; therefore at least six independent equations are necessary. Matrix **Q** can be used to describe the rank-3 region individually.

Example

A real egg was illuminated by three colored lights (orange, green, and blue). The image of the egg was obtained by a camcorder with 128 levels in each chromatic channel. The recording of the channels is shown in Fig. 3. The area bounded by a rectangle (12 × 12 pixels) was illuminated by all three light sources; it was selected for calculation of the matrix **Q**. The matrices of the responses are provided in Fig. 3 below the corresponding images. The solution of the system of linear equations [Eq. (7)] for all responses in the area is

$$\mathbf{Q} = 10^{-4} \begin{bmatrix} 3.312 & -3.525 & 0.088 \\ -3.525 & 8.852 & -4.888 \\ 0.088 & -4.888 & 5.552 \end{bmatrix} .$$

The system of equations is redundant: there are 144 equations for 6 unknowns. However, the norms of all response vectors calculated according to Eq. (7) are close

Table 2. Norm Values $r^T Q r$ of the Responses Taken in a Set of Points of a Real Color Image

1.00	1.01	0.98	0.99	0.98	0.98	1.01	0.97	0.98	0.99	0.97	0.96
1.07	0.98	0.97	0.94	1.01	1.05	1.01	1.00	1.00	1.01	1.01	0.97
1.02	1.01	0.96	1.03	0.99	1.05	1.03	1.01	1.02	1.05	1.04	1.02
1.04	1.00	0.98	1.00	0.96	1.03	1.06	1.00	1.03	1.06	1.05	1.02
1.02	0.97	1.01	0.96	1.03	1.05	1.08	1.00	1.05	1.01	0.98	0.97
0.99	1.00	1.02	0.97	1.01	1.02	1.04	0.98	1.06	1.04	1.02	0.97
0.98	0.95	0.97	0.95	0.99	1.01	1.03	1.01	0.99	0.99	0.98	0.99
0.94	0.98	0.98	0.97	1.00	0.99	1.01	0.99	1.00	1.03	0.97	0.96
0.97	0.94	1.00	0.98	0.99	1.00	1.00	1.05	0.95	0.96	1.00	0.98
0.98	1.01	0.96	0.99	0.99	1.01	0.97	0.96	1.00	0.98	1.01	0.97
1.00	0.98	0.99	0.99	0.97	0.98	1.00	1.02	1.01	1.00	0.98	0.98
1.00	1.00	1.01	1.04	1.03	1.01	1.03	1.00	0.99	1.03	1.00	0.93

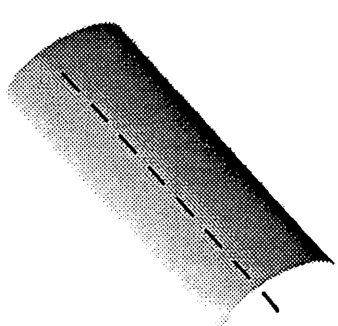


Fig. 4. Cylindrical surface illuminated by an arbitrary set of distant light sources. The luminance along the cylinder axis is constant.

to a value of 1 (see Table 2 for the individual results). The mean square error per point is as small as 2.84%.

Case of Rank-1 and Rank-2 Shapes

So far, we have considered only the case of rank-3 shapes. The reason for this is that segmentation for shapes of lower ranks is easy, though based on different principles. Let us now consider these cases.

We noted above that the rank-1 shapes are exact planes. According to formula (2), the response vectors for all points on a plane are equal for any rank of color. This property can be recognized on projection.

Rank-2 shapes are cylinders. Cylindrical surfaces can be recognized on the projection on the basis of their regularity, unrelated to the rank. A cylinder is invariant to spatial translations along its axis; therefore the projection of any cylinder is invariant to translations along the projection of its axis. Thus the region corresponding to a cylindrical surface can be defined by the direction of the cylinder axis: points of the surface on any line parallel to the cylinder axis have the same orientation, and therefore they have the same response vectors on the projection, as illustrated in Fig. 4.

4. CONCLUSION

We have introduced new concepts of color rank and surface rank and presented a working definition of region rank for color images. On the basis of these definitions we have presented a complete classification for the regions of a homogeneously colored Lambertian surface illuminated by an arbitrary, constant set of light sources, and we provided descriptors that uniquely specify each region. The segmentation based on the proposed descriptors is geometry insensitive but is sensitive to changes of set of illuminants and surface color. This segmentation is a necessary preparatory step to the shape-from-shading processing.

REFERENCES AND NOTES

1. B. K. P. Horn, *Robot Vision* (MIT Press, Cambridge, Mass., 1985, New York, 1985).
2. B. A. Wandell, "The synthesis and analysis of color images," IEEE Trans. Patt. Anal. Mach. Intell. **PAMI-A9,** 2–13 (1987).
3. K. Ikeuchi and B. K. P. Horn, "Numerical shape from shading and occluding boundaries," Artif. Intell. **17,** 141–184 (1981).
4. B. K. P. Horn and M. J. Brooks, "The variational approach to shape from shading," Comput. Vis. Graphics Image Process. **33,** 174–208 (1986).
5. A. Pentland, "Shape information from shading: a theory about human perception," Spatial Vision **4,** 165–182 (1989).
6. L. L. Kontsevich, A. P. Petrov, and I. S. Vergelskaya, "Reconstruction of shape from shading in color images," J. Opt. Soc. Am. A **11,** 1047–1052 (1994).
7. P. P. Nikolaev, "Algorithms for recognition of surface color," in *Modeling of Learning and Behavior*, M. S. Smirnov, ed. (Nauka, Moscow, 1975), pp. 121–151 (in Russian).
8. P. P. Nikolaev, "A model of color constancy for smooth spectral distributions," Biofizika **30,** 112–117 (1985) (in Russian).
9. P. P. Nikolaev, "Monocular color discrimination of nonplanar objects under various illumination conditions," Biofizika **33,** 140–144 (1988) (in Russian).
10. M. H. Brill, "Image segmentation by object color: a unifying framework and connection to color constancy," J. Opt. Soc. Am. A **7,** 2041–2047 (1990).
11. R. A. Schowengerdt, *Techniques for Image Processing and Classification in Remote Sensing* (Academic, New York, 1983), pp. 154–159.
12. J. M. Rubin and W. A. Richards, "Color vision and image intensity: when are changes material?" Biol. Cybern. **45,** 215–226 (1982).
13. S. A. Shafer, "Using color to separate reflection components," Color Res. Appl. **10,** 210–218 (1985).
14. G. J. Klinker, "A physical approach to color image understanding," Ph.D. dissertation, Doc. CMU-CS-88-161 (Carnegie-Mellon University, Pittsburgh, Pa., May 1988).
15. A. P. Petrov, "Structure of the set of perceived colors," Preprint IAE No. 4050/15 (Kurchatov Atomic Energy Institute, Moscow, 1984) (in Russian).
16. A. P. Petrov, "Color and Grassmann–Cayley coordinates of shape," in *Human Vision, Visual Processing, and Digital Display II*, B. E. Rogowitz, M. H. Brill, and J. P. Allebach, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **1453,** 353–364 (1991).
17. L. L. Kontsevich, "Recovering of shape for a uniformly colored and illuminated object from its two-dimensional projection," in *Data Processing in Complex Information Systems, Part 2* (Institute for Problems of Information Transmission, Academy of Sciences of the USSR, Moscow, 1986), pp. 16–19 (in Russian).
18. In some special cases, sets of response vectors can degenerate into simpler cases. For example, the right-circular cone (rank-3 shape) illuminated along the direction of its axis is not distinguishable from a plane on the image.
19. This induced metric relates orientations and responses. It should not be confused with the metric of color space, which is geometry independent.