Color Image Analysis with an Intrinsic Reflection Model

Gudrun J. Klinker, Steven A. Shafer, and Takeo Kanade

Computer Science Department*, Carnegie-Mellon University, Pittsburgh, PA 15213

Abstract
In this paper, we present an approach to color image understanding that can be used to segment and analyze surfaces with color variations due to highlights and shading. We begin with a theory - the Dichromatic Reflection Model - that relates the reflected light from dielectric materials, such as plastic, to fundamental physical reflection processes, and describes the color of the reflected light as a linear combination of the color of the light due to surface reflection (highlights) and body reflection (object color). This dichromatic theory is used in an algorithm that separates a color image into two parts: an image of just the highlights, and the original image with the highlights removed. In the past, we have applied this method to hand-segmented images. This paper shows how to perform automatic segmentation by applying this theory in stages to identify the object and highlight colors. The result is a combination of segmentation and reflection analysis that is better than traditional heuristic segmentation methods and provides important physical information about the surface geometry and material properties at the same time. This line of research can lead to physics-based image understanding methods that are both more reliable and more useful than traditional methods.

1. Introduction
Recently, work in image understanding has started to use intrinsic physical models to analyze intensity or color variations in images [1, 2, 3, 4, 5, 6]. However, such physics-based image analysis generally assumes that the image has previously been segmented into areas with consistent physical interpretations or that the physical influences have been separated. This paper describes an approach to image understanding that starts from the raw color image data and uses an intrinsic reflection model to automatically generate both an image segmentation and an image interpretation, in terms of the physical properties captured in the model. The system is based on the Dichromatic Reflection Model[2], a physical reflection model that accounts for color variations due to shading and highlights. Driven by the model, the algorithm segments color images along material boundaries, ignoring color changes due to highlights and shading. It incrementally identifies local and global properties of the scene, such as object and illumination colors, and uses them in interpreting pixels in an image. It adapts the image interpretation process to local scene characteristics and reacts differently to color and intensity changes at different places in the image. The result is a combination of segmentation and reflection analysis that is better than traditional heuristic segmentation methods that base their analysis on intensity or color differences or on a fixed set of user-defined features, such as intensity, hue and saturation[7, 8, 9]. In addition to generating a segmentation image that ignores shading and highlight changes, the approach generates important physical information about the scene. This line of research is aimed at developing physics-based image segmentation methods that are both more reliable and more useful than traditional segmentation methods.

*This research was supported by the Jet Propulsion Laboratory, California Institute of Technology, under NASA contract 957989, and by the Defense Advanced Research Projects Agency (DOD), ARPA Order No. 4976, monitored by the Air Force Avionics Laboratory (Air Force Wright Aeronautical Laboratories, Aeronautical Systems Division (AFSC), United States Air Force, Wright-Patterson AFB, OHIO 45433-6543) under contract F33615-87-C-1499 (Arpa Order # 4976, Amendment 20). Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the Jet Propulsion Laboratory, the Defense Advanced Research Projects Agency, or the US Government.

2. The Dichromatic Reflection Model
The Dichromatic Reflection Model describes the light, $L(\lambda,i,e,g)$, which is reflected from a point on a dielectric, non-uniform material as a mixture of the light $L_s(\lambda,i,e,g)$ reflected at the material surface and the light $L_b(\lambda,i,e,g)$ reflected from the material body (see Figure 1). The parameters i, e, and g describe the angles of the incident and emitted light and the phase angle; λ is the wavelength parameter. L_s is called the *surface reflection component*. It generally has approximately the same spectral power distribution as the illumination and appears as a highlight or as gloss on the object. L_b is called the *body reflection component*. It provides the characteristic object color and exhibits the properties of object shading.

$$L(\lambda,i,e,g) = L_s(\lambda,i,e,g) + L_b(\lambda,i,e,g) \tag{1}$$

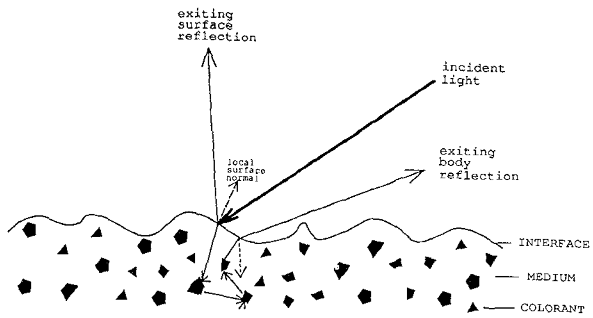

Figure 1: Light reflection of dielectric materials

The model separates the spectral reflection properties of L_s and L_b from their geometric reflection properties, modeling them as products of spectral power distributions, $c_s(\lambda)$ or $c_b(\lambda)$, and geometric scale factors, $m_s(i,e,g)$ or $m_b(i,e,g)$, which describe the intensity of the reflected light. Substituting these terms into equation (1) yields the Dichromatic Reflection Model equation:

$$L(\lambda,i,e,g) = m_s(i,e,g)c_s(\lambda) + m_b(i,e,g)c_b(\lambda) \tag{2}$$

The model thus describes the light that is reflected from an object point as a mixture of two distinct spectral power distributions, $c_s(\lambda)$ and $c_b(\lambda)$, each of which is scaled according to the geometric reflection properties of surface and body reflection. In the infinite-dimensional vector space of spectral power distributions (each wavelength defines an independent dimension in this vector space), the reflected light can thus be described as a linear combination of the two vectors $c_s(\lambda)$ and $c_b(\lambda)$.

When working with real images, the infinite-dimensional vector space of spectral color distributions is reduced through *spectral integration* to a three-dimensional *color space*. The process transforms the spectrum of an incoming light beam into a color triple $\mathbf{C} = [R,G,B]$. The Dichromatic Reflection Model is then describable as a function of two three-dimensional vectors, $\mathbf{C_s} = [R_s,G_s,B_s]$ and $\mathbf{C_b} = [R_b,G_b,B_b]$ which span a dichromatic plane in the three-dimensional color space:

$$\mathbf{C} = m_s\mathbf{C_s} + m_b\mathbf{C_b} \tag{3}$$

3. Object Shape and Color Variation

To use the Dichromatic Reflection Model in computer vision, it is important to discuss the relationship between the light mixtures of all points on an object. This involves studying the color variation over an entire object by projecting the colors of all object points into a color histogram and analyzing the histogram.

Since all points on one object depend on the same color vectors, $c_b(\lambda)$ and $c_s(\lambda)$, the light reflected from any such point is a linear combination of the same vectors. The light mixtures thus all fall into a *dichromatic plane* in the color space. Furthermore, an investigation of the geometrical properties of surface and body reflection reveals that the light mixtures form a dense color cluster in the dichromatic plane. The shape of this cluster is closely related to the shape of the object. The following discussion of color histograms uses perspective viewing and illumination geometry. For illustration purposes, it assumes that body reflection is approximately Lambertian and that surface reflection is describable by a function with a sharp peak around the angle of perfect mirror reflection. Note, however, that this analysis is not limited to a particular geometric reflection model. Figure 2 shows a sketch of a shiny cylinder. The left part of the figure displays the magnitudes of the body and surface reflection components. The curves show the loci of constant body or surface reflection. The darker curves are the loci of constant surface reflection. Since $m_s(i,e,g)$ decreases sharply around the object point with maximal surface reflection, m_{smax}, these curves are shown only in a small object area. The points in this area are called *highlight points*. The remaining object points are *matte points*. The right part of the figure shows the corresponding color histogram in the dichromatic plane. As will be described below, the object points form two linear clusters in the histogram.

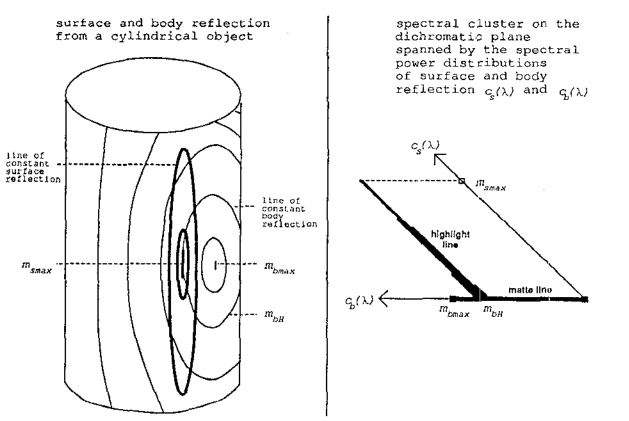

Figure 2: The shape of the color cluster for a cylindrical object

Light reflection at **matte points** is primarily determined by the body reflection process. Depending on the surface roughness, they may contain a small amount of surface reflection, as the result of light diffusion at the material surface. The following analysis assumes that the surface reflection component is small and constant. The observed light at matte points then depends mainly on $c_b(\lambda)$, scaled by $m_b(i,e,g)$ according to the geometrical relationship between the local surface normal of the object and the viewing and illumination directions. Consequently, the matte points form a *matte line* in the dichromatic plane in the direction of the body reflection vector, $c_b(\lambda)$, as shown in the right part of Figure 2.

Highlight points exhibit both body reflection and surface reflection. However, since $m_s(i,e,g)$ is much more sensitive to a small change in the photometric angles than $m_b(i,e,g)$, the body reflection component is generally approximately constant in a highlight area, as displayed by the curve with label m_{bH} in Figure 2. Accordingly, the second term of the Dichromatic Reflection Model equation (2) has an approximately constant value, $m_{bH}c_b(\lambda)$, and most of the color variation within the highlight comes from varying amounts of $m_s(i,e,g)$. The highlight points thus form a *highlight line* in the dichromatic plane in the direction of the surface reflection vector, $c_s(\lambda)$. The line departs from the matte line at position $m_{bH}c_b(\lambda)$, as shown in Figure 2. More precisely, the highlight cluster looks like a slim, skewed wedge because of the small variation of the body reflection component over the highlight.

The combined color cluster of matte and highlight points thus looks like a skewed T. The skewing angle of the T depends on the color difference between the body and surface reflection vectors while the position of the highlight line depends on the illumination geometry. Under reasonable assumptions, the highlight line starts in the upper 50 percent of the matte line[4]. We use this constraint, called the *50%-heuristic*, in the segmentation algorithm to distinguish color changes between matte and highlight points from color changes across material boundaries.

4. Influence of Camera Limitations

In order to apply the Dichromatic Reflection Model to real color images, we augment the model with a sensor model which accounts for several camera properties. First, the limited dynamic range of cameras adds *color clipping* and *blooming* to the sources of color variation in real images, causing the color clusters to bend and spread disproportionately at bright, saturated colors. Furthermore, the bands of a color image need to be *color balanced* to provide equal scaling along the axes of the color space. *Gamma-correction,* a non-linear transformation occurring in most cameras needs to be undone to restore the linear properties of body and surface reflection, which are essential to the Dichromatic Reflection Model. Finally, many cameras lenses exhibit chromatic aberration which causes pixel colors in areas of large color variation to deviate from the skewed T. A more detailed description, including color pictures with color clusters from real images that demonstrate these effects, can be found in[3, 4]. The effects are inherent in the picture taking process and need to be modeled in a successful color image analysis program.

5. Color Image Analysis

The Dichromatic Reflection Model and the sensor model can be used to analyze images of scenes that contain objects with highlights and shading. The goal of segmentation is to identify objects in an image, as delimited by material boundaries. Because most current color segmentation methods are based on a very simple interpretation of color changes in an image, they generally segment images not only along material boundaries but also along other lines exhibiting color or intensity variations, such as highlight and shadow boundaries, or internal object edges with significant shading changes. The Dichromatic Reflection Model provides a more sophisticated interpretation scheme relating the physics of light reflection to color changes in the image and in the color space. We use it to distinguish color changes at material boundaries from changes due to shading and highlights.

The Dichromatic Reflection Model and the sensor model describe how processes occurring in the scene and in the camera cause color variation in the image. A color image understanding algorithm has to invert this line of reasoning, inspecting the image and concluding from the color variations which processes have taken part in the image formation process. The problem is that the influence of any such process or combination of processes generally results in the same local image feature: color variation. To distinguish between the influences from several processes, the algorithm needs to inspect extended image areas. It needs to accumulate local color variations to determine distinguishing characteristics of the processes, such as the T-shape of a color cluster. But how can the optimal extent of an area be determined? There seems to be a circular problem of, on one hand, needing a prior segmentation to relate color variation to physical processes and, on the other hand, needing an understanding of the physical processes to provide a good segmentation.

Our segmentation method alternates between generating hypotheses from local image data and verifying whether the hypotheses fit the image, as shown in Figure 3. The hypotheses relate object color, shading, highlights and camera limitations to the shapes of color clusters in local image areas. The algorithm looks in a bottom-up process for color clusters from local image areas that exhibit the characteristic features of the body and/or surface reflection processes. When it finds a promising cluster in an image area, it generates a hypothesis. The algorithm then applies the new hypothesis to the image, using a region-growing approach to determine the precise extent of the image area to which the hypothesis applies. The physical knowledge embedded in the new hypothesis describes the object and highlight colors. It can be used to split every pixel in the new image area into its intrinsic body and surface reflection components. The resulting intrinsic images and the hypotheses together instantiate the general concepts of shading and highlights of the Dichromatic Reflection Model, describing the specific reflection processes that occur in this part of the scene.

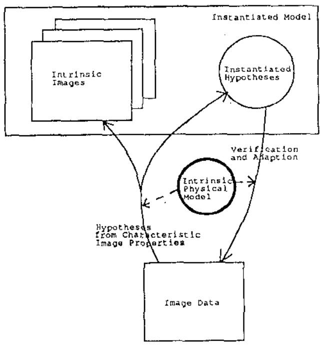

Figure 3: Using Dichromatic Reflection Model *for Color Image Segmentation*

The algorithm performs its generate-and-test analysis in several stages, each of which is related to a particular aspect of the Dichromatic Reflection Model or the sensor model. It starts out with the simplest and most common aspects and then uses the established knowledge to address the more complicated aspects. In this fashion, the algorithm performs the following steps:

1. Compute an initial, rough description of color variation in local image areas.
2. Generate hypotheses on matte color clusters. Exploit these hypotheses for image segmentation.
3. Extend the matte hypotheses into hypotheses on skewed T's in dichromatic planes. Resegment the image, exploiting these hypotheses.
4. Analyze the effects of blooming and color clipping.
5. Exploit the hypotheses about skewed T's of the segmented image areas to split all pixels into their reflection components.

This control structure exploits a physical model of light reflection to incrementally identify local and global properties of the scene, such as object and illumination colors. It uses these properties to interpret the pixels in the images. By using this control structure, the algorithm can adapt its image interpretation process to local scene characteristics and react differently to color and intensity changes at different places in the image. A detailed description of the algorithm is presented in[4, 10].

We start the image analysis process by generating, some initial, rough estimates about the color variations in the image. Such local color variations are described in terms of the principal components of the color distributions from small, non-overlapping windows of the image: depending on the number of non-zero eigenvalues, local color variations form a pointlike, linear, planar or volumetric cluster in the color space. This description can be related in a straightforward way to physical scene properties, such as linear color variation in matte object areas and dichromatic, color variation at highlights. We combine neighboring windows with similar color distributions into larger regions describing locally consistent color variation in the image.

The algorithm uses regions with linear descriptions to generate hypotheses on linear local color variation. The algorithm assumes that such linear hypotheses are related to matte shading on objects. We use the first eigenvector of the color distribution of such a region to approximate the direction of color variation, and thus the matte color vector, $\mathbf{C_b}$, of an object. A color cylinder with $\mathbf{C_b}$ as axis and a multiple of the estimated camera noise as width then describes a set of colors that are consistent with the hypothesized matte object color, under the presence of noise. Our algorithm uses this color cylinder to decide in a region growing step which pixels in the image to include into the region designated to the current hypothesis. The result is an image segmentation that outlines the matte object areas.

To include highlights into the segmented object area, the algorithm then extends the linear hypotheses into planar hypotheses, combining the hypothesized matte lines with highlight lines. To be combined with a highlight line, the color cluster of the matte object area needs to form a skewed T with the color cluster of a neighbor. Such a highlight area must be distinguished from a neighboring matte area of another object. Our main criterion for distinguishing these two types of neighbors is the 50%-heuristic mentioned in section 3: the color clusters of a matte and a highlight region intersect in the upper 50 percent of the matte cluster, while color clusters from neighboring matte regions converge at dark pixels. If a highlight cluster is found, it determines the highlight line and thus the surface reflection vector, $\mathbf{C_s}$, of an object. The matte and highlight lines form together a planar hypothesis. We then use the resulting planar hypothesis, extended into a planar slice to account for noise, to grow the linear object area into the highlight. After accounting for the effects of color clipping and blooming[4, 10], we have a color image segmentation that outlines the material boundaries, ignoring color changes due to shading or highlights.

As one application of the above method, we use the gathered information about the body and surface reflection vectors to split every color pixel into its two reflection components[3, 4, 10]. The result are two intrinsic images of the scene, one showing the objects as if they were completely matte, and the other showing only the highlights. To split the pixels into their reflection components, the algorithm uses the reflection vectors and their cross product to define a new (not necessarily orthogonal) coordinate system in the color cube. This coordinate system describes every color in the cube in terms of the amounts of body reflection, surface reflection and noise. There exists an affine transformation, and thus a linear transformation which transforms the color vector every pixel into an intrinsic vector describing the the reflection components and the associated noise of the pixel. The *body reflection image* of an image area consists then of the body reflection components of the transformed vectors of all pixels. The corresponding *surface reflection image* consists of the surface reflection components. The noise components provide the *intrinsic noise image*.

6. Discussion

Figure 4 shows the results of analyzing a color image under the guidance of the Dichromatic Reflection Model. The upper left quarter shows the original image. The upper right quarter shows the generated image segmentation. It demonstrates that our segmentation algorithm is able to outline material boundaries, while ignoring color variation due to highlights and shading. In contrast, a traditional color segmentation method, such as Phoenix[9], is unpredictable in its analysis of highlight areas on objects: on some objects, it includes the highlights into the object areas while it draws a boundary between matte and highlight areas on other objects[4]. The reason for such unpredictable behavior is that the segmentation decisions of Phoenix are not related to an analysis of physical reflection processes.

The lower quarters of Figure 4 show the intrinsic reflection images of the scene. They demonstrate that the algorithm is able to determine the body and surface reflection components of the various objects. The body reflection image may be a useful tool to determine object shapes from shading information[5]. The surface reflection image exhibits gradual changes between areas with no surface reflection and areas with very high amounts of surface reflection, demonstrating that surface reflection is a function of the photometric angles. It may be useful input for methods that determine object shapes from highlights[6].

Our approach has a more solid foundation than heuristic image segmentation methods; yet, it does have some key limitations[4]. The conceptional limitations of this approach are related to the basic principles of the Dichromatic Reflection Model. Since it attributes any linear color variation to the changing illumination geometry of a single material, the segmentation algorithm is unable to find material boundaries between objects with collinear matte clusters. A geometrical analysis, linking intensity gradients to object shape, will be needed to distinguish between such objects. The same will be necessary to analyze dark image areas which are currently excluded because their color information is too unreliable.

The model makes simplifying assumptions about the illumination conditions and the materials in the scene. A color cluster from an object in an unconstrained scene will generally not be a skewed T composed of linear subclusters because the illumination color may vary on different parts on the object surface, and the reflection properties of the object may also change, due to illumination changes and to pigment variations in the material body. The necessary extensions to the model will be the subject of future work.

The approach also has problems, if the T-shape of the cluster degrades into a line. Such degradation occurs, if the two linear clusters are collinear or if either of the two clusters is missing. Collinearity arises, if an object has the same color as the light source or if the object is a uniform reflector, such as grey objects or pale pastels. The matte cluster is missing, if the viewed object is very dark, or if the scene is illuminated with a narrow-band illuminant that does not overlap with the wavelengths at which the material reflects light. Matte clusters also do not exist for metallic objects. The highlight cluster may be missing, if an object does not reflect a highlight into the camera, due to its position in the scene and the illumination geometry. Finally, on objects with rough surfaces (on a fine scale), every pixel has both a significant body and surface reflection component and the resulting color clusters may fill out the entire dichromatic planes. A common special case of this are so-called "matte" materials - as opposed to glossy materials - which reflect a constant amount of surface reflection in every direction and thus never exhibit a highlight in the common sense of the word. The corresponding color clusters are linear clusters, translated from the origin of the color space according to the constant surface reflection component. The current method is not capable of distinguishing between all these cases. In combination with exploiting previously determined scene properties, such as the illumination color, it will be necessary to analyze the intensity gradients along the linear axes and relate them to the geometrical properties of the body and surface reflection components.

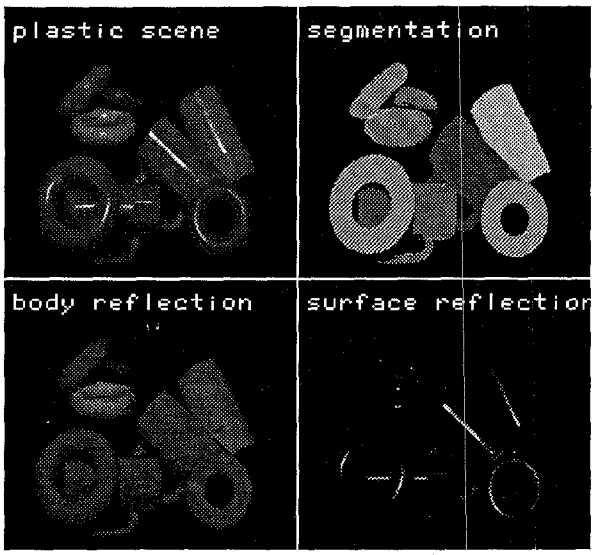

Figure 4: Color image analysis and segmentation on a color image with 8 plastic objects

7. Conclusions

In this paper, we have demonstrated that it is possible to analyze and segment real color images by using a physics-based color reflection model. Our model accounts for highlight reflection and matte shading, as well as for some characteristics of cameras. By developing a physical description of color variation in color images, we have developed a method to automatically segment an image while generating hypotheses about the scene. We then use the knowledge we have gained to separate highlight reflection from matte object reflection. The resulting intrinsic reflection images have a simpler relationship to the illumination geometry than the original image and may thus improve the results of many other computer vision algorithms, such as motion analysis, stereo vision, and shape from shading or highlights. Since the surface reflection component of dielectric materials generally has the same color as the illumination, we can also determine the illumination color from the intrinsic surface reflection image, information which is needed by color constancy algorithms[11].

Our hypothesis-based approach towards image segmentation may provide a new paradigm for low-level image understanding. Our method gains its strength from using an intrinsic model of physical processes that occur in the scene. The result are intrinsic images and hypotheses which are closely related in their interpretation to the intrinsic model, being instantiations of concepts formulated in the model. Our system alternates between a bottom-up step which generates hypotheses and a top-down step which applies the hypotheses to the images. Our analysis thus consists of many small interpretation cycles that combine bottom-up processing with feedback in top-down processing. This approach stands in contrast to traditional image segmentation methods which do not relate their analysis to intrinsic models and that also generally have a strictly bottom-up control structure and data flow. We feel that many low-level image understanding methods such as shape-from-x methods, stereo and motion analysis may be viewed and approached under this paradigm. We hope to extend our approach into a more complete low-level image analysis system which combines color analysis with a geometrical analysis of the scene, exploiting the body and surface reflection images. Along these lines, we may generate hypotheses about object shapes and about the object materials (see [12]). The highlight image may also provide strong evidence for the position of the light source.

References

1. H.G. Barrow, and J.M Tenenbaum, *Recovering Intrinsic Scene Characteristics from Images,* Academic Press, New York, 1978, pp. 3-26.

2. S.A. Shafer, "Using Color to Separate Reflection Components", *COLOR research and application*, Vol. 10, No. 4, Winter 1985, pp. 210-218, Also available as technical report TR-136, Computer Science Department, University of Rochester, NY, April 1984

3. G.J. Klinker, S.A. Shafer, and T. Kanade, "The Measurement of Highlights in Color Images", *International Journal on Computer Vision (IJCV)*, Vol. 2, No. 1, 1988, pp. 7-32.

4. G.J. Klinker, *A Physical Approach to Color Image Understanding,* PhD dissertation, Computer Science Department, Carnegie-Mellon University, May 1988, Available as technical report CMU-CS-88-161

5. B.K.P. Horn, "Understanding Image Intensities", *Artificial Intelligence*, Vol. 8, No. 11, 1977, pp. 201-231.

6. G. Healey and T.O. Binford, "Local Shape from Specularity", *Proceedings of the First International Conference on Computer Vision (ICCV),* IEEE, London, June 1987, pp. 151-161.

7. R. Ohlander, K. Price, and D.R. Reddy, "Picture Segmentation Using a Recursive Region Splitting Method", *Computer Graphics and Image Processing*, Vol. 8, 1978, pp. 313-333.

8. Y. Ohta, T. Kanade, and T. Sakai, "Color Information for Region Segmentation", *Computer Graphics and Image Processing*, Vol. 13, 1980, pp. 222-231.

9. S.A. Shafer and T. Kanade, "Recursive Region Segmentation by Analysis of Histograms", *Proceedings of the International Conference on Acoustics, Speech, and Signal Processing*, IEEE, Paris, France, May 1982, pp. 1166-1171.

10. G.J. Klinker, S.A. Shafer, and T. Kanade, "Image Segmentation and Reflection Analysis Through Color", *Proc. SPIE #937: Applications of Artificial Intelligence VI*, SPIE, Orlando, Florida, April 4-8 1988, pp. 229-244, Also in the proceedings of the DARPA Image Understanding Workshop, April 6-8, 1988, Boston, MA

11. L.T. Maloney and B.A. Wandell, "Color constancy: a method for recovering surface spectral reflectance", *Journal of the Optical Society of America A (JOSA-A)*, Vol. 3, No. 1, January 1986, pp. 29-33.

12. G. Healey, and T.O. Binford, "Predicting Material Classes", *Proc. of the DARPA-Image Understanding Workshop*, Morgan Kaufmann, Boston, MA, April 6-8 1988.