Photometric Invariant Region Detection in Multi-Spectral Images

Harro Stokman and Theo Gevers
ISIS, University of Amsterdam, Kruislaan 403
1098 SJ Amsterdam, The Netherlands
{stokman, gevers}@wins.uva.nl

Abstract

Our aim is to detect homogeneously colored regions invariant to surface orientation change, illumination, shadows and highlights in multi-spectral images where the spectral range corresponds to the visible wavelength interval. To this end, the influence of multi-spectral sensor space, normalized multi-spectral sensor space, and hue color space are examined, in theory, for the dichromatic reflection model and, in practice, for segmentation techniques based on k-means clustering. We show that homogeneously colored regions can be detected invariant to surface orientation change, shadows and highlights under the condition of equal-energy illumination where $e(\lambda) = e = constant$.

In this paper, we first present a method that achieves, in theory and in practice, an approximation of equal-energy illumination. The method requires that the spectral distribution of the illuminant is known. Secondly, we derive in theory three cluster models: points, lines and planes and show the invariance for each model to surface orientation change, illumination, shadows and highlights. We then present segmentation algorithms which incorporate these models and which are based on the k-means clustering technique. Experiments are conducted on multi-spectral images taken from colored objects in real-world scenes.

On the basis of the theoretical and experimental results on multi-spectral images it is concluded that the line and plane model detect regions invariant to a change in surface orientation, viewpoint of the camera, and illumination intensity. Furthermore, the plane model also detect regions independent of highlights. The point model provides segmentation results which is sensitive to surface orientation and illumination intensity.

1 Introduction

In the past, various color image segmentation methods have been proposed which do not account for the image formation process, [2], [6], for example. A major drawback of these color segmentation methods is that the values of the color features, on which the segmentation method relies, depend on the geometry of the object, the viewpoint of the camera or the illumination conditions. As a consequence, the obtained segmentation results are negatively affected by shadows, shading and highlights. In contrast, color segmentation methods that account for the process of image formation are studied recently. One of the first methods based on physics considerations is Klinker *et al.* [5]. They developed a color segmentation algorithm based on the dichromatic reflection model proposed by Shafer [7]. The method is based on evaluating "dog-leg" clusters in RGB-space. Then follows segmentation independent of the object's geometry, illumination and highlights. To achieve robust image segmentation, however, surface patches of objects in view must have a rather broad distribution of surface normals which may not hold for objects in general. Bajscy *et al.* [1] developed a similar image segmentation method using the $H-S$ color space instead of the RGB-color space. Healey[4] proposes a method to segment images on the basis of normalized color. However, as normalized color is sensitive to highlights, a separate highlight removal technique should be applied on the image prior to the actual segmentation. Various color features have been studied using a multi-resolution segmentation method [6]. However, no (reflection) model of the color imaging process is considered to explain the observed effects.

Our aim is to detect homogeneously colored regions invariant to surface orientation change, illumination, shadows and highlights in multi-spectral images. The spectral range is the visible wavelength interval. In order to obtain color constancy, we develop a method that achieves an approximation of equal-energy illumination where $e(\lambda) = e = constant$. We then analyze, in theory, the influence multi-spectral sensor space, normalized multi-spectral sensor space, and hue color space, for the dichromatic reflection model.

Based on these theoretical observations, we derive three cluster models in multi-spectral sensor space: points, lines and planes and show the sensitivity or in-

variance for each model to surface orientation change, illumination, shadows and highlights. We then develop segmentation algorithms which incorporate these models and which are based on the k-means clustering technique. Experiments are conducted on multi-spectral images taken from colored objects in real-world scenes.

The paper is organized as follows. In Section 2, the dichromatic reflection model under equal-energy illumination is given. The reflection model is used to study and analyze the multi-spectral subspace on which colors will be projected from an uniformly painted surface patch. In Section 3, segmentation methods are discussed. Experiments carried out with these segmentation methods will be given in Section 4. A summary is given in Section 5.

2 Photometric Color Invariance

2.1 The Reflection Model

Consider an image of an infinitesimal surface patch. Using N narrow-band filters with spectral sensitivities given by $f_1(\lambda)...f_N(\lambda)$ to obtain an image of the surface patch illuminated by a spectral power distribution (SPD) of the incident light denoted by $e(\lambda)$, the measured sensor values will be given by Shafer [7]:

$$C_n = m_b(\vec{n},\vec{s}) \int_\lambda f_n(\lambda)e(\lambda)c_b(\lambda)d\lambda + m_s(\vec{n},\vec{s},\vec{v}) \int_\lambda f_n(\lambda)e(\lambda)c_s(\lambda)d\lambda \quad (1)$$

for C_n giving the nth sensor response. Further, $c_b(\lambda)$ and $c_s(\lambda)$ are the albedo and Fresnel reflectance respectively. λ denotes the wavelength, $\vec{n}$ is the surface patch normal, $\vec{s}$ is the direction of the illumination source, and $\vec{v}$ is the direction of the viewer. Geometric terms m_b and m_s denote the geometric dependencies on the body and surface reflection respectively.

2.2 Illumination Independence

Consider the reflectance of a matte, white reference with constant spectral response, $c_b(\lambda) = \text{constant} = 1$. Assume that the surface normal is equal to the viewing direction, thus $m_b(\vec{n},\vec{s}) = 1$. As no specularities occur, the measured sensor values are, according to Eq. 1:

$$C_n = \int_\lambda f_n(\lambda)e(\lambda)d\lambda \quad (2)$$

for C_n giving the nth sensor response for the white reference sample. This way, the spectral transmittance of the illuminant is measured. Next, an image is captured of an arbitrary scene, and is divided by the previously obtained recording of the illuminant. We thus have:

$$C_n = \frac{m_b(\vec{n},\vec{s}) \int_\lambda f_n(\lambda)e(\lambda)c_b(\lambda)d\lambda}{\int_\lambda f_n(\lambda)e(\lambda)d\lambda} + \frac{m_s(\vec{n},\vec{s},\vec{v}) \int_\lambda f_n(\lambda)e(\lambda)c_s(\lambda)d\lambda}{\int_\lambda f_n(\lambda)e(\lambda)d\lambda} \quad (3)$$

for C_n giving the nth sensor response. Now assume that the filter $f_n(\lambda)$ is a narrow band filter modeled as a unit impulse that is shifted over n wavelengths: The transmission at $\lambda_n = \delta$ and zero elsewhere. This allows us to rewrite Eq. 3 as:

$$C_n = \frac{m_b(\vec{n},\vec{s})f_n(\lambda)e(\lambda)c_b(\lambda)}{f_n(\lambda)e(\lambda)} + \frac{m_s(\vec{n},\vec{s},\vec{v})f_n(\lambda)e(\lambda)c_s(\lambda)}{f_n(\lambda)e(\lambda)} \quad (4)$$

and thus

$$C_n = m_b(\vec{n},\vec{s})c_b(\lambda) + m_s(\vec{n},\vec{s},\vec{v})c_s(\lambda) \quad (5)$$

Eq. 5 implies that, under the assumption of a unit impulse band filter, the imagery can be made independent of the illuminant, if the original image is divided by the spectral transmittance of the illuminant.

However, in practice, the color filters are never unit impulse band filters. Eq. 5 now only holds if $e(\lambda)$, $c_b(\lambda)$ and $c_s(\lambda)$ have constant transmittance over the wavelength pass interval of the filter. Under the assumption that natural objects, due to physics restrictions, show somehow smooth spectral power distributions and under the assumption that the illuminant used has a gradual changing SPD, we might still be able to approximate Eq. 5.

2.3 Reflection with White Illumination

As no color exists without illumination, we illuminate with an *equal-energy spectrum* illuminant on the wavelength basis: $e(\lambda) = \text{e} = \text{constant}$. Considering dichromatic reflectance and equal-energy or "white" illumination, then $c_s(\lambda) = c_s$. The measured sensor values are then:

$$C_n = em_b(\vec{n},\vec{s})k_n + em_s(\vec{n},\vec{s},\vec{v})c_s \int_\lambda f_n(\lambda)d\lambda \quad (6)$$

for C_n giving the nth sensor response. under the assumption of an equal-energy or white light source. $k_n = \int_\lambda f_n(\lambda)c_b(\lambda)d\lambda$ is a compact formulation depending on the sensors and the surface albedo.

If the integrated white condition holds (as we assume throughout the paper) then:

$$\int_\lambda f_1(\lambda)d\lambda = \int_\lambda f_2(\lambda)d\lambda = ... = \int_\lambda f_N(\lambda)d\lambda = f \quad (7)$$

and thus:

$$C_n = em_b(\vec{n}, \vec{s})k_n + em_s(\vec{n}, \vec{s}, \vec{v})c_s f \quad (8)$$

In the next section, this reflection model will be used to study the behavior of the different color spaces in multi-spectral space.

2.3.1 Photometric Invariant Color Features for Matte, Dull Surfaces

Consider the body reflection term of Eq. 8:

$$C_{n,b} = em_b(\vec{n}, \vec{s})k_n \quad (9)$$

for $C_{n,b}$ giving the nth sensor response of a matte, dull surface patch under the assumption of equal-energy illumination.

According to the body reflection term, the color depends only on k_n (i.e. sensor and surface albedo) and the brightness on factor $em_b(\vec{n}, \vec{s})$. Then the observed colors of a matte, dull surface with fixed k_n can be represented by the color cluster vector $\vec{B}$, see Fig. 1, where the direction of $\vec{B}$ is based on k_n and its extent by $em_b(\vec{n}, \vec{s})$. As a consequence, a uniformly painted surface (i.e. with fixed k_n) may give rise to a broad variance of sensor values due to the varying circumstances induced by the image-forming process such as a change in object orientation, illumination intensity and position.

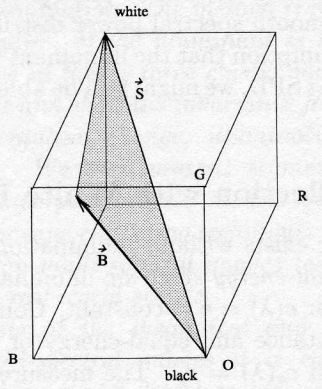

Figure 1: *Color cluster vector $\vec{B}$ of a matte surface in RGB, and color cluster vector $\vec{S}$ of a shiny surface in RGB.*

In contrast, normalized sensor space, for instance normalized rgb, or in the multi-spectral case, $c_1, c_2, ...c_{N-1}$ is insensitive to surface orientation, illumination direction and intensity as can be seen from:

$$c_1(C_1, C_2, ..C_N) = \frac{em_b(\vec{n}, \vec{s})k_1}{em_b(\vec{n}, \vec{s})(k_1 + k_1 + ... + k_N)} = \frac{k_1}{k_1 + k_1 + ... + k_N} \quad (10)$$

only dependent on the sensors and the surface albedo. Equal arguments hold for $c_2..c_{N-1}$.

Similarly, hue H is an invariant for matte, dull surfaces. Assume that three color filters C_1, C_2, C_3 correspond to the RGB filters of a color camera, so $C_1 = R$, $C_2 = G$, $C_3 = B$ Let H then be defined as

$$H(C_1, C_2, C_3) = \arctan\left(\frac{\sqrt{3}(C_2 - C_3)}{(C_1 - C_2) + (C_1 - C_3)}\right) \quad (11)$$

Substitution of Eq. 9 in Eq. 11 produces:

$$H(C_{1,b}, C_{2,b}, C_{3,b}) = \arctan\left(\frac{\sqrt{3}em_b(\vec{n}, \vec{s})(k_{C_1} - k_{C_3})}{em_b(\vec{n}, \vec{s})((k_{C_1} - k_{C_2}) + (k_{C_1} - k_{C_2}))}\right) = \arctan\left(\frac{\sqrt{3}(k_{C_2} - k_{C_3})}{(k_{C_1} - k_{C_2}) + (k_{C_1} - k_{C_3})}\right) \quad (12)$$

Obviously, in practice, the assumption of objects composed of matte, dull surfaces is not always realistic. To that end, the effect of surface reflection (highlights) is discussed in the following section.

2.3.2 Photometric Invariant Color Features for Both Matte and Shiny Surfaces

Consider the surface reflection term of Eq. 8:

$$C_s = em_s(\vec{n}, \vec{s}, \vec{v})f \quad (13)$$

for a RGB camera where $C_1 = R$, $C_2 = G$, $C_3 = B$ and thus $C_s \in \{C_{1,s}, C_{2,s}, C_{3,s}\}$ giving the red, green and blue sensor response for a highlighted surface patch with white illumination.

Note that under the given conditions, the color of highlights is not related to the color of the surface on which they appear, but only on the color of the light source. Thus for the white light source, the surface reflection color cluster $\vec{S}$ is on the achromatic axis of the color space corresponding to intensity I, see Figure 1.

For a given point on a shiny surface, the contribution of the body reflection component C_b and surface reflection component C_s are added together $C_w = C_s + C_b$. Hence, the observed colors of the surface must be inside the triangular color cluster of a three dimensional color space formed by the two reflection components, see Fig. 1.

Because H is a function of the angle between the *main diagonal* or achromatic axis where $C_1 = C_2 = ... = C_N$ and the color point, all possible colors of the same (shiny) surface region (i.e. with fixed albedo) have

to be of the same hue as follows from substituting Eq. 8 in Eq. 11:

$$\begin{aligned} H(C_{1,w}, C_{2,w}, C_{3,w}) &= \\ \arctan\left(\frac{\sqrt{3}(C_{2,w} - C_{3,w})}{(C_{1,w} - C_{2,w}) + (C_{1,w} - C_{2,w})}\right) &= \\ \arctan\left(\frac{\sqrt{3}em_b(\vec{n},\vec{s})(k_{C_2} - k_{C_3})}{em_b(\vec{n},\vec{s})((k_{C_1} - k_{C_2}) + (k_{C_1} - k_{C_3}))}\right) &= \\ \arctan\left(\frac{\sqrt{3}(k_{C_2} - k_{C_3})}{(k_{C_1} - k_{C_2}) + (k_{C_1} - k_{C_3})}\right) \quad (14) \end{aligned}$$

factoring out dependencies on illuminant e, object geometry $m_b(\vec{n},\vec{s})$, viewpoint $m_s(\vec{n},\vec{s},\vec{v})$, and specular reflection coefficient c_s and hence only dependent on the sensors and the surface albedo. Note that $C_{1,w} = em_b(\vec{n},\vec{s})k_{C_1} + em_s(\vec{n},\vec{s},\vec{v})c_s f$, $C_{2,w} = em_b(\vec{n},\vec{s})k_{C_2} + em_s(\vec{n},\vec{s},\vec{v})c_s f$, and $C_{3,w} = em_b(\vec{n},\vec{s})k_{C_3} + em_s(\vec{n},\vec{s},\vec{v})c_s f$.

It is important to note that if the assumption of white illumination is *not* valid, that then the hue is *not* invariant for highlights. Consider an infinitesimal surface patch with $m_b(\vec{n},\vec{s}) = 0$ (for instance transparent glass), where the Fresnel reflectance is constant $c_s(\lambda) = c_s$. The measured sensor values are

$$C_n = c_s m_s(\vec{n},\vec{s},\vec{v}) \int_\lambda f_n(\lambda)e(\lambda)d\lambda \quad (15)$$

Comparison of Eq. 15 and Eq. 2 tells that the specular reflectance has the same spectral distribution as the illuminant, though not the same intensity. Thus, the main diagonal spanned by the surface patch under varying $\vec{n},\vec{s},\vec{v}$ has the same orientation as the illuminant. As the illuminant is not white, the main diagonal does not correspond to the achromatic axis. Because H is a function of the angle between the *main diagonal* or achromatic axis where $C_1 = C_2 = ... = C_N$ and the color point, it is necessary to obtain equal-energy illumination using Eq. 3 for hue to be valid.

3 Segmentation Method

The k-means method is based on the minimization of the sum of squared distances of all pixels to their cluster means. First an initial partition is given, obtained by assigning each pixel to its closest cluster mean, on basis of the Euclidean distance measure, and new partitions are obtained by moving pixels from one cluster to another until the summed square distance is reduced to its minimum. For a more detailed description, see for example [3].

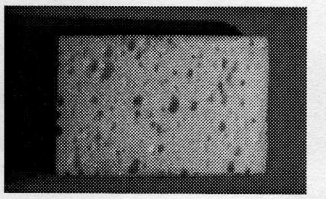

Figure 2: *Object exhibiting a large variety of body reflection.*

3.1 Clustering by Fitting Points, Lines (Matte Objects) or Planes (Shiny Objects) Through *RGB*-data

A naive approach for image segmentation is to fit k points (k-*means*) in *RGB* space. As was shown in Section 2.3.1, a uniformly painted surface may give rise to broad variance of *RGB* values due to varying circumstances induced by the image-forming process such as a change in object orientation, illumination and position.

An alternative way is to cluster on k straight lines (k-*lines*) starting from the origin in *RGB*-color space. This can be seen as follows. According to the body reflection model given by Eq. 9, the observed colors of a matte, dull surface with fixed k_n are represented by the color cluster vector $\vec{B}$, see Fig. 1. Therefore, we propose that the alternative distance measure is defined as the shortest distance of the observed color to a line starting from the origin. To illustrate this, consider the color image in Fig. 2. The corresponding multi-spectral data shown in Fig. 3 shows two distinct colors. The k-lines clustering method accurately detects lines starting at the origin through the multi-spectral data.

Similarly, to find homogeneous colored *shiny* surfaces, we propose an alternative distance measure which is based on fitting planes (k-*planes*), originating at the achromatic axis, to the observed colors in the *RGB*-color cube. In other words, a region is considered to be homogeneous if the *RGB*-colors form a plane, see Figure 1, in the *RGB*-color cube at one side of the achromatic axis.

4 Experiments

In this section, we evaluate and compare the performance of the various cluster models for the purpose of color invariant image segmentation.

4.1 Preprocessing

First, in Eq. 3 it was shown that if the spectral transmission of the illuminant is known, division of each

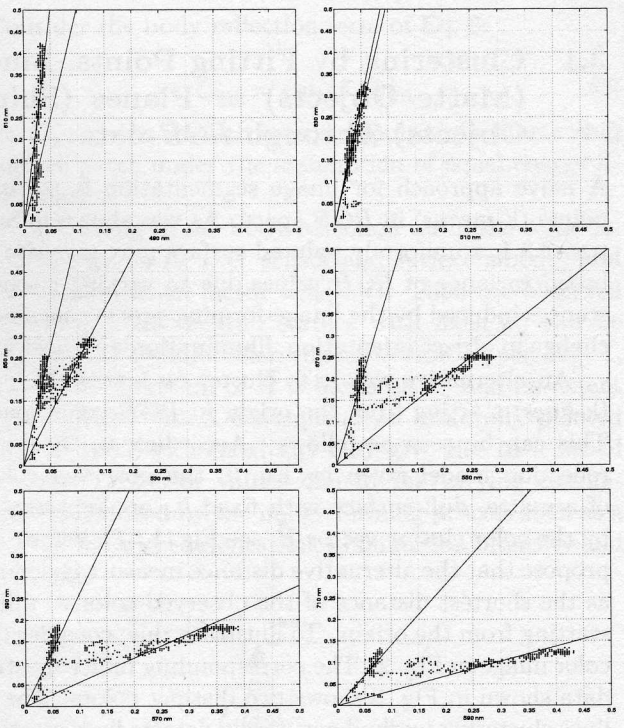

Figure 3: *From the image shown in Fig. 2, multi-spectral data is obtained. The spectra are sampled at bandwidths of 20 nm. Next, the clustering algorithm fits lines through the multi-spectral pixels. The graphs shown here contain the plots of the same pixels, each time for two different spectra. The graphs show that pixels form lines starting at the origin, due to varying geometry of the objects. By the fitting of lines through the pixel clusters, segmentation can be achieved invariant to illumination, geometry and shadows.*

color band by the spectral transmission of the illuminant for that color band, will approximately factor out the illuminant. This way, equal-energy illumination is achieved, on which the segmentation algorithm relies. This pre-processing step is demonstrated in Fig. 4, 5. The spectra were recorded using an Imspector V7 spectrograph with the aid of a Jain CV-M300 gray value camera the Matrix Magic Color frame grabber. Illuminants are halogen and a Dolan & Jenner Fiber-Lite PL800 Illuminator. The digitization was done in 8 bits. In Fig. 4a, the spectral transmission of two different illuminants is shown. The bandwidth of the narrow band filters is 5 nm. In Fig. 4b, the spectral transmission of the orange yellow patch of the MacBeth ColorChecker is shown, illuminated by the two light sources. In Fig. 4c, the reconstructed spectral transmission of the sample under equal-energy illumination is shown. Between wavelengths 500 nm and 650 nm, both reconstructed spectra appear similar. As is clear from this result, Eq. 3 is sensitive to the spectral power distribution of the illuminants. For example, the transformation shows to be unstable between 400 and 500 nm wavelength.

In Color Fig. 5, the experiment is repeated for an *RGB* image depicting a red, green and a blue pen. The objects were recorded with the aid of the SONY XC-003P CCD color camera (3 chips) and the Matrox Magic Color frame grabber. The digitization was done in 8 bits per color. In Fig. 5a, the scene is illuminated by a halogen light source. The camera is white balanced, yielding equal-energy illumination. Fig. 5b contains the same scene illuminated by yellow filtered tungsten light source. Fig. 5c shows the reconstruction of Fig. 5b as if the scene was illuminated by the equal-energy halogen light source. By comparison of Fig. 5a and Fig. 5c, the overall color appearance is approximately similar. However, the tungsten light source is tubular shaped, while the halogen light source is a point source. Therefore, highlights (dependent on direction of the light source) show up differently. In Fig. 5c, the transformation to equal-energy illumination is shown graphically: The dashed color cube contains the *RGB* space due to the yellow filtered tungsten light source. With a change in illumination, both the size and shape of the color cube change. The equally sided larger cube corresponds to white, equal-energy illumination. Scaling of the axes of the yellow filtered tungsten light cube thus yields an approximation of the equal-energy illumination color cube.

4.2 Segmentation

Having the equal-energy illumination condition approximately fixed, we now consider the segmentation results

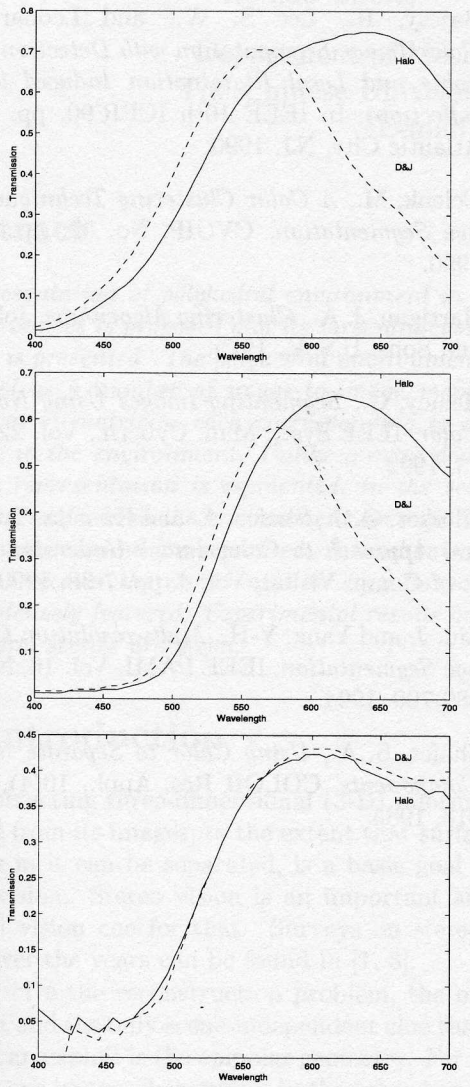

Figure 4: *a. Spectral transmission of the two illuminants. b.Spectral transmission of the orange yellow patch of the MacBeth ColorChecker, illuminated by the two illuminants. c. Reconstructed spectral transmission of the orange yellow patch under equal-energy illumination.*

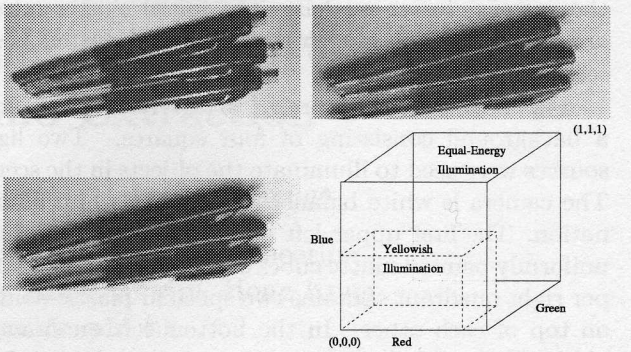

Figure 5: *a. Color image under white illuminant. b. Color image under yellow filtered tungsten illuminant c. Reconstructed white illuminant image from yellow illuminant image d. Transformation to equal-energy illumination: The dashed color cube contains the RGB space due to the yellow filtered tungsten light source. With a change in illumination, both the size and shape of the color cube change. The equally sided larger cube corresponds to white, equal-energy illumination. Scaling of the axes of the yellow filtered tungsten light cube thus yields an approximation of the equal-energy illumination color cube.*

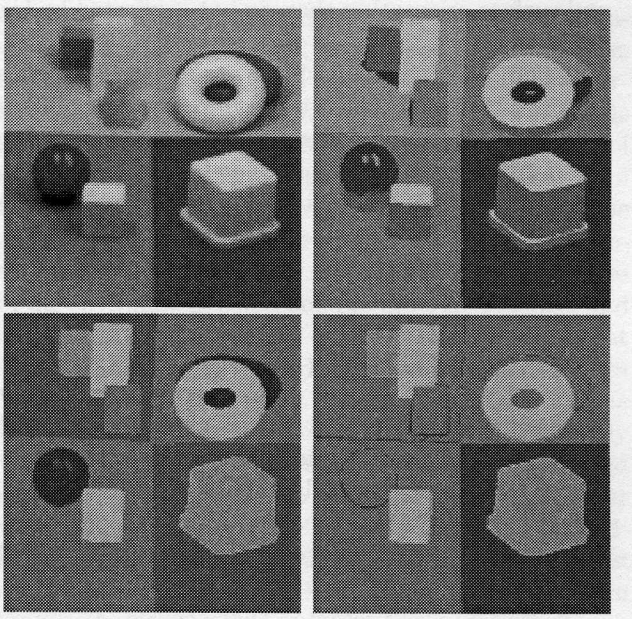

Figure 6: *Segmentation results for the k-means clustering method. a. Color image b. Cluster model is a point, region detection is sensitive to intensity changes, shadows, geometry, highlights and color transitions. c. Cluster model is a line, region detection is sensitive to highlights and color transitions. d. Cluster model is a plane, region detection is sensitive only to color transitions.*

obtained by the k-means algorithm differentiated for the various cluster models. The initial cluster means are similar for each model, the number of clusters is set to five.

Fig. 6 a is a color image of several toys against a background consisting of four squares. Two light sources were used to illuminate the objects in the scene. The camera is white balanced to achieve white illumination. The first upper left quadrant consists of three uniformly painted matte cubes of wood. The second upper right quadrant contains two specular plastic donuts on top of each other. In the bottom left quadrant a red highlighted ball and a matte cube are shown while the last quadrant contains two matte cubes. Note that each individual object is painted uniformly with a distinct color. The image is contaminated by noise, shadows, shading, inter-reflections and specularities. Inter-reflections occur when an object receives the reflected light from other objects. In Fig. 6b the segmentation result is shown obtained by the k-means clustering method where the cluster model is a point. Clearly, false regions are introduced by abrupt surface orientations, shadows, inter-reflections and highlights.

In contrast, the result of fitting straight lines through RGB data by the method described in Section 3.1, shown in Fig. 6c, is insensitive for shadows and surface orientation changes but is affected by highlights.

Good performance is shown for the results that are obtained by fitting planes through RGB-data, shown in Fig. 6d by the method described in Section 3.1. Here, computed region edges correspond to material boundaries discounting the disturbing influences of surface orientation, illumination, shadows and highlights.

5 Conclusion

We analyzed, in theory, multi-spectral sensor space, normalized multi-spectral sensor space, and hue color space for the dichromatic reflection model. It was shown that homogeneously colored regions can be detected invariant to surface orientation change, shadows and highlights under the condition of equal-energy illumination where $e(\lambda) = c = \text{constant}$.

A method was presented that achieves an approximation of equal-energy illumination. The method requires that the spectral distribution of the illuminant is known. Then, three cluster models were derived in theory: points, lines and planes. On the basis of the theoretical and experimental results on multi-spectral images it is concluded that the line and plane model detect regions invariant to a change in surface orientation, viewpoint of the camera, and illumination intensity. Furthermore, the plane model also detect regions independent of highlights. The point model provides segmentation results which is sensitive to surface orientation and illumination intensity.

References

[1] Bajcsy, R., Lee S. W., and Leonardis, A., *Color Image Segmentation with Detection of Highlights and Local Illumination Induced by Inter-reflections*, In IEEE 10th ICPR'90, pp. 785-790, Atlantic City, NJ, 1990.

[2] Celenk, M., *A Color Clustering Technique for Image Segmentation*, CVGIP, No. 52, pp. 145-170, 1990.

[3] Hartigan, J. A., *Clustering Algorithms*, John Wiley and Sons, U.S.A., 1975.

[4] Healey, G., *Segmenting Images Using Normalized Color*, IEEE Syst., Man. Cybern., Vol. 22, pp. 64-73, 1992.

[5] Klinker, G. J., Shafer, A. and Kanada, T. *A Physical Approach to Color Image Understanding*, Int. J. of Comp. Vision, Vol. 4, pp. 7-38, 1990.

[6] Liu, J. and Yang, Y-H., *Multi-resolution Color Image Segmentation*, IEEE PAMI, Vol. 16, No. 7, pp. 689-700, 1994.

[7] Shafer, S. A., *Using Color to Separate Reflection Components*, COLOR Res. Appl., 10(4), pp 210-218, 1985.