Segmentation of Color Images

T. Gevers and F.C.A. Groen

Faculty of Mathematics & Computer Science

University of Amsterdam

Kruislaan 409

1098 SJ Amsterdam

The Netherlands

november 1990

Abstract

Because segmentation results of grey-value images might be negatively affected by the presence of intensity changes due to shadows and surface curvature in those images, we have investigated to what extent these segmentation results can be improved by using color information.
The influence of different color systems has been examined as well as two segmentation methods were investigated: a clustering and a region growing technique. The clustering technique is based on the k-means algorithm, and is iterative. The region growing technique is derived from the split&merge algorithm proposed by [4].In the region growing technique, homogeneity criteria play an important role. Two homogeneity criteria have been implemented: a homogeneity criterion based on functional approximation and the mean&variance homogeneity criterion.
The results show a reasonable insensibility to changes in intensity due to shadows and surface curvature, whereby the region growing technique produces better results than the clustering algorithm. Color systems which are linear combinations of the same RGB basis provide similar segmentation results. The IHS, rgb and the $L^*a^*b^*$ color systems provide poor segmentation results when intensity is small. The rgb color system provides the best segmentation results if the image contains intensity changes.

keywords: color feature, color system, metamerism, color ratio, color gamut, primary color, chromaticity coordinates, chromaticity diagram, clustering, region growing, functional approximation, spatial color progress coherence, witness.

1 Introduction

Segmentation results obtained from grey-value images might be negatively affected by the presence of intensity changes due to shadows and surface curvature in those images. The goal of this project is to investigate to what extent segmentation results based on grey-value information can be improved by using color information. In order to express color information, different color systems are available. These color systems describe color by means of three parameters.
Publications on segmentation algorithms based on color information are not so very common. Sarabi and Aggarwal [11] describe an interactive system which segments color images. Huntsberger and Descalzi [5] present a color edge detector using fuzzy set theory to model inherent noise and correlation between color coordinates. Nevatia [7] extended the Hueckel operator for color edge extraction and Ohlander [8] employed color information for color image segmentation. Ohta [9,10] proposes a segmentation algorithm based on recursively thresholding of histograms of color features. Kender [6] discusses the problems of color transformations. Skotnicki [13] discusses color features and a 2D region growing algorithm. Schacter, Davis and Rosenfeld [12] used a 3D color clustering technique to segment images and Celenk [1] describes a color clustering algorithm, which operates in the 1976 *C.I.E.* ($L^*a^*b^*$) uniform color coordinate system.
In this paper, we first mention the problems encountered during the segmentation of grey-value images. Secondly, different color systems are discussed. Next, two segmentation methods are presented and the experiments which have been carried out with the two segmentation methods on a set of color images, will be discussed. Finally, several conclusions will be drawn.

2 Problems with Grey-value Segmentation

The following problems are encountered with the segmentation of grey-value images:

- **Shadows.**
  Segmentation methods partition an image array into low-level parts corresponding to groups of pixels with particular properties. These parts are homogeneous according to a certain homogeneity criterion. If a shadow falls on a homogeneous surface then a segmentation method based on grey- value information will most likely partition the surface into two parts, because the homogeneous surface might be considered as two surfaces with different grey-values. The correct segmentation result should be an entire region corresponding to the surface, ignoring shadows.
- **Surface Curvature.**
  Also segmentation results based only on grey-value information might be negatively affected by the presence of changes in intensity caused by surface curvature. It is also most likely that grey-value-based segmentation results will contain redundant regions caused by this phenomenon.
- **Metamerism.**
  A segmentation based on grey-value information might not detect the difference between two adjacent different colored objects, reflecting the same amount of light (i.e. having the same grey-values whereas being differently colored).

It may be expected that segmentation results of grey-value images can be improved by using color information. This can be seen as follows. The intensity of reflected light for any point on a surface, depends on the optical properties of the surface, the relative position of the surface and the orientation of the surface with respect to light sources.
In fact, the intensity of light reflected by a point on a matte surface (diffuse reflection) illuminated by ambient light and one point source is calculated by, [3]:

$$I_r = k_r I_a + (1-s)\frac{k_r I_s}{d+d_0}(N \cdot L)$$

$$I_g = k_g I_a + (1-s)\frac{k_g I_s}{d+d_0}(N \cdot L)$$

$$I_b = k_b I_a + (1-s)\frac{k_b I_s}{d+d_0}(N \cdot L)$$

Parameters k_r, k_g and k_b are the coefficients of reflection for respectively the red, green and the blue component for a point on the surface. I_a is the amount of ambient light falling on the point and I_s is the intensity of the source. Parameter d

is the distance from the point to the point source and d_0 is adjusted if light sources are larger than point sources. The unit vector N is the normal vector describing the orientation of the point on the surface and the direction to the light source is described by the unit vector L. Parameter s is the amount of shading falling on the point (i.e. a range from 0 to 1, where s=0 no shading and s=1 total shading).
This equation can be expanded if more than one point source is present. If we take this simple shading model (diffuse reflection) as starting point, the following arguments motivate our expectation:

Argument 1 Consider the color ratio's $I_r/(I_r + I_g + I_b)$, $I_g/(I_r + I_g + I_b)$ and $I_b/(I_r + I_g + I_b)$. These color ratio's are not sensitive to intensity changes due to surface curvature and shadows, because:

$$\frac{I_r}{I_r + I_g + I_b} =$$

$$\frac{k_r(I_a + (1-s)\frac{I_s}{d+d_0}(N \cdot L))}{(k_r + k_g + k_b)(I_a + (1-s)\frac{I_s}{d+d_0}(N \cdot L))} =$$

$$\frac{k_r}{k_r + k_g + k_b}$$

Equal argument holds for green and blue.
Hence, intensity information is factored out of the system (i.e. normalized with respect to intensity).

Argument 2 The main cause of intensity changes are due to the curvature of the surface. Hence, $cos\theta = N \cdot L$ is a very important aspect, where the angle of incidence θ is the angle between the unit normal vector N of the surface and L describing the direction to the light source. If neighbor pixels have intensity differences caused by variations of θ then their values exhibit a continuous progress. Therefore information is contained in the discontinuities.

Argument 3 The metamerism problem will be solved by including color information.

A segmentation result which is unaffected by the presence of shadows and changes in intensity due to surface curvature in the image , is considered as a good segmentation result as opposed to a segmentation result with split regions caused by the presence of these phenomena. In order to give a more formal description, the *spatial color progress coherence* concept is introduced. A set of connected pixels is $SCPC$ if the pixel values exhibit a color distribution, which can be interpreted as being originated from a homogeneous colored surface contaminated by shadows and changes in intensity due to surface curvature.
We mentioned that the intensity of light reflected by a point on a matte surface can be calculated. Let the value of pixel $\vec{a}_p$ correspond to the intensity of reflected light, measured by a color camera, from a point p on a physical surface. Now we can give a definition of the $SCPC$ concept.
A set S of pixels is $SCPC$ if:

$$\forall \vec{a}_p \in S \ \vec{a}_p = \begin{pmatrix} k_r I_a + (1-s_p)\frac{k_r I_s}{d_p+d_0}(N_p \cdot L) \\ k_g I_a + (1-s_p)\frac{k_g I_s}{d_p+d_0}(N_p \cdot L) \\ k_b I_a + (1-s_p)\frac{k_b I_s}{d_p+d_0}(N_p \cdot L) \end{pmatrix}$$

Hence, a set S of connected pixels is $SCPC$ if they are originated from the same homogeneous colored surface.
Let I denoting the image and let $SCPC$ be a logical predicate defined over a set of contiguous pixels. Then a good segmentation is defined as a partition of I into disjoint regions $I1, I2, \ldots, Ik$ such that

(i) $\bigcup_{j=1}^{k} I_j = I$;

(ii) I_j, $j = 1, 2, \ldots, k$ is connected

(iii) $SCPC(I_j) =$ TRUE for $j = 1, 2, \ldots, k$

(iv) $SCPC(I_i \vee I_j) =$ FALSE $\forall i \neq j$, where I_i and I_j are adjacent.

Condition (i) indicates that the segmentation must be complete. Condition (ii) requires that pixels in a region should be connected. The third condition (iii) indicates that each region satisfies the $SCPC$ criterion. Condition (iv) is that no adjacent regions may satisfy the $SCPC$ condition if they are grouped together.
In order to investigate to what extent segmentation results of grey-value images can be improved by using color information, different color systems are discussed now.

3 Color

Each of the color systems, which are presented in this paper, describes color by means of three color features. The color features , recorded by a color camera, are respectively R, G and B and are represented by the red, green and blue images. The RGB color system is based on the additive color mixture principle. This means that a range of colors (*color gamut*) is produced by the mixture of the three *primary colors* with corresponding intensities.

Because there exists no unique set of three primary colors which can describe all perceived colors, an international standard was established in 1931 by the *Commission International de l'Éclairage* (*C.I.E.*). The chosen standard primary colors are not real but imaginary colors, because they are too saturated to be seen by a human eye or produced. The advantage , however, is that those primary colors can describe each perceived color mathematically. These standard primary colors are denoted by X, Y, and Z, yielding the XYZ color system.

The XYZ color system has the additional advantage that intensity information (illumination) is determined only by its Y value. If we are only interested in the ratio of the standard primary colors the xyz color system is obtained. x, y, and z are called the *chromaticity coordinates* and are calculated by dividing respectively the X, Y, and Z coordinates by their total sum. It is obvious that the intensity information is factored out of the system, because the chromaticity coordinates reflect only the ratio of the three standard primary colors.

Since the sum of the chromaticity coordinates equals unity, two of these three quantities are sufficient to describe a color. When the x and y values are represented in a plane with the visible spectrum, the *C.I.E.* 1931 *chromaticity diagram* is obtained, see Fig. 1.

All colors can be represented by the diagram. The spectral colors are lying on the tongue-shaped curve according to their wavelengths.

By dividing the R, G, and B coordinates by their total sum, the r, g, and b quantities are obtained, yielding the rgb color system. This color system is normalized with respect to intensity, like the xyz color system.

The rgb color systems become unstable when intensity is small. Also color feature b is correlated with r and g. In order to overcome these problems, color feature b is discarded and intensity is incorporated in the color system, yielding the Irg color system.

Figure 1: The *C.I.E.* 1931 chromaticity diagram.

An other color system is the $I1I2I3$ color system proposed by Otha [9,10] and is based on the Karhunen-Loève transformation of eight different color images.

The IHS color system is non-linear, where I is intensity, H is the hue and S is saturation, Smith [14].

An important characteristic of a color system is that the system is visual uniform, where numerical distances could be related to human perceptual differences.

MacAdam proved the visual nonuniformity of the XYZ color system by doing experiments. These experiments are based on the *just noticeable difference* (JND) concept. The JND is defined as the minimal visual difference of two colors: let Q_0 be a color then color Q_1 is just noticeable different from Q_0, if there is no color Q_2, lying on the line from Q_0 through Q_1, which is closer to and noticeable different from Q_0. The results of the experiments made by MacAdam is that colors with coordinates (x_i, y_i) in the chromaticity diagram, which are not just noticeable different from color center (x_0, y_0), are lying on an ellipsis with center (x_0, y_0) and all colors which are (just) noticeable different from (x_0, y_0) are lying outside the ellipsis. The ellipses of an arbitrary number of colors were calculated and analyzed. MacAdam proved that the differences between the axis-diameters of the ellipses were rather (too) large and therefore nonuniformity of the XYZ color system can easily be seen.

The next two color systems are visual uniform because their color feature spaces have the property that the magnitude of the perceived color differ-

ence of two colors corresponds to the euclidean distance between the two colors in their color feature spaces. The two color systems are derived from the XYZ color system and are called the $U^*V^*W^*$ and the $L^*a^*b^*$ color system proposed by *C.I.E.*.
Other color systems, such as the $I_{ntsc}QY$, UVW and the Vuv, will not be examined, because they are similar to one of the above discussed color systems.
The properties of the different color systems are summarized in Fig. 2.

Figure 2: Color systems properties.

4 Segmentation Methods

4.1 Clustering Technique

Two methods have been implemented in order to examine color segmentation: a clustering and a region growing technique. The clustering technique clusters pixels on basis of color features, according to the k-means algorithm, [2]. The k-means method is based on the minimization of the so-called *performance index*, which is defined as 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 performance index is reduced to its minimum.
The determination of initial cluster means plays a crucial part, because the initial partition is obtained by assigning pixels to their closest cluster means and the better the initial partition is, the faster the k-means method will converge.
A more detailed version is as follows:

- **STEP 1:**

  $s = 1$;
  determine k initial cluster means $\vec{c}_1(s)$, $\vec{c}_2(s)$, ..., $\vec{c}_k(s)$, corresponding to k important patterns in the data;

- **STEP 2:**

  During the s^{th} iteration each pixel $\vec{a}$ is assigned to one of the k clusters $C_1(s)$, $C_2(s)$, ..., $C_k(s)$:

  $\vec{a}$ is assigned to cluster $C_j(s)$ if

  $$||\vec{a} - \vec{c}_j(s)|| < ||\vec{a} - \vec{c}_i(s)||,$$

  $$\forall i, j = 1, 2, ..., k \wedge i \neq j;$$

- **STEP 3:**

  Recomputation of k new cluster means:

  $$\vec{c}_j(s+1) = \frac{1}{N_j} \sum_{\vec{a} \in C_j(s)} \vec{a}, \quad j = 1, 2, ..., k;$$

- **STEP 4:**

  if $\exists j \in \{1, \ldots, k\}$ $\vec{c}_j(s+1) \neq \vec{c}_j(s)$
  then $s = s + 1$; **GOTO STEP 2**;
  else **STOP**.

$\vec{c}_j(s)$ indicates the cluster mean of the j^{th} cluster after s iterations, $C_j(s)$ denotes the j^{th} cluster after s iterations with cluster mean $\vec{c}_j(s)$, $\vec{a}$ is a pixel described by a n-dimensional feature vector and N_j is the number of pixels of cluster $C_j(s)$.

In order to determine the initial cluster means, a n-dimensional histogram is generated and k cluster means, corresponding to k conspicuous peaks in the histogram, are computed. The number of clusters must be known in advance.

Figure 3: Stages of the surface-based region growing method.

4.2 Region Growing Technique

The other approach is derived from the split& merge algorithm proposed by Horowitz and Pavlidis [4], and is shown in Fig. 3.
First, color feature images are computed from the original R, G and B values of the red, green, and blue images. During the second step the image, where each pixel is a n-dimensional color feature vector, will be partitioned into square regions of equal size, corresponding to the conversion of the image into the initialization level of the quadtree, see Fig. 4.

Figure 4: The structure of a quadtree.

The third step will merge four quadrants together into one quadrant when the resulting quadrant satisfies a certain homogeneity condition and one quadrant will be split into four quadrants if the quadrant does not satisfy a homogeneity condition. The fourth step in the processing sequence is necessary in order to group adjacent quadrants, because intrinsic homogeneous objects or object surfaces might not exactly fit within a quadrant, see Fig. 5.

Figure 5: The fitting problem.

During the 5th step small regions will be eliminated because there may be still a large number of small regions to be left. The final segmentation result will be computed in the 6th step, yielding a contour image where each contour corresponds to objects or object surface boundaries.

4.2.1 Homogeneity Criteria

Homogeneity criteria play an important role during the execution of the proposed region growing algorithm, because they determine whether a region is homogeneous or not.
Two homogeneity criteria have been implemented: a homogeneity criterion based on functional approximation and the mean&variance homogeneity criterion. The mean&variance homogeneity criterion assumes that the intensity of the pixels, discarding noise, over a region is constant. The hypothesis of the functional approximating criterion is that the intensity over a region may either be constant or variable due to intensity changes caused by shadows and surface curvature.

4.2.2 Homogeneity Criterion Based on Functional Approximation

The principle of the functional approximating criterion is that relevant information, such as edges,

in a region is detected when the data under consideration can't be adequately approximated by a set of approximating functions, and no relevant information, such as intensity changes due to shadows and surface curvature, is ignored when the data is adequately approximated by the set of approximating functions. We have chosen low-order bivariate polynomial approximating functions as the set of approximating functions, because these functions detect relatively well useful information (i.e. abrupt changes in the color features) and ignore misleading information (i.e. changes in intensity caused by shadows and surface curvature) when the order is not too high.
This set of low-order polynomials can be written as:

$$f_m(x,y) \quad = \quad \sum_{i+j \leq m} a_{ij} x^i y^j,\ m \leq 4$$

The planar polynomial is obtained by $m = 1$, the biquadratic polynomial by $m = 2$, etc.

Example 1. For m=2, the result is as follows:

$$\begin{aligned} f_2(x,y) \quad = \quad & a_{00} + a_{10}x + a_{01}y + \\ & a_{11}xy + a_{21}x^2 + a_{12}y^2 \end{aligned}$$

A least-square solver calculates the vector $\vec{a}$ and the fitting error ϵ for a region R.

$$\epsilon \quad = \quad \sqrt{\sum_{x,y \in R} \frac{(g(x,y) - f_m(x,y))^2}{N}}$$

Where N is the total of surface interior pixels contributing to the sum.
If the mean noise variance in the image can be estimated, the mean fitting error ϵ can be compared to this. If ϵ equals the noise, region R can be adequately approximated and if not, useful information (i.e. an edge or boundary is present) is detected. An estimate of the mean noise variance can be obtained by a local 3x3 planar fit, and not including pixels with high slopes in the calculation. Thus, pixels with gradients above a certain threshold will not be considered in the calculation of the noise.

- So the mean noise variance σ^2 is estimated over a 3x3 neighborhood by:

$$\sigma^2 \quad = \quad 1/9 \sum (g(x,y) - (a_{00} + a_{10} + a_{01}))^2$$

4.2.3 The Mean&Variance Homogeneity Criterion

The mean&variance homogeneity criterion is based on the mean and variance of a particular region R, which is the case for $m = 0$. Hence, the hypothesis of the mean&variance criterion is that the pixels over the region R have constant values as opposed to the functional approximating criterion.

$$f_0(x,y) \quad = \quad a_{00}$$

This is equal to the mean pixel value in region R.

$$\epsilon \quad = \quad \sqrt{\sum_{x,y \in R} \frac{(g(x,y) - a_{00})^2}{N}}$$

5 Experiments

Experiments have been carried out with the two segmentation methods on a set of three color images in order to investigate the results of including color. The definition of a good segmentation result was discussed in Section 2. To measure the quality of a segmentation result, an ideal segmentation , called the *witness* of the image, is given first. The witness should satisfy the definition of the earlier defined $SCPC$ segmentation. The composition of the regions in the witness is made by an expert.
The quality of the segmentation results is obtained by comparing the results to their witnesses. As criterion the number of split regions is used. The more split regions a segmentation result contains as opposed to its witness the worser the result. The evaluation is done by a human observer. The overview of segmentation results of the k-means algorithm is shown in Fig. 6. and those of the split&merge in Fig. 7.
The Rod, Disc and the Manhat images are shown in Fig. 8.
The Rod image is a rod with seven different colors along it, placed on a dark background. Segmentation problems can be caused by changes in brightness due to the curvature of the rod, and small intensity values of the background. The Disc image consists of three discs. Segmentation problems can be caused by the brightness differences on the background and the unclear part of the demarcation of the outer and the middle disc. The Manhat image is composed of objects on a surface. The image is clearly contaminated

Figure 6: Overview k-means results.

Figure 7: Overview split&merge results.

by intensity changes due to shadows and surface curvature. The witnesses are shown in Fig. 9.
Only the most characteristic segmentation results will be presented and are summarized in Fig. 6 and 7.
These general results are illustrated below.
Fig. 10 shows that segmentations based on color information are better than those based on grey-value information.
Non-linear color systems become unstable when intensity is small as was discussed before (Fig. 11).
The normalized color systems rgb and xyz provide the best segmentation results if the image contains intensity changes (Fig. 12).
Linear combinations perform the same, as could be expected from the homogeneity criteria used (Fig. 13).
Results of the split&merge algorithm are better than those of the k-means algorithm (Fig. 14).
The mean&variance criterion assumes a constant value over the regions, which might not be the case for grey-value images and color systems containing intensity information (e.g. RGB, $I1 I2I3$ and IHS). Therefore, the functional approximation criterion, which assumes that the value of the regions may either be constant or variable, is more suited for grey-value images and color systems containing intensity information.
If the intensity is factored out of the system, such as rgb and xyz, the mean&variance criterion is sufficient (Fig. 15).

6 Conclusion

The goal of this research project, is to investigate to what extent color information can improve segmentation results based only on grey-value information. The influence of different color systems has been examined for two segmentation methods. Also the influence of the two segmentation methods has been investigated. The conclusions of the experiments are:

- Segmentation results based on color information are better than those based on grey-value information.
- If the intensity in the image is small, the rgb, IHS and the $L^*a^*b^*$ color systems provide poor segmentation results.
- Normalized color systems provide the best segmentation results if the image contains intensity changes due to surface curvature and shadows.
- The segmentation results of the RGB, $XY Z$ and the $I1I2I3$ color systems are similar, what could be expected for linear combinations of the same RGB basis.

Figure 8: 8.a The Rod. 8.b The Disc. 8.c The Manhat.

Figure 9: 9.a The Rod witness. 9.b The Disc witness. 9.c The Manhat witness.

- Segmentation results of the split & merge algorithm are better than those of the k-means algorithm.
- Using the mean&variance homogeneity criterion, segmentation results obtained from grey-value images and color systems containing intensity information, are somewhat more sensitive to changes in intensity as opposed to the functional approximation criterion.

Figure 10: 10.a Grey 10.b *RGB*.

Acknowledgements

We express our appreciation to Prof. A.W.M. Smeulders for his helpful comments and suggestions and we thank the Technical University of Delft for providing the color images.

References

[1] Celenk, M.,
A Color Clustering Technique for Image Segmentation,
Computer Vision, Graphics, and Image Processing, 52, 1990, pp. 145-170.

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

[3] Hearn, D., Baker, M. P.,
Computer Graphics,
Prentice-Hall, Englewood Cliffs, New Yersey, 1986.

[4] Horowitz, S. L., and Pavlidis, T.,
Picture Segmentation by a Directed Split-and-Merge Procedure,
Proceedings Second International Joint Conference on Pattern Recognition, Copenhagen, 1974, pp. 424-433.

Figure 11: 11.a rgb 11.b IHS

Figure 13: 13.a RGB 13.b XYZ 13.c $I1I2I3$

Figure 12: 12.a rgb 12.b Grey

Figure 14: 14.a K-means Grey 14.b Split&Merge Grey

[5] Huntsberger, T. L., and Descalzi, M. F., **Color Edge Detection**, Pattern Recognition Letters 3, 1985, pp. 205-209.

[6] Kender, J. R., **Saturation, Hue, and Normalized Colors: Calculation, Digitization Effects, and Use**, Technical Report, Department of Computer Science, Carnegie-Mellon University, 1976.

[7] Nevatia, R., **A Color Edge Detector**, Proc. IJCPR-III, 1976, pp. 829-832.

[8] Ohlander, R., **Analysis of Natural Scenes**, Ph.D. Thesis, Department of Computer Science, Carnegie-Mellon University, 1975.

[9] Ohta, Y., **Knowledge-Based Interpretation of Outdoor Natural Scenes**, Pitman Publishing, London, 1985.

[10] Ohta, Y., Kanade, T., and Sakai, T., **Color Information for Region Segmentation**, Computer Graphics and Image Processing 13, 1980, pp. 222-241.

[11] Sarabi, A., and Aggarwal, J. K., **Segmentation of Chromatic Images**, Pattern Recognition, Vol. 13, 1981, pp. 417-427.

[12] Schacter, B. J., Davis, L. S. and Rosenfeld, A., **Scene Segmentation by Cluster Detection in Color Spaces**, SIGART Newsl., No. 58, 1976, pp. 16-17.

Figure 15:

[13] Skotnicki, P. J.,
Segmentation of Color Images,
Signal and System Technique Group, Pattern Recognition Section, Delft University of Technology, Faculty of Applied Physics, Delft, 1989, pp. 1-81.

[14] Smith, A. R.,
Color Gamut Transform Pairs,
Computer Graphics 12, No. 3, 1978, pp. 12-19.