PERGAMON

Pattern Recognition 34 (2001) 2259–2281

PATTERN RECOGNITION
THE JOURNAL OF THE PATTERN RECOGNITION SOCIETY

www.elsevier.com/locate/patcog

Color image segmentation: advances and prospects

H.D. Cheng*, X.H. Jiang, Y. Sun, Jingli Wang

Department of Computer Science, Utah State University, Logan, UT 84322-4205, USA

Received 23 February 2000; accepted 12 September 2000

Abstract

Image segmentation is very essential and critical to image processing and pattern recognition. This survey provides a summary of color image segmentation techniques available now. Basically, color segmentation approaches are based on monochrome segmentation approaches operating in different color spaces. Therefore, we first discuss the major segmentation approaches for segmenting monochrome images: histogram thresholding, characteristic feature clustering, edge detection, region-based methods, fuzzy techniques, neural networks, etc.; then review some major color representation methods and their advantages/disadvantages; finally summarize the color image segmentation techniques using different color representations. The usage of color models for image segmentation is also discussed. Some novel approaches such as fuzzy method and physics-based method are investigated as well.

Keywords: Color image segmentation; Color representations; Color space transformations; Neural networks; Thresholding; Clustering; Edge detection; Region-based approach; Physics based approach; Fuzzy logic

1. Introduction

Image segmentation is the first step in image analysis and pattern recognition. It is a critical and essential component of image analysis and/or pattern recognition system, is one of the most difficult tasks in image processing, and determines the quality of the final result of analysis. Image segmentation is a process of dividing an image into different regions such that each region is, but the union of any two adjacent regions is not, homogeneous. A formal definition of image segmentation is as follows [1]: If $P(\,)$ is a homogeneity predicate defined on groups of connected pixels, then segmentation is a partition of the set F into connected subsets or regions $(S_1, S_2, \ldots, S_n)$ such that

$$\bigcup_{i=1}^{n} S_i = F \quad \text{with } S_i \cap S_j = \Phi \quad (i \neq j).$$

The uniformity predicate $P(S_i) = \text{true}$ for all regions, S_i, and $P(S_i \cup S_j) = \text{false}$, when $i \neq j$ and S_i and S_j are neighbors.

According to [2], "the image segmentation problem is basically one of psychophysical perception, and therefore not susceptible to a purely analytical solution". There are many papers and several surveys on monochrome image segmentation techniques. Color image segmentation attracts more and more attention mainly due to the following reasons: (1) color images can provide more information than gray level images; (2) the power of personal computers is increasing rapidly, and PCs can be used to process color images now. The segmentation techniques for monochrome images can be extended to segment color images by using R, G and B or their transformations (linear/non-linear). However, comprehensive surveys on color image segmentation are few. Ref. [3] analyzed the problem when applying edge-based

* Corresponding author. Tel.: + 1-435-797-2054; fax: + 1-435-797-3265.

E-mail addresses: cheng@hengda.cs.usu.edu (H.D. Cheng), xihua.jiang@citrix.com (X.H. Jiang), angela.sun@motorola.com (Y. Sun), jingliwang@cc.usu.edu (J. Wang).

0031-3203/01/$20.00
PII: S0031-3203(00)00149-7

and region-based segmentation techniques to color images with complex texture, and Ref. [4] discussed the properties of several color representations, the segmentation methods and color spaces.

This paper provides a summary of color image segmentation techniques available at present, and describes the properties of different kinds of color representation methods and problems encountered when applying the color models to image segmentation. Some novel approaches such as fuzzy and physics-based approaches will be discussed as well.

Section 2 briefly introduces the major segmentation approaches for processing monochrome images: histogram thresholding, characteristic feature clustering, edge detection, region-based methods, fuzzy techniques and neural networks. Section 3 reviews some major color representations and their advantages/disadvantages. Section 4 investigates the segmentation techniques applied to color images using different color representations, and the summary is given in Section 5.

2. Monochrome image segmentation

Monochrome image segmentation approaches are based on either discontinuity and/or homogeneity of gray level values in a region. The approach based on discontinuity tends to partition an image by detecting isolated points, lines and edges according to abrupt changes in gray levels. The approaches based on homogeneity include thresholding, clustering, region growing, and region splitting and merging. Several survey papers on monochrome image segmentation [1,2,5–8] cover the major image segmentation techniques available.

Ref. [2] discussed segmentation from the view point of cytology image processing. It categorized various existing segmentation techniques before 1980s into three classes: (1) characteristic feature thresholding or clustering, (2) edge detection, and (3) region extraction. The segmentation techniques were summarized and comments were given on the advantages and disadvantages of each approach. The threshold selection schemes based on gray level histogram and local properties, and based on structural (textural) and syntactic techniques were described. Clustering techniques were regarded as "the multidimensional extension of the concept of thresholding". Some clustering schemes utilizing different kinds of features (multi-spectral information, mean/variation of gray level, texture, color) were discussed. Various edge detection techniques were presented, which were categorized into two classes: parallel and sequential techniques. The parallel edge detection technique implies that the decision of whether a set of points are on an edge or not is dependent on the gray level of the set and some set of its neighbors, which includes high-emphasis spatial frequency filtering, gradient operators, adaptive local operator [9,10], functional approximations [11,12], heuristic search and dynamic programming, relaxation, and line and curve fitting, while the sequential techniques make decision based on the results of the previously examined points. A brief description of the major component of a sequential edge detection procedure was given. It also briefly introduced region merging, region splitting and combination of region merging and splitting approaches.

Ref. [5] classified image segmentation techniques into six major groups: (1) measurement space guided spatial clustering, (2) single linkage region growing schemes, (3) hybrid linkage region growing schemes, (4) centroid linkage region growing schemes, (5) spatial clustering schemes, and (6) split-and-merge schemes. These groups are compared on the problem of region merge error, blocky region boundary and memory usage. The hybrid linkage region growing schemes appear to offer the best compromise between having smooth boundaries and few unwanted region merges. One of the drawbacks of feature space clustering is that the cluster analysis does not utilize any spatial information. The survey presented some spatial clustering approaches which combine clustering in feature space with region growing or spatial linkage techniques. It gives a good summary of kinds of linkage region growing schemes. The problem of high correlation and spatial redundancy of multi-band image histograms and the difficulty of clustering using multi-dimensional histograms are also discussed.

Ref. [6] surveyed segmentation algorithms based on thresholding and attempted to evaluate the performance of some thresholding techniques using uniformity and shape measures. It categorized global thresholding techniques into two classes: point-dependent techniques (gray level histogram based) and region-dependent techniques (modified histogram or co-occurrence based). Discussion on probabilistic relaxation and several methods of multi-thresholding techniques was also given.

Ref. [7] regarded image segmentation as the bridge in a machine vision system between a low-level vision subsystem including image processing operations (such as noise reduction, edge extraction) to enhance the input image, and a high-level vision subsystem including object recognition and scene interpretation. After segmentation, the enhanced input image is mapped into a description involving regions with some common features for the high-level vision tasks. The segmentation techniques are categorized into three main classes: pixel-based, edge-based, and region-based schemes. Some common image segmentation approaches are studied, such as Gaussian filtering, Otsu's thresholding method, Chow–Kaneko's adaptive thresholding, Yanowitz and Bruckstein's adaptive edge-based thresholding, Parker's local intensity gradient approach, and Horowitz and Pavlidis's split, merge, and group (SMG) approach [7]. Ref. [1] reviewed gray level thresholding, edge detection and many other approaches such as fuzzy set segmentation approaches, neural network based approaches, Markov Random Field (MRF) based approaches, and surface based approaches. It also considers range image and magnetic resonance image (MRI) besides the most common light intensity image. A brief introduction to color image segmentation and fuzzy segmentation approach is discussed. The development based on the applications of fuzzy operators, properties and mathematics are presented, and segmentation based on IF–THEN rules is predicted as a promising research area in the near future. After the discussion of these segmentation approaches, the authors made a comparison of six histogram based methods and two iterative pixel classification methods: relaxation and MAP (maximum a posteriori) estimation. Finally, the attempts for objective evaluation of segmentation results are studied.

For detail descriptions of various monochrome image segmentation approaches, readers may refer to Refs. [1,2,5–7].

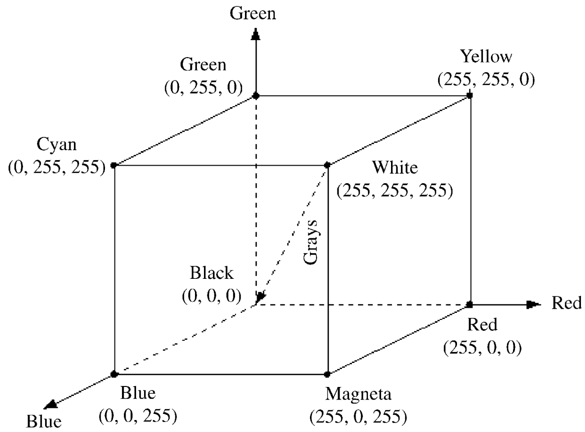

Fig. 1. *RGB* color space represented in a 3-dimensional cube [14].

3. Color features

Color is perceived by humans as a combination of tristimuli R (red), G (green), and B (blue) which are usually called three primary colors. From R, G, B representation, we can derive other kinds of color representations (spaces) by using either linear or nonlinear transformations. Several color spaces, such as RGB, HSI, $CIE\ L{*}u{*}v{*}$ are utilized in color image segmentation, but none of them can dominate the others for all kinds of color images. Selecting the best color space still is one of the difficulties in color image segmentation [13].

Red, green, and blue components can be represented by the brightness values of the scene obtained through three separate filters (red, green, and blue filters) based on the following equations:

$$R = \int_{\lambda} E(\lambda) S_R(\lambda)\,\mathrm{d}\lambda,$$

$$G = \int_{\lambda} E(\lambda) S_G(\lambda)\,\mathrm{d}\lambda,$$

$$B = \int_{\lambda} E(\lambda) S_B(\lambda)\,\mathrm{d}\lambda,$$

where S_R, S_G, S_B are the color filters on the incoming light or radiance $E(\lambda)$, and λ is the wavelength.

The RGB color space can be geometrically represented in a 3-dimensional cube (Fig. 1) [14]. The coordinates of each point inside the cube represent the values of red, green and blue components, respectively.

The laws of colorimetry are [15]: (1) any color can be created by these three colors and the combination of the three colors is unique; (2) if two colors are equivalent, they will be again equivalent after multiplying or dividing the three components by the same number; (3) the luminance of a mixture of colors is equal to the sum of the luminance of each color. The tristimulus values that served as the color basis are: 425.8 nm for blue, 546.1 nm for green, and 700.0 nm for red. Any color can be expressed by these three color bases.

RGB is the most commonly used model for the television system and pictures acquired by digital cameras. Video monitors display color images by modulating the intensity of the three primary colors (red, green, and blue) at each pixel of the image [16,17]. RGB is suitable for color display, but not good for color scene segmentation and analysis because of the high correlation among the R, G, and B components [18,19]. By high correlation, we mean that if the intensity changes, all the three components will change accordingly. Also, the measurement of a color in RGB space does not represent color differences in a uniform scale, hence, it is impossible to evaluate the similarity of two colors from their distance in RGB space.

3.1. Linear transformations

3.1.1. YIQ

YIQ is used to encode color information in TV signal for American system. It is obtained from the RGB model by a linear transformation:

$$\begin{pmatrix} Y \\ I \\ Q \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.274 & -0.322 \\ 0.211 & -0.253 & -0.312 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

where $0 \leqslant R \leqslant 1$, $0 \leqslant G \leqslant 1$, $0 \leqslant B \leqslant 1$.

The Y component, is a measure of the luminance of the color, and is a likely candidate for edge detection in a color image. The I and Q components jointly describe the hue and saturation of the image [20]. The YIQ space can partly get rid of the correlation of the red, green and blue components in an image. The linear transformation needs less computation time than nonlineal ones, which makes the YIQ space more preferable to nonlinear systems.

3.1.2. YUV

YUV is also a kind of TV color representation suitable for European TV system. The transformation is:

$$\begin{pmatrix} Y \\ U \\ V \end{pmatrix} = \begin{pmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

where $0 \leqslant R \leqslant 1$, $0 \leqslant G \leqslant 1$, $0 \leqslant B \leqslant 1$.

3.1.3. $I_1 I_2 I_3$

Ref. [21] performed systematic experiments of region segmentation to derive a set of effective color features. At each step of the recursive region splitting, new color features are calculated by Karhunen–Loeve transformation of R, G, and B. It applied eight kinds of color pictures, analyzed over 100 color features, and found a set of effective color features as follows:

$$I_1 = (R + G + B)/3,$$

$$I_2 = (R - B)/2,$$

$$I_3 = (2G - R - B)/4.$$

Comparing $I_1 I_2 I_3$ with 7 other color spaces ($RGB, YIQ, HSI, Nrgb, CIE, XYZ, CIE(L^*u^*v^*)$, and $CIE(L^*a^*b^*)$), Ref. [21] claimed that $I_1 I_2 I_3$ was more effective in terms of the quality of segmentation and the computational complexity of the transformation.

3.2. Nonlinear transformations

3.2.1. Normalized $RGB(Nrgb)$

When performing color image segmentation, we need to make colors not dependent on the change in lighting intensities. An efficient method is to get the variations of intensities uniformly across the spectral distribution. The normalized color space is formulated as:

$$r = R/(R + G + B),$$

$$g = G/(R + G + B),$$

$$b = B/(R + G + B).$$

Since $r + g + b = 1$, when two components are given, the third component can be determined. We may use only two out of three [22].

Ref. [23] created another normalized color representation space which was defined as

$$Y = c_1 R + c_2 G + c_3 B,$$

$$T_1 = \frac{R}{R + G + B},$$

$$T_2 = \frac{G}{R + G + B}.$$

where $c_1 + c_2 + c_3 = 1$, c_1, c_2, and c_3 are constants and can be combined to produce the illumination of the image pixel. We can see that T_1 and T_2 are decided only by the percentage of the RGB components, therefore, they can represent the real color information of an image, and they are independent of the brightness of the image which is one of the advantages of the normalized RGB color system [24]. It is more convenient to represent the color plane since the color values are set to a narrow limitation.

Normalization reduces the sensitivity of the distribution to the color variability [25]. It is relatively robust to the change of the illumination. But an obvious shortcoming of normalized RGB is that the normalized colors are very noisy if they are under low intensities. This is due to the nonlinear transformation from the RGB space to the normalized RGB space.

3.2.2. HSI

The HSI (hue–saturation–intensity) system is another commonly used color space in image processing, which is more intuitive to human vision [26–29]. There are some variants of HSI systems, such as HSB (hue–saturation–brightness), HSL (hue–saturation–lightness), and HSV (hue–saturation–value) [30–32].

The HSI system separates color information of an image from its intensity information. Color information is represented by hue and saturation values, while intensity, which describes the brightness of an image, is determined by the amount of the light. Hue represents basic colors, and is determined by the dominant wavelength in the spectral distribution of light wavelengths. It is the location of the peak in the spectral distribution. The saturation is a measure of the purity of the color, and signifies the amount of white light mixed with the hue. It is the height of the peak relative to the entire spectral distribution.

The HSI color space can be described geometrically as in Fig. 2 [14]. Generally, hue is considered as an angle between a reference line and the color point in RGB space. The range of the hue value is from 0° to 360°, for example, blue is 240°, yellow is 60°, green is 120°, and magenta is 300°. The saturation component represents

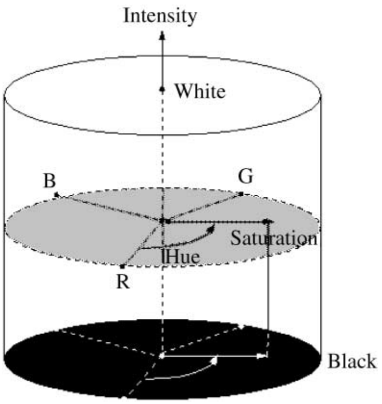

Fig. 2. *HSI* color space [14].

the radial distance from the cylinder center. The nearer the point is to the center, the lighter is the color. Intensity is the height in the axis direction. The axis of the cylinder describes the gray levels, for instance, zero (minimum) intensity is black, full (maximum) intensity is white. Each slice of the cylinder perpendicular to the intensity axis is a plane with the same intensity.

The *HSI* color system has a good capability of representing the colors of human perception, because human vision system can distinguish different hues easily, whereas the perception of different intensity or saturation does not imply the recognition of different colors.

The *HSI* coordinates can be transformed from the *RGB* space. The formulas for hue, saturation, and intensity are:

$$H = \arctan\left(\frac{\sqrt{3}(G-B)}{(R-G)+(R-B)}\right),$$

$$Int = \frac{(R+G+B)}{3},$$

$$Sat = 1 - \frac{\min(R,G,B)}{I}.$$

The hue is undefined if the saturation is zero, and the saturation is undefined when the intensity is zero.

We may use gray-level algorithms to operate on the intensity component of *HSI* description. To segment objects with different colors, we may apply the segmentation algorithms to the hue component only. For example, we may set thresholds on the range of hues that separate different objects easily, but it is difficult to transform these thresholds into *RGB* values, since hue, saturation, and intensity values are all encoded into *RGB* values. It is especially efficient when the images have non-uniform illumination such as shade, since hue is independent on intensity values. We may also efficiently apply thresholds to hue, saturation, and intensity components, respectively, to form some regions that can be fit for various region growing algorithms. Hue is particularly useful in the cases where the illumination level varies from point-to-point or image-to-image. If the integrated white condition holds, hue is invariant to certain types of highlights, shading, and shadows. Besides, segmentation in the 1-D hue space is computationally less expensive than in the 3-D *RGB* space. One of the disadvantages of hue is that it has a nonremovable singularity near the axis of the color cylinder, where a slight change of input R, G, and B values can cause a large jump in the transformed values. The singularities may create discontinuities in the representation of colors [15]. The hue values near the singularity are numerically unstable. That is why pixels having low saturation are left unassigned to any regions in many segmentation algorithms. Also, if the intensity of the color lies close to white or black, hue and saturation play little role in distinguishing colors.

3.2.3. *CIE spaces*

CIE (Commission International de l'Eclairage) color system was developed to represent perceptual uniformity, and thus meets the psychophysical need for a human observer. It has three primaries denoted as X, Y, and Z. Any color can be specified by the combination of X, Y and Z. The values of X, Y, and Z can be computered by a linear transformation from *RGB* tristimulus coordinates. In particular, the transformation matrix for the NTSC (National Television System Commission, United States) receiver primary system is:

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.607 & 0.174 & 0.200 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.116 \end{pmatrix} \begin{pmatrix} R \\ G \\ B \end{pmatrix}.$$

There are a number of *CIE* spaces can be created once the *XYZ* tristimulus coordinates are known. *CIE* ($L^*a^*b^*$) space and *CIE* ($L^*u^*v^*$) space are two typical examples. They can all be obtained through nonlinear transformations of X, Y, and Z values.

CIE ($L^*a^*b^*$) is defined as:

$$L^* = 116\left(\sqrt[3]{\frac{Y}{Y_0}}\right) - 16,$$

$$a^* = 500\left[\sqrt[3]{\frac{X}{X_0}} - \sqrt[3]{\frac{Y}{Y_0}}\right],$$

$$b^* = 200\left[\sqrt[3]{\frac{Y}{Y_0}} - \sqrt[3]{\frac{Z}{Z_0}}\right].$$

where frac $YY_0 > 0.01$, $X/X_0 > 0.01$, and $Z/Z_0 > 0.01$. (X_0, Y_0, Z_0) are X, Y, Z values for the standard white.

CIE ($L^*u^*v^*$) is defined as:

$$L^* = 116\sqrt[3]{\frac{Y}{Y_0}} - 16, \qquad u^* = 13L^*(u' - u_0),$$

$$v^* = 13L^*(v' - v_0).$$

where $Y/Y_0 > 0.01$, Y_0, u_0, and v_0 are the values for the standard white, and

$$u' = \frac{4X}{X + 15Y + 3Z}, \qquad v' = \frac{6Y}{X + 15Y + 3Z}.$$

Each point in these two color spaces can be regarded as a point in the (L^*, a^*, b^*) or (L^*, u^*, v^*) three-dimensional color space, so that the difference of two colors can be calculated as the Euclidean distance between two color points. The formulas for the color difference are described below:

For CIE $(L^*a^*b^*)$ space:

$$\Delta E_{ab} = \sqrt{(\Delta L^*)^2 + (\Delta a^*)^2 + (\Delta b^*)^2}.$$

For CIE $(L^*u^*v^*)$ space:

$$\Delta E_{uv} = \sqrt{(\Delta L^*)^2 + (\Delta u^*)^2 + (\Delta v^*)^2}.$$

The ability to express color difference of human perception by Euclidean distance is very important to color segmentation. (L^*, a^*, b^*) or (L^*, u^*, v^*) is approximately uniform chromaticity scale. That is, they match the sensitivity of human eyes with computer processing [33], whereas the RGB or XYZ color space does not have such a property. Therefore, we can derive the perceptual color attributes such as intensity, hue and saturation conveniently. We may use one of the two CIE color spaces and the associated color difference formulas to define HSI space which is mapped to the cylindrical coordinates of (L^*, u^*, v^*) or (L^*, a^*, b^*) space, and also consistent to the definition of HSI space.

For CIE $(L^*a^*b^*)$ space:

$$I = L^*,$$

$$H = \arctan(a^*/b^*),$$

$$S = \sqrt{(a^*)^2 + (b^*)^2}.$$

For CIE $(L^*u^*v^*)$ space:

$$I = L^*,$$

$$H = \arctan(u^*/v^*),$$

$$S = \sqrt{(u^*)^2 + (v^*)^2}.$$

The two CIE spaces share same L^* value, which defines the lightness, or the intensity, of a color.

CIE spaces can control color and intensity information more independently and simply than RGB primary colors. Direct color comparison can be performed based on geometric separation within the color space, therefore, it is especially efficient in the measurement of small color difference. However, it still has the same problem of singularity as other nonlinear transformations.

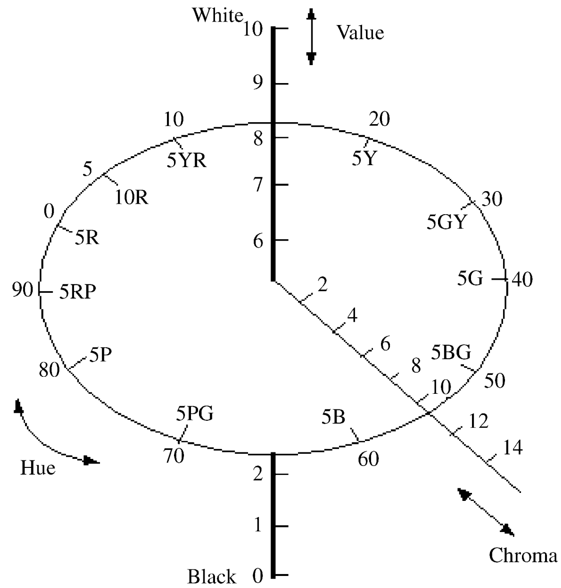

Fig. 3. Munsell color space [34].

3.2.4. Munsell

Munsell color system was created in 1969 [34]. It is one of the earliest methods to specify colors. The Munsell color system uses three attributes of color perception, which are Munsell Hue, Value, and Chroma. Fig. 3 describes the Munsell color space.

There are five colors regarded as the major hues, they are red (R), yellow (Y), green (G), blue (B) and purple (P). The combinations of colors: YR, GY, BG, PB, and RP are five half-way hues. The hue circle is divided equally into 100 parts with the five major hues and the five half-way hues located on the circle with equal spaces. All hue values are arranged on the hue circle, and can be described in two ways. One is to indicate their relative associations with a specified hue, for example, $R, 2R$, and $10R$. The other way is to assign a number to a hue: $5R = 0$, $6R = 1$, and increase the number clockwise along the hue circle to get a number for each hue.

Munsell value (V) describes the lightness of a color. It defines the value of black as 0 and the value of white as 10. An equation describes the relationship between V and Luminance Y:

$$Y = 1.2219V - 0.23111V^2 + 0.23951V^3 - 0.021009V^4 + 0.0008404V^5,$$

Chroma (C) is like the saturation component in the CIE representation. It describes the purity of a color. When C is equal to zero, it is an achromatic color.

Similar to the HSI space, the Munsell color space can also be represented by a cylindrical coordinate system consisting of H, V and C. However, we cannot convert the Munsell color system into CIE standard descriptions by a formula. Therefore, a method must be created to map the real color signals into the Munsell space [34]. Three levels of image representation were demonstrated in the experiment. The lowest level, density, can be achieved by Black/White component of the images; the next level, three primary colors, can be implemented on RGB images; and the highest level, three perceptual attributes can be confirmed with the images of Munsell HVC system.

3.3. Hybrid color space

Ref. [35] proposed a method to classify the pixels in Hybrid Color Space (HCS) which is composed by a set of three color features. Taking into account the K available color features ($R, G, B, R_n, G_n, B_n, I, H, S, X, Y, Z, X_n, Y_n$, Z_n, L^*, a^*, b^*, u^*, v^*, I_2, I_3, A, C_1, C_2, Y', I', Q',U', V', C^*_{uv}, h^0_{uv}, S^*_{uv}, $C^*_{a,b}$, h^0_{ab}). It used a specific informational criterion to select a set of three most discriminating color features. Experimental results show that: I_3, A and C^*_{uv} color features are selected. Finally, the classified LUV and HCS images are compared. This method depends on the application. It does not give the optimal solution but consumes less computational time.

3.4. Physics-based models

The traditional methods of color image segmentation suffer from too many erroneous regions because they have not accounted for the influence of optical effects on object colors. Image segmentation should base on material surface changes that include the material changes and variation due to shading, shadows and highlights. Besides objects receive the lights directly from illumination sources, they reflect lights from other objects. If the physical models were included at the segmentation stage, many of these regions could be classified correctly.

Physical models in color image processing are quite different from the conventional color information representations. These models aim at eliminating the effect of highlights, shadows and shading, and segmenting a color image at boundaries of objects. Shafer's "dichromatic reflection model" [36] and Healey's "approximate color-reflectance model (ACRM)" [37] are two typical examples.

Reflection is highly related to the nature of the materials. Ref. [37] divided materials into different classes: one includes optically homogeneous materials like metals, glass and crystals, and another includes the optically inhomogeneous materials such as plastics, paper, textiles and paints. Usually it is very helpful to identify or classify the material in the scene of an image before an algorithm is applied. For example, we should distinguish metals from dielectrics since they interact with lights in different ways. Thus, they require different algorithms in image understanding.

Based upon a dichromatic reflection model, Ref. [36] used a method to determine the amount of interface reflection and body reflection in a color image pixel by pixel. It also presented an algorithm for analyzing color values in an image, which is very useful in color image understanding.

3.5. Discussion on color space

Ref. [38] showed that nonlinear color transformations such as HSI and the normalized color space had essential, nonremovable singularities and there were spurious modes in the distribution of values. Ref. [38] suggested that linear spaces, such as YIQ, be used, rather than nonlinear spaces.

The major problem of linear color spaces is the high correlation of the three components, which makes the three components dependent upon each other and associate strongly with intensity. Hence, linear spaces are very difficult to discriminate highlights, shadows and shading in color images. Besides, if a linear color space is used, image segmentation has to be performed in a 3-D space, usually on one component at a time, because it is difficult to combine the information inherent in these components. However, nonlinear color spaces do not have such problems. In HSI space, hue can be used for segmentation in 1-D space if the saturation is not low, where certain types of highlights, shadows and shading can be discounted [30].

4. Color image segmentation

It has long been recognized that human eye can discern thousands of color shades and intensities but only two-dozen shades of gray. It is quite often when the objects cannot be extracted using gray scale but can be extracted using color information. Compared to gray scale, color provides information in addition to intensity. Color is useful or even necessary for pattern recognition and computer vision. Also the acquisition and processing hardwares for color images have become more available and accessible to deal with the computational complexity caused by the high-dimensional color space. Hence, color image processing has become increasingly more practical.

As mentioned before, the literature on color image segmentation is not as extensively present as that on monochrome image segmentation. Most published results of color image segmentation are based on gray level image segmentation approaches with different color representations, as shown in Fig. 4.

Ref. [1] gives a brief introduction to color image segmentation, and mentions that color images can be considered as a special case of multi-spectral images and any segmentation method for multi-spectral images can be

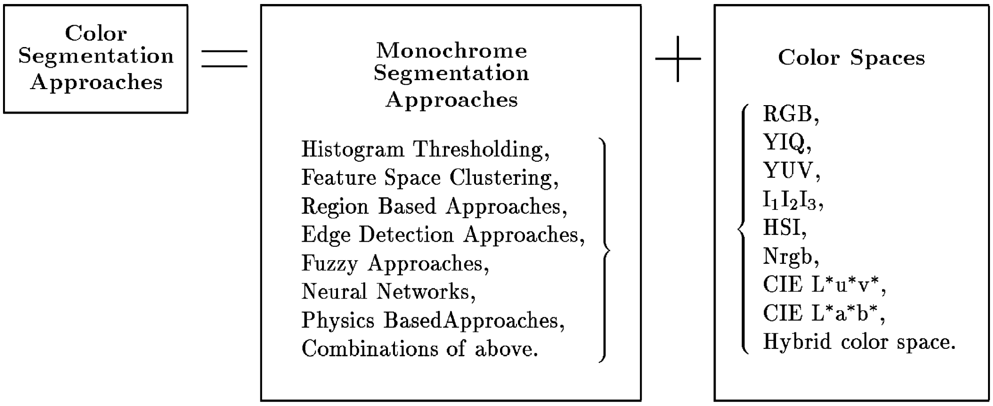

Fig. 4. Commonly used color image segmentation approaches.

applied to color images. Ref. [3] analyzed the complexities encountered in segmenting color images with complex texture. Only two color spaces, RGB and HSI, are discussed. Texture is considered to be the major problem for all segmentation techniques, thus much more discussion was made on texture analysis than on color representation, and the problems of feature extraction in images with textural variations are discussed particularly. The paper focuses on two major groups of segmentation techniques: boundary formation and region formation. The problem of how and to what degree the semantic information should be employed in image segmentation is discussed briefly. Experimental results are also presented. Ref. [4] discussed several color spaces (RGB, normalized RGB space, $HSI, CIE\ L^*u^*v^*$, YIQ, YUV) and their properties. Segmentation approaches are categorized into four classes: pixel based segmentation, area based segmentation, edge based segmentation and physics based segmentation. A brief conclusion is drawn based on the analysis of the literature available.

Most gray level image segmentation techniques can be extended to color images, such as histogram thresholding, clustering, region growing, edge detection, fuzzy approaches and neural networks. Gray level segmentation methods can be directly applied to each component of a color space, then the results can be combined in some way to obtain a final segmentation result [39]. However, one of the problems is how to employ the color information as a whole for each pixel. When the color is projected onto three components, the color information is so scattered that the color image becomes simply a multispectral image and the color information that humans can perceive is lost. Another problem is how to choose the color representation for segmentation. As discussed in the previous section, each color representation has its advantages and disadvantages. There is no single color representation that can surpasses others for segmenting all kinds of color images.

In most of the existing color image segmentation approaches, the definition of a region is based on similarity of color. This assumption often makes it difficult for any algorithms to separate the objects with highlights, shadows, shadings or texture which cause inhomogeneity of colors of the objects' surface. Using HSI can solve this problem to some extent except that hue is unstable at low saturation. Some physics based models have been proposed to solve this problem [36,37,40–52]. These models take into account the color image formation [17], but they have too many restrictions which prevent them from being extensively applied.

Segmentation may be also viewed as image classification problem based on color and spatial features. Therefore, segmentation methods can be categorized as supervised or unsupervised learning/classification procedures. Ref. [53] compared different color spaces (RGB, normalized RGB, HIS, hybrid color space) and supervised learning algorithms for segmenting fruit images. Supervised algorithms include: maximum likelihood, decision tree, K nearest neighbor, neural networks, etc. Ref. [54] explored six unsupervised color image segmentation approaches: adaptive thresholding, fuzzy C-means, SCT/center split, PCT (principal components transform)/median cut, split and merge, multiresolution segmentation. Ref. [54] showed that combining different methods results in further improvement in the number of correctly identified tumor borders, and by incorporating additional heuristics in merging the segmented object information, the success rate can be further increased. Some algorithms combine unsupervised with supervised methods to segment color images. Ref. [55] used unsupervised learning and classification based on the FCM algorithm and nearest prototype rule. The classified pixels are used to generate a set of prototypes which are optimized using a multilayer neural network. The supervised learning is utilized because the optimized prototypes are subsequently used to classify other image pixels. Ref. [56] employed a neural network for supervised segmentation and a fuzzy clustering algorithm for unsupervised segmentation. The detail description of various unsupervised and supervised approaches to color image segmentation can be found in Refs. [124,130,131].

4.1. *Histogram thresholding (mode method) and color space clustering*

Histogram thresholding is one of the widely used techniques for monochrome image segmentation [19]. It assumes that images are composed of regions with different gray level ranges, the histogram of an image can be separated into a number of peaks (modes), each corresponding to one region, and there exists a threshold value corresponding to valley between the two adjacent peaks. As for color images, the situation is different from monochrome image because of multi-features. Multiple histogram-based thresholding divided color space by thresholding each component histogram. There is some limitation when dividing multiple dimensions because thresholding is a technique for gray scale images. For example, the shape of the cluster is rectangle. Ref. [57] used a tool for clustering color space based on the least sum-of-squares criterion and gave an approximate solution. Since the color information is represented by tristimulus R, G and B or their linear/nonlinear transformations, representing the histogram of a color image in a 3-dimensional array and selecting threshold in this histogram is not a trivial job [5], and detecting the clusters of points within this space will be computationally expensive. One way to solve this problem is to develop efficient methods for storing and processing the information of the image in the 3-D color space. Due to the capability of the RAM is increasing rapidly, this problem alleviates. Ref. [58] used a binary tree to store the 3-D histogram of a color image, where each node of the tree includes RGB values as the key and the number of points whose RGB values are within a range, with this key being the center of the range. Ref. [59] also utilized the same data structure and similar method to detect clusters in the 3-D normalized color space (X, Y, I). Another way is to project the 3-D space onto a lower-dimensional space, such as 2-D or even 1-D. Ref. [60] used projections of 3-D normalized color space (X, Y, I) onto the 2-D planes ($X - Y, X - I$, and $Y - I$) to interactively detect insect infestations in citrus orchards from aerial color infrared photographs. Ref. [61] provided segmentation approaches using 2-D projections of a color space. Ref. [62] suggested a multidimensional histogram thresholding scheme using threshold values obtained from three different color spaces (RGB, YIQ, and HSI). This method uses a mask for region splitting and the initial mask includes all pixels in the image. For any mask, histograms of the nine redundant features (R, G, B, Y, I, Q, H, S, and I) of the masked image are computed, all peaks in these histograms are located, the histogram with the best peak is selected and a threshold is determined in this histogram to split the masked image into two subregions for which two new masks are generated for further splitting. This operation is repeated until no mask left unprocessed, which means none of the nine histograms of existing regions can be further thresholded and each region is homogeneous. Ref. [63] segments a color image into regions with perceptually uniform colors by means of Munsell three-color attributes (H, V and C). A recursive thresholding method similar to the one in Ref. [62] is executed to segment a color image into some meaningful regions. A criterion for locating the best peak is defined as

$$f = (S_p \times 100)/(T_a \times F_h)$$

where S_p represents a peak area between two successive valleys, T_a is the overall area of the histogram, and F_h denotes the full width at half-maximum of the peak. In order to obtain the maximum information between two sources, mode (regions with high densities) and valley (regions with low densities), Ref. [64] adopted entropy based thresholding method. Mode seeking is decided by the multimodel probability density function (pdf) estimation, and the mode can be found by thresholding the pdf.

In the above approaches, thresholding is performed on only one color component at a time. Thus the regions extracted are not based on the information available from all three components simultaneously because the correlation among the three components is neglected. This problem can be solved if we can find such a line that when the points in the 3-D space are projected onto it, and the projected points can be well separated. Ref. [65] utilized the Fisher Linear Discriminant to find such a line for 1-D thresholding. The method operates in the CIE ($L^*a^*b^*$) color space. The cluster distribution in the 3-D space is first estimated using only 1-D histograms. Then, for 1-D thresholding, the clusters are projected onto the line determined by the Fisher discriminant method which can minimize the clustering error rate. This permits the simultaneous utilization of all color information.

Ref. [21] employed the segmentation approach in Ref. [62] to extract a set of effective color features through experiments. Instead of using redundant features for thresholding, it applied Karhunen–Loeve transformation to RGB color space to extract features with large discriminant power to isolate the clusters in a given region. Given a region S, let Σ be the covariance matrix of the distributions of R, G, and B in S, and λ_1, λ_2 and λ_3 be the eigenvalues of Σ and $\lambda_1 \geqslant \lambda_2 \geqslant \lambda_3$. Let $V_i = (v_{Ri},$

$v_{Gi}, v_{Bi})^t$ be the eigenvectors of Σ corresponding to λ_i, i.e.,

$$\Sigma \cdot V_i = \lambda_i \cdot V_i,$$

where $i = 1, 2, 3$. Then, the color features F_1, F_2, and F_3 are defined as:

$$F_i = v_{Ri} \times R + v_{Gi} \times G + v_{Bi} \times B$$

where $\|V_i\| = 1$, and $i = 1, 2, 3$. It can be proved that F_1, F_2, and F_3 are uncorrelated, and F_1 has the largest variance equal to λ_1, thus has the largest discriminant power. F_2 has the largest discriminant power among the vectors orthogonal to F_1. At each step of segmentation, three new color features F_1, F_2, and F_3 are adaptively calculated and used to compute the histograms for thresholding. However, performing K–L transformation at each segmentation step involves too much computational time for real application. Ref. [21] just used it as a tool to find a set of color features as effective as those extracted by K–L transformation (see Section 3.2.3).

Ref. [66] also used the same idea to extract the principal components of a color distribution for detecting clusters. A detail algorithm is given in Ref. [63] for determining the peaks in a histogram. "Clustering of characteristic features applied to image segmentation is the multidimensional extension of the concept of thresholding" [2]. Generally, two or more characteristic features form a feature space and each class of regions is assumed to form a separate cluster in the space. The reason to use multiple characteristic features to perform image segmentation is that, sometimes, there are problems not solvable with one feature but solvable with multiple features. The characteristic features may be any features that could be used for the segmentation problem, such as the gray level value of multi-spectral images, gray level histogram, mean, deviation, texture, etc. For color images, a color space is a natural feature space, and applying the clustering approach to color image segmentation is a straightforward idea, because the colors tend to form clusters in the color space. Clustering has been used as an important pattern recognition technique for many years. The biggest problem that it suffers from is how to determine the number of clusters in an unsupervised clustering scheme, which is known as cluster validity. As for a color image, the selection of color space is quite critical to this approach. For example, if R, G, B is selected for color clustering, because of the high correlation among R, G and B, the object with a uniform color but different intensities could be segmented into different objects. In other words, color image with shadows or shading cannot be segmented properly.

Ref. [67] presented a two-stage segmentation algorithm for color images based on 1-D histogram thresholding and the fuzzy c-means (FCM) techniques which are viewed as coarse and fine segmentations, respectively. At the coarse stage, 1-D histogram thresholding is applied to one color component a time (RGB, YIQ, and $I_1 I_2 I_3$), then scale-space filter is used to locate the thresholds in each color component's histogram. These thresholds are then used to partition the color space into several hexahedra. The hexahedra that contain a number of pixels exceeding a predefined threshold are declared as valid clusters. All peaks in the 1-D histograms could be the cluster centers. The coarse segmentation attempts to automatically find the number of classes and the center of each class. At the fine stage, FCM is used to assign the unclassified pixels to the closest class using the cluster centers detected at the coarse stage. This technique tried to solve the problem of cluster validity by using a coarse–fine concept. But it is still based on the assumption that the histograms are not unimodal. It has the same disadvantage as most thresholding techniques have. Ref. [26] proposed an iterative multispectral image segmentation approach using FCM, and performed experiments in RGB and Ohta color spaces. The Ohta feature space is a linear transformation of RGB space. The three features are $(R + G + B)/3$, $(R - B)$ and $(2*R - G - B)/2$. They are significant in this order, and in many cases good segmentation can be achieved by using only the first two.

Ref. [68] proposed a method based on K-nearest neighbor (K-NN) technique for detecting fruit and leaves in a color scene, which is used to build a vision system for a robotic citric harvesting device. The feature vectors are based on the YUV color space. In order to incorporate information about shape and surface of the fruit and leaves, not only the color components of the pixel itself but also the components of its four neighbors are included in the vector $X(i,j)$

$$\begin{aligned} X(i,j) = [&U(i,j), V(i,j), U(i+h,j), V(i+h,j), U(i-h,j), \\ &V(i-h,j), U(i,j+h), V(i,j+h), U(i,j-h), \\ &V(i,j-h)] \end{aligned}$$

where $U(i,j)$ and $V(i,j)$ are the U and V components of YUV color coordinates and h represents the neighbor. How to choose reference sample becomes important because only finite reference can be used in practice. Secondly, mislabeled reference and/or "outlyer" can reduce the classification accuracy. Thirdly, the more the number of references, the more noise or erroneous data. Finally, large reference sets can bring a problem of computational speed. The combined technique of multiediting and condensing is applied to reduce the size of the reference set. By doing so, still with very high accuracy, the computational time can be reduced dramatically.

4.2. Region based approaches

Region based approaches, including region growing, region splitting, region merging and their combination, attempt to group pixels into homogeneous regions. In the region growing approach, a seed region is first selected, then expanded to include all homogeneous neighbors, and this process is repeated until all pixels in the image are classified. One problem with region growing is its inherent dependence on the selection of seed region and the order in which pixels and regions are examined. In the region splitting approach, the initial seed region is simply the whole image. If the seed region is not homogeneous, it is usually divided into four squared sub-regions, which become new seed regions. This process is repeated until all sub-regions are homogeneous. The major disadvantage of region splitting is that the resulting image tends to mimic the data structure used to represent the image and comes out too square. The region merging approach is often combined with region growing or region splitting to merge the similar regions for making a homogeneous region as large as possible.

These techniques work best on images with an obvious homogeneity criterion and tend to be less sensitive to noise because homogeneity is typically determined statistically. They are better than feature space thresholding or clustering techniques by taking into account both feature space and the spatial relation between pixels simultaneously. However, all region based approaches are by nature sequential, and another problem with these techniques is their inherent dependence on the selection of seed region and the order in which pixels and regions are examined. Refs. [21,62,63,65] used region splitting approach to segment color images. The homogeneous criteria utilized by them are based on 1-D histogram thresholding on the features of color components or extracted from color spaces.

Ref. [69] proposed a color segmentation approach which combines region growing and region merging techniques. It starts with the region growing process using the criteria based on both color similarity and spatial proximity. Euclidean distance over R, G, B color space is used to define the color similarity, which defines three criteria of color homogeneity: the local homogeneity criterion (LHC) corresponding to a local comparison between adjacent pixels; the first average homogeneity criterion (AHC1) corresponding to a local and regional comparison between a pixel and its neighborhood, considering only the region under study; and the second average homogeneity criterion (AHC2) corresponding to a global and regional comparison between a pixel and the studied region. The regions generated by region growing process are then merged on the basis of a global homogeneity criterion based on color similarity to generate a nonpartitioned segmentation consisting of spatially disconnected but similar regions. The problem with this method is that the selection of the thresholds for these criteria is rather subjective and the thresholds are image dependent. Another problem is that it is not applicable to images with shadows or shading.

In order to recognize the small object or local variance of color image, Ref. [70] proposed a hierarchical segmentation which identifies the uniform region via a thresholding operation on a homogeneity histogram. The homogeneity is defined as a composition of two components: variance and discontinuity of the intensity, $(R + G + B)/3$, the local information as well as global information is taken into account, therefore, the quality of the segmentation result is much improved. The region based approach is widely used in color image segmentation because it considers the color information and spatial details at the same time.

4.3. Edge detection

Edge detection is extensively utilized for gray level image segmentation, which is based on the detection of discontinuity in gray level, trying to locate points with abrupt changes in gray level. Edge detection techniques are usually classified into two categories: sequential and parallel [1,2]. A parallel edge detection technique means that the decision of whether or not a set of points are on an edge is not dependent on whether other sets of points lie on an edge or not. In principle, the edge detection operator can be applied simultaneously all over the image. One technique is high-emphasis spatial frequency filtering. Since high spatial frequencies are associated with sharp changes in intensity, one can enhance or extract edges by performing high-pass filtering using the Fourier operator. The problem here is how to design a relevant filter. There are many types of parallel differential operators such as Roberts, Sobel, and Prewitt operators, which are called the first-difference operators, and the Laplacian operator, which is called the second-difference operator. The main differences between these operators are the weights assigned to each element of the mask. These operators require that there is a distinct change in gray level between two adjacent points, and only very abrupt edges between two regions could be detected. They cannot detect ill-defined edges that are formed by a gradual change in gray level across the edge. Since the computation is based on a small window, the result is quite susceptible to noise. Sequential edge detection means that the result at a point is dependent on the result of the previously examined points. There are a number of sequential techniques utilizing heuristic search and dynamic programming. The performance of a sequential edge detection algorithm will depend on the choice of a good initial point, and it is not easy to define a termination criterion.

In a monochrome image, edge is defined as a discontinuity in the gray level, and can be detected only when there is a difference of the brightness between two regions. However, in color images, the information about edge is much richer than that in monochrome case. For example, edges between two objects with the same brightness but different hue can be detected in color images [71]. Accordingly, in a color image, an edge should be defined by a discontinuity in a three-dimensional color space. Ref. [23] gave three alternatives for the definition of a color edge: (i) Define a metric distance in some color space and use discontinuities in the distance to determine edges. This makes color edge detection still be performed in 1-D space. Hence the result cannot be expected to be better than that achieved by edge detection in an equivalent monochrome image. (ii) Regard a color image as composed of three monochrome images formed by the three color components, respectively, and perform gray level edge detection on these three images separately. Then the edges detected in the three images might be merged by some specified procedures. This is still essentially a gray-level edge detection technique and may be unsatisfactory in some cases, for example, when gradient edge detectors are employed, the three gradients for one pixel may have the same strength but in opposite directions [72,73]. (iii) Impose some uniformity constraints on the edges in the three color components to utilize all of the three color components simultaneously, but allow the edges in the three color components to be largely independent. Actually, these constraints directly affect the computation of the three color components, which makes definition (iii) essentially different from definition (ii).

Ref. [23] used definition (iii) to define a color edge, which is based on the edge operator provided by [11,12]. This definition allows the edges in the three components to be independent, except that the spatial angles of the edges have to be the same. Y, T_1, and T_2 are used to compute color edges using an extension of the edge operator in Refs. [11,12], and the edges in hue and saturation components are derived from the computations performed on the Y, T_1, T_2 components. The results of the experiments on two images show that most of the edges in the chromatic components correspond to the boundaries of desired objects, and the number of edges detected in chromatic components (T_1, T_2, hue and saturation) is smaller than that of the edges detected in the brightness component (Y) because the color in an image is relatively invariant over the object surface while the brightness may vary due to nonuniform illumination and reflections, and for a large percentage of the cases, the brightness edge also exists where a chromatic edge exists. However situations exist in images with low contrast or poor illumination where the brightness edges are absent but chromatic edges are present. The explanation is that in natural scenes, it is unlikely that objects of different hue will accidentally have the same brightness component. Based on the above observation, Ref. [23] suggested that the use of color is more likely to aid in obtaining reliable initial data in a multilevel segmentation scheme because the chromatic edges seem to contain fewer spurious edges, and also concluded that ignoring the color information results in only a graceful degradation in performance because the brightness edges and chromatic edges tend to be highly correlated. This is true only for the two images employed in that paper, because there are hardly any shadows or shading in these two images. For the images with considerable amount of shadows and shading, this conclusion will not be appropriate any more. The hue and saturation information turns out to be very important, because in such a case the brightness information cannot, but the hue information can, be used to detect proper edges at the boundaries of shadows or shading. Therefore, if there exists a large amount of shadows or shading, the degradation of performance resulted from ignoring chromatic information will no longer be graceful.

Ref. [20] discussed the choice of the color spaces for edge detection. It considered *RGB*, *YIQ*, *CIE* ($L^*a^*b^*$), $G_1^*G_2^*G_3^*$ and $P_1P_2P_3$, where $G_1^*G_2^*G_3^*$ is obtained using the model of the human visual system, and P_1, P_2 and P_3 are the three principal components obtained using the normalized covariance matrix of the red, green and blue components of a color image. The normalized covariance matrices of the three components in the color spaces of a girl picture are given. For each color space, edge detection is performed using the compass gradient edge detection method. The color edge is determined by the maximum value of the 24 gradient values in three components and eight directions at a pixel. By analyzing the energy content and edge activity index of the color components, it was concluded that the inherent cross-correlation of the three components in a color space should be considered for color edge detection, and the most effective components for color edge detection are G, P_1, Y, L^* and G_1^* which are related to the brightness components in the five color spaces. However, Ref. [20] did not take into account the effect of shadows or shading.

Ref. [74] argued that the edges detected in hue correlate more directly with material boundaries than those detected in brightness, *RGB*, or normalized space, and the color segmentation in the *HSI* space could make the segmentation results enhanced in images with high saturation, even in the presence of confounding cues due to shading, shadow, transparency, and highlights. It compared several color spaces such as *RGB*, Normalized *RGB*, *HSI* and ($L^*a^*b^*$), discussed the properties of hue, and demonstrated that due to the additive/shift invariance and multiplicative/scale invariance properties, hue is invariant to transparencies and certain types of highlights, shadows and shading. The method for edge detection in the paper is: first, use Canny edge operator [77] to generate an intensity edge map; then gradually eliminate the edges if the hue changes are small. The major problem of the *HSI* space is its singularity at the axis of the color cylinder where $R = G = B$ or saturation $= 0$. Hue is unstable at low saturations, thus, using hue to segment regions with low saturations is not reliable. Ref. [74]

utilized a first-order membrane type stabilizer based on Markov random fields to smooth the unstable hue regions of low saturations or low intensities and obtained much improvement in color segmentation. The experiments show that starting with an intensity edge map and applying the first-order smoothness operator to the hue map generated by a modulo edge detector can result in a hue edge map that discounts confounding cues due to shading, shadow, transparency, and highlights. Based on this work, Ref. [74] designed and fabricated an analog CMOS VLSI circuit with on-board phototransistor input that computes normalized color and hue. Ref. [75] adopted an improved watershed algorithm which uses gradient of gray level to segment image into homogeneous regions and oversegmentation is also considered. Ref. [76] incorporated the zero-, first-, and second-order derivatives. Some advantages of the algorithm are: 1. Precise region boundaries are detected. 2. Overcome undesirable undersegmentation caused by the oversmoothing and oversegmentation of ramp-like region. 3. Small segments caused by noise variations in statistically coherent regions are removed.

We want to emphasize here that edge detection cannot segment an image by itself. It can only provide useful information about the region boundaries for the higher-level systems, or it can be combined with other approaches, e.g., region based approaches [78–83], to complete the segmentation tasks.

4.4. Fuzzy techniques

The segmentation approaches mentioned above take crisp decisions about regions. Nevertheless, the regions in an image are not always crisply defined, and uncertainty can arise within each level of image analysis and pattern recognition. It can occur at the low level in the raw sensor output, and extend all the way through intermediate and higher levels. Since decisions at any level are based on the results of previous levels, any decision made at a previous level will have an impact on all higher-level activities. A recognition or computer vision system must have sufficient flexibility for processing of uncertainty in any of these levels so that the system could retain as much information as possible at each level. In this way, the final output of the system may not be biased too much by lower level decisions, unlike the classical approaches. As the first essential step of a recognition or vision system, image segmentation particularly should have such a provision for representing and manipulating the uncertainties.

Fuzzy set theory provides a mechanism to represent and manipulate uncertainty and ambiguity. Fuzzy operators, properties, mathematics, and inference rules (IF–THEN rules) have found considerable applications in image segmentation [1,84–107]. Prewitt first suggested that the output of image segmentation should be fuzzy subsets rather than ordinary subsets [1]. In fuzzy subsets, each pixel in an image has a degree to which it belongs to a region or a boundary, characterized by a membership value. By doing so, we can avoid making a crisp decision earlier and keep the information through the higher processing levels as much as possible.

Recently, there has been an increasing use of fuzzy logic theory for color image segmentation [20,26,63, 108–120]. As mentioned above, for color images, the colors tend to form clusters in color space which can be regarded as a natural feature space. One problem with traditional clustering techniques is that there are only two values, either 1 or 0, to indicate to what degree a data point belongs to a cluster. This requires well-defined boundaries between clusters, which is not the usual case for real images. This problem can be solved by using fuzzy set methods. FCM is a method that can allow ambiguous boundaries between clusters, and has received much attention [26]. It is an iterative optimization method. It calculates the memberships of a data point in each of the clusters based on the distances between the point and the cluster centers. The cluster centers are then updated based on the resulting clusters. In the iteration, an objective criterion function is used to minimize the distance between the data point in a cluster and the cluster center, and to maximize the distance between cluster centers.

In a two-stage (coarse and fine) segmentation algorithm, Ref. [67] used FCM to assign the unclassified pixels at the coarse stage to the closest class using the detected cluster centers. This technique tried to solve the problem of cluster validity by using a coarse–fine concept. But its applicability is still limited by the histogram thresholding technique. Ref. [26] proposed an iterative multispectral image segmentation approach based on FCM and performed experiments on a color image in *RGB* and Ohta color spaces. There are two major problems with the FCM method: (1) How to determine the number of clusters remains unsolved. (2) The computational cost is quite high for large data sets.

Ref. [108] defined a segment as "a collection of touching pixels having almost the same color while the change in color is gradual" and utilized a fuzzy approach to attack this fuzzy concept. It defined a contrast fuzzy membership function on a *RGB* space as follows:

$$\mu_c = \begin{cases} 0 & \text{if } Contrast \leqslant a_1 \\ 1 & \text{if } Contrast > a_2 \\ (Contrast - a_1)/(a_2 - a_1) & \text{otherwise} \end{cases}$$

where $\text{Contrast} = \sqrt{(R_v - R_w)^2 + (G_v - G_w)^2 + (B_v - B_w)^2}$ is the contrast between two pixels v and w, a_1 and a_2 are two predetermined thresholds. The concept of contrast is used to determine the homogeneity criterion for the region-growing approach, which is defined as: the absolute contrast (contrast between pixel and region) is low and the local or relative contrast (contrast between pixel and its neighbors in the growing direction) is also low. This criterion can keep the segment growing even when there is a gradual change of color on the surface of an object and also prevents two regions with similar color but separated by an edge from merging. Using fuzzy membership function, Ref. [108] also defined degree of farness μ_f, which is determined by the distance in spatial domain and the contrast in color space, i.e., $\mu_{\mathrm{f}} = \mu_d \times \mu_c$ where mu_d is the membership function of distance and μ_c is the membership function of contrast. The farness is then used to assign the unclassified pixels to the closest clusters. The problem is that *R*, *G* and *B* are highly correlated and subject to the illumination, which makes the contrast defined by *RGB* not reliable.

Ref. [109] proposed a color edge detection approach based on fuzzy IF–THEN rules in the *HSI* color space. Linear fuzzy membership functions are used to describe the absolute difference in the three components of *HSI* between two pixels, and, particularly, to describe the relevance of hue due to the unstability of hue at low saturation. Several 3×3 masks are used to describe the potential edge structures, and for each mask, a fuzzy IF–THEN rule is developed. These rules are then combined using fuzzy "OR" operator to infer fuzzy subsets representing the potential edges. The fusion of the edges detected in hue, saturation and intensity is quite natural using fuzzy logic representation. The final fuzzy set *PEP–HSI* (Potential Edge Pixel in *HSI* space) is obtained by a weighted combination of three sets representing three types of edge pixels (S_3: edge pixels in three components, S_2: edge pixels in two components, and S_1: edge pixels in only one component). The membership function for *PEP–HIS* is defined as

$$\mu_{pep\text{-}his} = 1/2 \times \mu_{s3} + 1/3 \times \mu_{s2} + 1/6 \times \mu_{s1}.$$

Ref. [110] proposed a color image segmentation algorithm based on fuzzy homogeneity. The fuzzy set theory and maximum fuzzy entropy principle are used to map the color image from space domain to fuzzy domain, which will keep the maximum information. Both the global and local information is taken into account while calculating fuzzy homogeneity histogram. The scale-space filter (SSF) is utilized to analyze the homogeneity histogram to find the appropriate segments. The final result is transformed from fuzzy domain to space domain using the inverse S-function. The experimental results demonstrate the effectiveness of the proposed approach. Ref. [111] employed the concept of homogram to extract homogeneous regions in a color image. Utilizing the fuzzy homogeneity approach to find thresholds for each color component. Then the segmentation results for the three-color components are combined. In order to solve the problem of over-segmentation, a region merging process is performed based on color similarity. This approach is more effective in finding homogeneous regions than histogram based approach.

4.5. Physics based approaches

Physics based segmentation approaches aim at solving this problem by employing physical models to locate the objects' boundaries while eliminating the spurious edges of shadow or highlights in a color image. Among the physics models, "dichromatic reflection model" [36] and "approximate color-reflectance model (ACRM)" [37] are the most common ones.

Reflection is highly related to the nature of the materials. Ref. [37] divided materials into different classes: optically homogeneous materials like metals, glass and crystals, and optically inhomogeneous materials such as plastics, paper, textiles and paints. Usually it is very helpful to identify or classify the material in the scene of an image before the algorithm is applied. For example, we should distinguish metals from dielectrics since they interact with lights in different ways and require different algorithms for image understanding.

Using the color feature of human face, Ref. [51] built a skin color model to capture the chromatic characteristics. Human skin color fall in a small region in color space and can be approximated by a Gaussian distribution. According to color and shape features of human faces, Ref. [51] concluded: regardless of the size, orientation and viewpoint, human face in color image can be identified. In order to segment images containing multi-colored object and multiple materials, Ref. [52] proposed a model consisting of three elements: surface, illumination, and the light transfer function. The parameter space for every element is divided into a set of subspaces, which allow reasoning about the relationships of adjacent hypothesis region. Based on physical parameters, Ref. [52] gave all possible combinations of the subsets and some rules of merging. This approach is easily expendable and allows greater complexity in images.

Based on "dichromatic reflection model", Ref. [36] created a method to determine the amount of interface reflection and body reflection in a color image pixel by pixel, and presented an algorithm for analyzing color values, which is very useful in color-image understanding. "ACRM" [37] demonstrated the independence of the spectral composition and geometrical scaling of the light reflected. This model is consistent with dichromatic reflection model when the materials are inhomogeneous dielectrics.

Physics based segmentation approaches use the same segmentation techniques discussed before. For example, [42] used Canny's edge detector [77] to segment an image of a valve based on the ACRM model, and [121] applied clustering method to color image segmentation based on the dichromatic reflection model. The characteristic of these approaches lies in that they use different reflection models for color images. The existing physics based models are efficient only in image processing for the materials whose reflection properties are known and easy to model. There are too many rigid assumptions of these physics models regarding the material type, the light source and illumination. These conditions may not be satisfied in the real world. Therefore, these models can be used only in a very limited scope of applications.

4.6. *Neural networks approaches*

Artificial Neural Networks (ANN) are widely applied for pattern recognition. Their extended parallel processing capability and nonlinear characteristics are used for classification and clustering. ANN explore many competing hypotheses simultaneously through parallel nets instead of performing a program of instructions sequentially, hence ANN can be feasible for parallel processing. Neural networks are composed of many computational elements connected by links with variable weights. The complete network, therefore, represents a very complex set of interdependencies which may incorporate any degree of nonlinearity, allowing very general function to be modeled. Training time are usually very long, but after training, the classification using ANN is rapid.

4.6.1. *Hopfield Neural Networks* (*HNN*)

Ref. [122] presented the segmentation problem for gray-level image as minimizing a suitable energy function for Hopfield networks. It derived the network architecture from the energy function. Based on the idea in Refs. [122,123] described two algorithms to segment the color images using HNN. The first algorithm locates the significant peaks by applying histogram analysis and designs three different networks (one for each color feature). The segmentation results of three color components are combined to get the final image. *RGB* color features, the I_1, I_2 and I_3 color features, and the Karhunen–Loeve transformation of the *RGB* color features (*KL–RGB*) have been used in the experiments. Ref. [123] gave another algorithm built a single Hopfield network with $M \times N \times S$ neurons to segment color images. M, N are the image size, and S is the number of selected clusters obtained by analyzing the histogram. For both algorithms, histogram analysis is very important since it produces the coarse segmentations and determines both network structures and their initializations. In the algorithms, the spatial information is also considered in order to produce consistent color pixel labeling.

An unsupervised algorithm [124] used HNN to segment the color image of liver tissues prepared and stained by standard method. The results show that the *RGB* color space representation of the color images is more suitable than *HSV* and *HLS* color spaces. This algorithm can automatically extract the nuclei region and cytoplasm region which are useful for diagnosis. Ref. [125] attempted to segment sputum color images in order to build an automatic diagnosis system for lung cancer. After masking nonsputum cell, HNN can make a crisp classification of cells by labelling pixels as background, cytoplasm and nucleus. The technique has yielded correct segmentations of complex scene, however, more work needs to be done to solve the overlap of the cells.

In Ref. [126], every boundary pixel is assigned to an element of the HNN and a local minimum can be found by HNN. In order to reduce the computing time, the pyramid images are used to perform fast segmentation and obtain a global optimal solution. The active region segmentation method is based on a regularization approach in which region growing is incorporated into edge detection.

4.6.2. *The self-organizing map* (*SOM*)

The self-organizing map (SOM) projects input space on prototypes of low-dimensional regular grids that can be effectively utilized to visualize and explore properties of the data [127]. Self-organization of Kohonen Feature Map (SOFM) network is a powerful tool for data clustering. Ref. [128] employed the watershed segmentation to the luminance component of color image. In order to solve the problem of oversegmentation, the Kohonen self-organizing map (SOM) network is used. The area of the watershed segments, the chrominance components and the luminance component are input into the network for obtaining the information about how to merge the image. In Ref. [129], first, SOFM is used for quantization of the input color image in order to reduce computational time, and get an indexed image. Second, local histogram is calculated by using a moving window and index-count vectors are obtained. Third, the index-count vectors are used as the training data for SOFM. Finally, each cluster is mapped from the index-count space to the original image. Ref. [129] stated that SOMF was a fast training method and parallel hardware structure was also given.

4.6.3. *Other neural networks*

Backpropagation (BP) algorithm can be used to segment color images as well. Color features are successively input into BP, one pixel at a time, and three input nodes are necessary. Ref. [130] presents the BP architecture and how to train BP network. C-mean and Learning Vector Quantization (LVQ) algorithms are compared with the back-propagation neural approach. Ref. [131] studied local linear map network (LLM), which is related to self-organizing maps. LLM network, representing the to-be-learned mappings as a collection of locally valid linear mappings that are learned by separate units, is a feedforward neural network. LLM is trained in a supervised-learning scheme to yield a probability value for each image pixel belonging to the face region. In order to

Table 1
Monochrome image segmentation techniques

Segmentation technique	Method description	Advantages	Disadvantages
Histogram thresholding (mode method)	Requires that the histogram of an image has a number of peaks, each corresponds to a region	It does not need a prior information of the image. For a wide class of images satisfying the requirement, this method works very well with low computation complexity	(1) Does not work well for an image without any obvious peaks or with broad and flat valleys (2) Does not consider the spatial details, so cannot guarantee that the segmented regions are contiguous
Feature space clustering	Assumes that each region in the image forms a separate cluster in the feature space. Can be generally broken into two steps: (1) categorize the points in the feature space into clusters; (2) map the clusters back to the spatial domain to form separate regions	Straightforward for classification and easy for implementation	(1) How to determine the number of clusters (known as cluster validity) (2) Features are often image dependent and how to select features so as to obtain satisfactory segmentation results remains unclear (3) Does not utilize spatial information
Region-based approaches	Group pixels into homogeneous regions. Including region growing, region splitting, region merging or their combination	Work best when the region homogeneity criterion is easy to define. They are also more noise immune than edge detection approach	(1) Are by nature sequential and quite expensive both in computational time and memory (2) Region growing has inherent dependence on the selection of seed region and the order in which pixels and regions are examined (3) The resulting segments by region splitting appear too square due to the splitting scheme
Edge detection approaches	Based on the detection of discontinuity, normally tries to locate points with more or less abrupt changes in gray level. Usually classified into two categories: sequential and parallel	Edge detecting technique is the way in which human perceives objects and works well for images having good contrast between regions	(1) Does not work well with images in which the edges are ill-defined or there are too many edges (2) It is not a trivial job to produce a closed curve or boundary (3) Less immune to noise than other techniques, e.g., thresholding and clustering
Fuzzy approaches	Apply fuzzy operators, properties, mathematics, and inference rules (IF– THEN rules), provide a way to handle the uncertainty inherent in a variety of problems due to ambiguity rather than randomness	Fuzzy membership function can be used to represent the degree of some properties or linguistic phrase, and fuzzy IF–THEN rules can be used to perform approximate inference	(1) The determination of fuzzy membership is not a trivial job (2) The computation involved in fuzzy approaches could be intensive
Neural network approaches	Using neural networks to perform classification or clustering	No need to write complicated programs. Can fully utilize the parallel nature of neural networks	(1) Training time is long (2) Initialization may affect the results (3) Overtraining should be avoided

Table 2
Characteristics of color spaces

Color space	Advantages	Disadvantages
RGB	Convenient for display	Not good for color image processing due to the high correlation
YIQ	Can be used to efficiently encode color information in the TV signal of American system; Partly gets rid of the correlation of *RGB*; Involves less computation time; *Y* is good for edge detection	Correlation still exists due to the linear transformation, though not as high as *RGB*
YUV	Can be used to efficiently encode color information in the TV signal of European system; Partly gets rid of the correlation of *RGB*; Involves less computation time	Correlation still exists due to the linear transformation, though not as high as *RGB*
$I_1I_2I_3$	Partly gets rid of the correlation of *RGB*; Involves less computation time; Can be useful for color image processing	Correlation still exists due to the linear transformation, though not as high as *RGB*
HSI	Based on human color perception; Useful in some cases where the illumination level varies, because hue is invariant to certain types of highlights, shading, and shadows; Hue can be useful for separating objects with different colors	Nonremovable singularity and numerically unstable at low saturation due to nonlinear transformation
Nrgb (Normalized *rgb*)	The individual color components are independent on the brightness of the image; Convenient to represent the color plane; Robust to the change of the illumination	Very noisy at low intensities due to nonlinear transformation.
CIE spaces ($L^*u^*v^*$ or $L^*a^*b^*$)	Can control color and intensity information independently; Direct color comparison can be performed based on geometric separation within CIE space, and efficient in measuring small color difference	Have the same singularity problem as other nonlinear transformations do

get better segmentation results, a training database including different faces and views is used. An oscillatory cellular neural network (OCNN) is employed in Ref. [132] to segment color images. Its architecture consists of an array of simple neural oscillators with inter-connections limited to the nearest neighborhood. The advantage of OCNN is that it solves a bottleneck created by the immense number of interconnections between a global separator and the oscillators. Connectedness between neighboring color pixels is defined as color connectedness matrix (CCM) to group the correlated segments. Simulation results demonstrate the validity and performance of OCNN. Ref. [133] used a Constrain Satisfaction Neural Network (CSNN) to perform MAP segmentation which utilizes the advantages of GMRF (Gauss–Markov Random Field). Real world and synthetic images are tested based on *RGB* color space.

4.7. Discussions on color image segmentation

There are two critical issues for color image segmentation: (1) what segmentation method should be utilized; and (2) what color space should be adopted. At present, color image segmentation methods are generally extended from monochrome segmentation approaches. Several approaches applied to color image are discussed in this section, including histogram thresholding, region based approaches, edge detection and fuzzy techniques. A combination of these approaches is often utilized for color image segmentation [21,33,62,63,65,78–83,134]. Table 1 is a summary of these approaches. Other approaches, such as segmentation using Markov Random Field [135–137] and segmentation based on texture [28,138,139] can also be found from the literature.

The selection of a color space for image processing is image/application dependent. There is no any color space which is better than others and suitable to all kinds of images yet. Table 2 lists the characteristics of different color spaces for color image segmentation.

5. Summary

There is no universal theory on color image segmentation yet. All of the existing color image segmentation approaches are, by nature, ad hoc. They are strongly application dependent, in other words, there are no general algorithms and color space that are good for all color images. An image segmentation problem is basically one of psychophysical perception, and it is essential to supplement any mathematical solutions by a priori knowledge about the picture knowledge. Most gray level image segmentation techniques could be extended to color image, such as histogram thresholding, clustering, region growing, edge detection and fuzzy based approaches. They can be directly applied to each component of a color space, then the results can be combined in some way to obtain the final segmentation result. However, one of the problems is how to employ the color information as a whole for each pixel. When color is projected onto three components, the color information is so scattered that the color image becomes simply a multispectral image and the color information that human can perceive is lost. Another problem is how to choose the color representation for segmentation, since each color representation has its advantages and disadvantages.

In most of the existing color image segmentation approaches, the definition of a region is based on similar color. This assumption often makes it difficult for many algorithms to separate the objects with highlights, shadows, shading or texture which cause inhomogeneous colors of the objects' surface. Using HSI can solve this problem to some extent except that hue is unstable at low saturation. Some physics based models have been proposed to find the objects' boundaries based on the type of materials, but these models have too many restrictions which limit them to be extensively employed.

The fuzzy set theory has attracted more and more attention in the area of image processing. Fuzzy set theory provides us with a suitable tool which can represent the uncertainties arising in image segmentation and can model the cognitive activity of the human beings [114]. Fuzzy operators, properties, mathematics, and inference rules (IF–THEN rules) have found more and more applications in image segmentation. Despite the computational cost, fuzzy approaches perform as well as or better than their crisp counterparts. The more important advantage of a fuzzy methodology lies in that the fuzzy membership function provides a natural means to model the uncertainty in an image. Subsequently, fuzzy segmentation results can be utilized in feature extraction and object recognition phases of image processing and computer vision. Fuzzy approach provides a promising means for color image segmentation.

References

[1] S.K. Pal et al., A review on image segmentation techniques, Pattern Recognition 29 (1993) 1277–1294.
[2] K.S. Fu, J.K. Mui, A survey on image segmentation, Pattern Recognition 13 (1981) 3–16.
[3] E.M. Riseman, M.A. Arbib, Computational techniques in the visual segmentation of static scenes, Comput. Vision Graphics Image Process. 6 (1977) 221–276.
[4] W. Skarbek, A. Koschan, Colour image segmentation – a survey, Technical Report, Tech. Univ. of Berlin, October 1994.
[5] R.M. Haralick, L.G. Shapiro, Image segmentation techniques, Comput. Vision Graphics Image Process. 29 (1985) 100–132.
[6] P.K. Sahoo et al., A survey of thresholding techniques, Comput. Vision Graphics Image Process. 41 (1988) 233–260.
[7] L. Spirkovska, A summary of image segmentation techniques, NASA Technical Memorandum 104022, June 1993.
[8] C.H. Chen, On the statistical image segmentation techniques, Proceedings of IEEE Conference on Pattern Recognition and Image Processing, 1981, pp. 262–266.
[9] A. Rosenfeld, M. Thurston, Edge and curve detection for visual scene analysis, IEEE Trans. Comput. C-20 (1971) 562–569.
[10] A. Rosenfeld, M. Thurston, Y. Lee, Edge and curve detection further experiments, IEEE Trans. Comput. C-21 (1972) 677–715.
[11] M. Hueckel, An operator which locates edges in digitized pictures, J. Assoc. Comput. Mach. 18 (1) (1971) 113–125.
[12] M. Hueckel, A local visual operator which recognizes edges and lines, J. Assoc. Comput. Mach. 20 (4) (1973) 634–647.
[13] J. Gauch, Chi-Wan Hsia, A comparison of three color image segmentation algorithm in four color spaces, SPIE Vol. 1818 Visual Communications and Image Processing '92, 1992, pp. 1168–1181.
[14] D.E.P. Hoy, On the use of color imaging in experimental applications, Exp. Tech. 21 (4) (1997) 17–19.
[15] M. Chapron, A new chromatic edge detector used for color image segmentation, IEEE International Conference on Pattern Recognition, A, 1992, pp. 311–314.

[16] M.T. Orchard, C.A. Bouman, Color quantization of images, IEEE Trans. Signal Process. 39 (12) (1991) 2677–2690.

[17] D. Comaniciu, P. Meer, Robust analysis of feature spaces: color image segmentation, IEEE Conference on Computer Vision and Pattern Recognition, 1997, pp. 750–755.

[18] M. Pietikainen et al., Accurate color discrimination with classification based on feature distributions International Conference on Pattern Recognition, C, 1996, pp. 833–838.

[19] E. Littmann, H. Ritter, Adaptive color segmentation – a comparison of neural and statistical methods, IEEE Trans. Neural Network 8 (1) (1997) 175–185.

[20] G.S. Robinson, Color edge detection, Opt. Eng. 16 (5) (1977) 479–484.

[21] Y. Ohta, T. Kanade, T. Sakai, Color information for region segmentation, Comput. Graphics Image Process. 13 (1980) 222–241.

[22] P. Golland, A.M. Bruckstein, Why R.G.B.? Or How to design color displays for martians, Graphical Models Image Process. 58 (5) (1996) 405–412.

[23] Nevatia, A color edge detector and its use in scene segmentation, IEEE Trans. System Man Cybernet. SMC-7 (11) (1977) 820–826.

[24] I. Andreadis, M.A. Browne, J.A. Swift, Image pixel classification by chromaticity analysis, Pattern Recognition Lett. 11 (1990) 51–58.

[25] J.C. Terrillon, M. David, S. Akamatsu, Detection of human faces in complex scene images by use of a skin color model and of invariant Fourier-Mellin moments, IEEE International Conference on Pattern Recognition, 1998, pp. 1350–1355.

[26] T.L. Huntsberger, C.L. Jacobs, R.L. Cannon, Iterative fuzzy image segmentation, Pattern Recognition 18 (2) (1985) 131–138.

[27] T. Carron, P. Lambert, Color edge detector using jointly hue, saturation and intensity, IEEE International Conference on Image Processing, Austin, USA, 1994, pp. 977–1081.

[28] Y. Rui, A.C. She, T.S. Huang, Automated region segmentation using attraction-based grouping in spatial-color-texture space, International Conference on Image Processing, A, 1996, pp. 53–56.

[29] W.S. Kim, R.H. Park, Color image palette construction based on the HSI color system for minimizing the reconstruction error, IEEE International Conference on Image Processing, C, 1996, pp. 1041–1044.

[30] P.W.M. Tsang, W.H. Tsang, Edge detection on object color, IEEE International Conference on Image Processing, C, 1996, pp. 1049–1052.

[31] E. Tepichin, J.G. Suarez-Romero, G. Ramirez, Hue, brightness, and saturation manipulation of diffractive colors, Opt. Eng. 34 (10) (1995) 2886–2890.

[32] K.M. Kim, C.S. Lee, Y.H. Ha, Color image quantization using weighted distortion measure of HVS color activity, IEEE International Conference on Pattern Recognition, 1996, pp. 1035–1039.

[33] D.C. Tseng, C.H. Chang, Color segmentation using perceptual attributes, IEEE International Conference on Pattern Recognition, A, 1992, pp. 228–231.

[34] S. Tominaga, Expansion of color images using three perceptual attributes, Pattern Recognition Lett. 6 (1987) 77–85.

[35] N. Vandenbroucke, L. Macaire, J.G. Posaire, Color pixels classification in an hybrid color space, IEEE International Conference on Image Processing, 1998, pp. 176–180.

[36] S.A. Shafer, Using color to separate reflection components, Color Res. Appl. 10 (4) (1985) 210–218.

[37] G. Healey, Using color for geometry-insensitive segmentation, Opt. Soc. Am. 22 (1) (1989) 920–937.

[38] J. Kender, Saturation, hue, and normalized color: calculation, digitization effects, and use, Computer Science Technical Report, Carnegie Mellon University, 1976.

[39] C.K. Yang, W.H. Tsai, Reduction of color space dimensionality by moment-preserving thresholding and its application for edge detection in color images, Pattern Recognition Lett. 17 (1996) 481–490.

[40] G.J. Klinker, S.A. Shafer, T. Kanade, A physical approach to color image understanding, Int. J. Comput. Vision 4 (1990) 7–38.

[41] M.H. Brill, Image segmentation by object color: a unifying framework and connection to color constancy, Opt. Soc. Am. 7 (10) (1990) 2041–2047.

[42] G. Healey, Segmenting images using normalized color, IEEE Trans. System Man Cybernet. 22 (1) (1992) 64–73.

[43] A.P. Petrov, L.L. Kontsevich, Properties of color images of surfaces under multiple illuminants, Opt. Soc. Am. 11 (10) (1994) 2745–2749.

[44] B.A. Maxwell, S.A. Shafer, Physics-based segmentation: moving beyond color, IEEE International Conference on Computer Vision and Pattern Recognition, 1996, pp. 742–749.

[45] G.J. Klinker, S.A. Shafer, Takeo Kanade, Image segmentation and reflection analysis through color, Proceedings of the Image Understanding Workshop, Morgan Kaufmann, San Mateo, CA, 1988, pp. 835–838.

[46] S.L. Guth, Model for color vision and light adaptation, Opt. Soc. Am. 8 (6) (1991) 976–993.

[47] S.A. Shafer, T. Kanade, Using shadows in finding surface orientations, Comput. Vision Graphics Image Process. 22 (1983) 145–176.

[48] G.J. Klinker, S.A. Shafer, T. Kanade, The measurement of highlights in color images, Int. J. Comput. Vision 2 (1) (1988) 7–32.

[49] S. Shafer, T. Kanade, G. Klinker, C. Novak, Physics-based models for early vision by machine SPIE, Perceiving, Measuring, and Using Color, Vol. 1250, Santa Clara, February 1990, pp. 222–235.

[50] R. Bajcsy, S. Wooklee, A. Leonardis, Detection of diffuse and specular interface reflections and inter-reflections by color image segmentation, Int. J. Comput. Vision 17 (3) (1996) 241–272.

[51] M.H. Yang, N. Ahuja, Detecting human faces in color images, IEEE International Conference on Image Processing, 1998, pp. 127–130.

[52] B.A. Maxwell, S.A. Shafer, Physics-based segmentation: looking beyond color, IEEE Computer Vision and Pattern Recognition, June 16–20, 1996, pp. 867–878.

[53] W. Power, R. Clist, Comparison of supervised learning techniques applied to color segmentation of fruit image,

Proceedings of SPIE, Intelligent Roberts and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling, Boston, Massachusetts, November 19–21, 1996, pp. 370–381.

[54] G.A. Hance, S.E. Umbaugh, R.H. Moss, W.V. Stoecker, Unsupervised color image segmentation with application to skin tumor borders, IEEE Eng. Med. Biol. 15 (1) (1996) 104–111.

[55] J. Wu, H. Yan, A.N. Chalmers, Color image segmentation using fuzzy clustering and supervised learning, J. Electron. Image 3 (4) (1994) 397–403.

[56] K.B. Eom, Segmentation of monochrome and color texture using moving average modeling approach, Image Vision Comput. 17 (1999) 233–244.

[57] T. Uchiyama, M.A. Arbib, Color image segmentation using competitive learning, IEEE Trans. Pattern Anal. Mach. Intell. 16 (12) (1994) 1197–1206.

[58] B. Schacter, L. Davis, A. Rosenfeld, Scene segmentation by cluster detection in color space, Department of Computer Science, University of Maryland, College Park, MD, November 1975.

[59] A. Sarabi, J.K. Aggarwal, Segmentation of chromatic images, Pattern Recognition 13 (6) (1981) 417–427.

[60] S.A. Underwood, J.K. Aggarwal, Interactive computer analysis of aerial color infrared photographs, Comput. Graphics Image Process. 6 (1977) 1–24.

[61] J.M. Tenenbaum, T.D. Garvey, S. Weyl, H.C. Wolf, An interactive facility for scene analysis research, Technical Note 87, Artificial Intelligence Center, Stanford Research Institute, Menlo Park, CA, 1974, 230–239.

[62] R. Ohlander, K. Price, D.R. Reddy, Picture segmentation using a recursive region splitting method, Comput. Graphics Image Process. 8 (1978) 313–333.

[63] S. Tominaga, Color image segmentation using three perceptual attributes, IEEE Proceedings of the Conference on Computer Vision and Pattern Recognition, Los Alamos, CA, 1986, pp. 628–630.

[64] G.D. Guo, S. Yu, S.D. Ma, Unsupervised segmentation of color images, IEEE International Conference on Image Processing, 1998, pp. 299–302.

[65] M. Celenk, A color clustering technique for image segmentation, Graphical Models Image Process. 52 (3) (1990) 145–170.

[66] S. Tominaga, Color classification of natural color images, Color Res. Appl. 17 (4) (1992).

[67] Y.W. Lim, S.U. Lee, On the color image segmentation algorithm based on the thresholding and the fuzzy c-means techniques, Pattern Recognition 23 (9) (1990) 935–952.

[68] F. Ferri, E. Vidal, Color image segmentation and labeling through multi-edit condensing, Pattern Recognition Lett. 13 (1992) 561–568.

[69] A. Tremeau, N. Borel, A region growing and merging algorithm to color segmentation, Pattern Recognition 30 (7) (1997) 1191–1203.

[70] H.D. Cheng, Y. Sun, A hierarchical approach to color image segmentation using homogeneity, IEEE Trans. Image Process. (2001), in press.

[71] L. Macaire, V. Ultre, J.-G. Postaire, Determination of compatibility coefficients for color edge detection by relaxation, International Conference on Image Processing, C, 1996, pp. 1045–1048.

[72] P.E. Trahanias, A.N. Venetsanopoulos, Color edge detection using vector order statistics, IEEE Trans Image Process. 2 (2) (1993) 259–265.

[73] P.E. Trahanias, A.N. Venetsanopoulos, Vector order statistics operators as color edge detectors, IEEE Trans. Systems Man Cybernet.-Part B: Cybernetics 26 (1) (1996) 135–143.

[74] F. Perez, C. Koch, Toward color image segmentation in analog VLSI: algorithm and hardware, Int. J. Comput. Vision 12 (1) (1994) 17–42.

[75] S. Ji, H.W. Park, Image segmentation of color image based on region coherency, IEEE International Conference on Image Processing, 1998, pp. 80–83.

[76] J. Luo, R.T. Gray, H.C. Lee, Incorporation of derivative priors in adaptive bayesian color image segmentation, IEEE International Conference on Image Processing, 1998, pp. 780–784.

[77] J. Canny, A computational approach to edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 8 (6) (1986) 679–698.

[78] N. Ito et al., The combination of edge detection and region extraction in non-parametric color image segmentation, Inform. Sci. 92 (1996) 277–294.

[79] Y. Xiaohan, J. Yla, Image segmentation combining region growing with edge detection, 11th International Conference on Pattern Recognition, The Netherland, August 30–September 3 1992, pp. 481–484.

[80] T. Pavlidis, Y.T. Liow, Integrating region growing and edge detection, IEEE Trans. Pattern Anal. Mach. Intell. 12 (3) (1990) 225–233.

[81] Q. Huang et al., Foreground/background segmentation of color images by integration of multiple cues, International Conference on Image Processing, A, 1995, pp. 246–249.

[82] Q. Huang, B. Dom, N. Megiddo, W. Niblack, Segmenting and representing background in color images, International Conference on Pattern Recognition, C, 1996, pp. 13–17.

[83] R.I. Taylor, P.H. Lewis, Color image segmentation using boundary relaxation, IEEE International Conference on Pattern Recognition, C, 1992, pp. 721–724.

[84] J.M. Keller, M.R. Gray, J.A. Givens, A fuzzy K-nearest neighbor algorithm, IEEE Trans. Systems Sci. Cybernet. SMC-15 (1985) 580–585.

[85] J.M. Keller, C.L. Carpenter, Image segmentation in the presence of uncertainty, Int. J. Intell. Systems SMC-15 (1990) 193–208.

[86] L.K. Huang, M.J. Wang, Image thresholding by minimizing the measures of fuzziness, Pattern Recognition 28 (1) (1995) 41–51.

[87] A. De Luca, S. Termini, A definition of a nonprobabilistic entropy in the setting of fuzzy set theory, Inform. Control 20 (1972) 301–312.

[88] N.R. Pal, S.K. Pal, Object-background segmentation using new definition of entropy, IEE Proc. Part E 136 (4) (1989) 284–295.

[89] M.P. Windham, Geometrical fuzzy clustering algorithms, Fuzzy Sets and Systems 10 (1983) 271–279.

[90] S.K. Pal, R.A. King, On edge detection of X-ray images using fuzzy sets, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-5 (1) (1983) 69–77.

[91] D.N. Chun, H.S. Yang, Robust image segmentation using genetic algorithm with a fuzzy measure, Pattern Recognition 29 (7) (1996) 1195–1211.

[92] S.K. Pal, A. Rosenfeld, Image enhancement and thresholding by optimization of fuzzy compactness, Pattern Recognition Lett. 7 (1988) 77–86.

[93] S. Tamura, S. Higuchi, K. Tanaka, Pattern classification based on fuzzy relations, IEEE Trans. System Man Cybernet SMC-1 (1) (1971) 61–66.

[94] R.L. Cannon, J.V. Dave, J.C. Bezdek, Efficient implementation of the fuzzy c-means clustering algorithms, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8 (2) (1986) 249–255.

[95] J.C. Bezdek, P.F. Castelaz, Prototype classification and feature selection with fuzzy sets, Pattern Recognition Lett. 14 (1993) 483–488.

[96] Gath, A.B. Geva, Unsupervised optimal fuzzy clustering, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-11 (7) (1989) 773–781.

[97] T.L. Huntsberger, C. Rangarajan, S.N. Jayaramamurphy, Representation of the uncertainty in computer vision using fuzzy sets, IEEE Trans. Comput. C-35 (2) (1993) 145–156.

[98] M.P. Windham, Geometrical fuzzy clustering algorithms, Fuzzy Sets and Systems 10 (1983) 271–279.

[99] S.K. Pal, R.A. King, Prototype classification and feature selection with fuzzy sets, Electron. Lett. 16 (10) (1993) 376–378.

[100] X.L. Xie, G. Beni, A validity measure for fuzzy clustering, IEEE Trans. Pattern Anal. Mach. Intell. 13 (8) (1991) 841–847.

[101] J.K. Udupa, S. Samarasekera, Fuzzy connectedness and object definition: theory, algorithms and applications in image segmentation, Graphical Models Image Process. 58 (3) (1996) 246–261.

[102] I. Bloch, Fuzzy connectivity and mathematical morphology, Pattern Recognition Lett. 14 (1993) 483–488.

[103] A. Rosenfeld, Fuzzy digital topology, Inform. Control 40 (1) (1979) 76–87.

[104] A. Rosenfeld, The fuzzy geometry of image subsets, Pattern Recognition Lett. 2 (1984) 311–317.

[105] K. Tsuda, M. Minoh, K. Ikeda, Extracting straight lines by sequential fuzzy clustering, Pattern Recognition Lett. 17 (1996) 643–649.

[106] S.Z. Selim, M.A. Ismail, On the local optimality of the fuzzy isodata clustering algorithm, IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8 (2) (1986) 284–288.

[107] S.K. Pal, Image segmentation using fuzzy correlation, Inform. Sci. 62 (1992) 223–250.

[108] A. Moghaddamzadeh, N. Bourbakis, A fuzzy region growing approach for segmentation of color images, Pattern Recognition 30 (6) (1997) 867–881.

[109] T. Carron, P. Lambert, Fuzzy color edge extraction by inference rules quantitative study and evaluation of performances, International Conference on Image Processing, B, 1995, pp. 181–184.

[110] H.D. Cheng, J. Li, Fuzzy homogeneity and scale space approach to color image segmentation, International Conference on Computer Vision, Pattern Recognition and Image Processing, Atlantic City, February 27–March 3, 2000.

[111] H.D. Cheng, X.H. Jiang, Homogram thresholding approach to color image segmentation, International Conference on Computer Vision, Pattern Recognition and Image Processing, Atlantic City, February 27–March 3, 2000.

[112] T. Carron, P. Lambert, Symbolic fusion of hue-chroma-intensity features for region segmentation, International Conference on Image Processing, B, 1996, pp. 971–974.

[113] A.E. Pienkowski, T.J. Dennis, Applications of fuzzy logic to artificial color vision, SPIE Comput. Vision Robots 595 (1985) 50–55.

[114] W. Pedrycz, Fuzzy sets in pattern recognition: methodology and methods, Pattern Recognition 23 (1/2) (1990) 121–146.

[115] C.W. Tao, W.E. Thompson, A fuzzy if–then approach to edge detection. FUZZY-IEEE 93, San-Francisco, USA, 1993, pp. 1356–1360.

[116] A. Moghaddamzadeh, N. Bourbakis, A fuzzy technique for image segmentation of color images, IEEE Word Congress on Computational Intelligence: FUZZY-IEEE, Orlando, Florida, June 1994.

[117] Y.S. Chen, H.Y. Hwang, B.T. Chen, Color image analysis using fuzzy set theory, International Conference on Image Processing, A, 1995, pp. 242–245.

[118] A. Moghaddamzadeh, N.G. Bourbakis, Segmentation of color images with highlights and shadows using fuzzy reasoning, IS & T/SPIEs Symposium, Electronic Imaging: Science and Technology, San Jose, CA, February 1995, pp. 5–10.

[119] A. Moghaddamzadeh, N. Bourbakis, A fuzzy approach for smoothing and edge detection in color images, IS & T/SPIE Symposium, Electronic Imaging: Science and Technology, San Jose, CA, February 1995.

[120] M. Mari, S. Dellepiane, A segmentation method based on fuzzy topology and clustering, IEEE International Conference on Pattern Recognition, B, 1996, pp. 565–569.

[121] C.K. Ong, T. Matsuyama, Robust color segmentation using the dichromatic reflection model, IEEE International Conference on Pattern Recognition, 1998, pp. 780–784.

[122] C.-L Huang, Pattern image segmentation using modified Hopfield model, Pattern Recognition Lett. 13 (1999) 345–353.

[123] P. Campadelli, D. Medici, R. Schettini, Color image segmentation using Hopfield networks, Image Vision Comput. 15 (1997) 161–166.

[124] M. Sammouda, R. Sammouda, N. Niki, K. Mukai, Segmentation and analysis of liver cancer pathological color image based on artificial neural networks, IEEE 1999 International Conference on Image Processing, October 24–28, 1999, Kobe, Japan, pp. 392–396.

[125] R. Sammouda, N. Niki, H. Nishitani, Segmentation of sputum color image for lung cancer diagnosis based on neural networks, IEICE Trans. Inform. Systems E81-D (8) (1998) 862–871.

[126] H. Iwata, H. Nagahashi, Active region segmentation of color images using neural networks, Systems Comput. J. 29 (4) (1998) 1–10.

[127] J. Vesanto, E. Alhoniemi, Clustering of the self-organizing map, IEEE Trans. Neural Networks 11 (3) (2000) 586–600.

[128] S. Ji, H.W. Park, Image segmentation of color image based on region coherency, 1998 International Conference on Image Processing, Chicago, Illinois, USA, October 4–7 1998, pp. 80–83.

[129] Y.S. Lo, S.C. Pei, Color image segmentation using local histogram and self-organization of Kohonen feature map, 1999 International Conference on Image Processing, KOBE, Japan, October 24–28 1999, pp. 232–239.

[130] P. Lescure, V. Meas-Yedid, H. Dupoisot, G. Stamon, Color segmentation on biological microscope images, Proceeding of SPIE, Application of Artificial Neural Networks in Image Processing IV, San Jose, California, January 28–29, 1999, pp. 182–193.

[131] R. Rae, H.J. Ritter, Recognition of human head orientation based on artificial neural networks, IEEE Trans. Neural Network 9 (2) (1998) 257–265.

[132] C.Y. Ho, H. Kurokawa, A learning algorithm for oscillatory cellular neural networks, Neural Networks 12 (1999) 825–836.

[133] F. Kurugollu, B. Sankur, MAP segmentation of color images using constraint satisfaction neural network, IEEE 1999 International Conference on Image Processing, KOBE, Japan, October 24–28 1999, pp. 236–239.

[134] R. Schettini, A segmentation algorithm for color images, Pattern Recognition Lett. 14 (1993) 499–506.

[135] C.L. Huang, T.Y. Cheng, C.C. Chen, Color images segmentation using scale space filter and markov random field, Pattern Recognition 25 (10) (1992) 1217–1229.

[136] Y.N. Sun, C.S. Wu, X.Z. Lin, N.H. Chou, Color image analysis for liver tissue classification, Opt. Eng. 32 (7) (1993) 1609–1614.

[137] J.Q. Liu, Y.H. Yang, Multiresolution color image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. 16 (7) (1994) 689–700.

[138] T. Caelli, D. Reye, On the classification of image regions by color, texture and shape, Pattern Recognition 26 (4) (1993) 461–470.

[139] K. Valkealahti, E. Oja, Reduced multidimensional histograms in color texture description, International Conference on Pattern Recognition, 1998, pp. 1057–1061.

About the Author—HENG-DA CHENG received Ph.D. degree in Electrical Engineering from Purdue University, West Lafayette, IN, in 1985 (supervisor: K.S. Fu). Now he is Full Professor, Department of Computer Science, and an Adjunct Full Professor, Department of Electrical Engineering, Utah State University, Logan, UT. Dr. Cheng also holds a Visiting Professor position of Northern Jiaotong University, China.

Dr. Cheng has published more than 180 technical papers, is the co-editor of the book, *Pattern Recognition: Algorithms, Architectures and Applications* (World Scientific, 1991), and the editor for a couple conference proceedings. His research interests include image processing, pattern recognition, computer vision, artificial intelligence, medical information processing, fuzzy logic, genetic algorithms, neural networks, parallel processing, parallel algorithms, and VLSI architectures.

Dr. Cheng was the general chair and program chair of the Third International Conference on Computer Vision, Pattern Recognition and Image Precessing (CVPRIP'2000), 2000; was the general chair and program chair of the First International Workshop on Computer Vision, Pattern Recognition and Image Processing (CVPRIP'98), 1998, and was the Program Co-Chairman of Vision Interface '90, 1990. He served as a program committee member and session chair for many conferences, and as a reviewer for many scientifical journals and conferences.

Dr. Cheng has been listed in *Who's Who in the World*, *Who's Who in America*, *Who's Who in Communications and Media*, *Who's Who in Finance and Industry*, *Who's Who in Science and Engineering*, *Men of Achievement*, 2000 *Notable American Men*, *International Leaders in Achievement*, *Five Hundred Leaders of Influence*, *International Dictionary of Distinguished Leadership*, *etc*. He is appointed as a *Member of the Advisory Council*, the International Biographical Center, England, and a *Member of the Board of Advisors*, the American Biographical Institute, USA.

Dr. Cheng is a Senior Member of the IEEE Society, and a Member of the Association of Computing Machinery. Dr. Cheng is also an Associate Editor of *Pattern Recognition* and an Associate Editor of *Information Sciences*.

About the Author—XIHUA JIANG received the Bachelor of Science degree from the Department of Computer and System Science, Nankai University, Tianjin, China, in 1988, the Master of Science degree from the Department of Computer Science and Technology, Peking University, Beijing, China, in 1997, and the Master of Science degree from the Department of Computer Science, Utah State University, Logan, UT, in 1999.

From 1988 to 1994, he was a research assistant and software engineer in the Department of Computer and System Science, Nankai University, Tianjin, China, and participated in the design, development, and implementation of several robot control systems for the national hi-tech plan in China. Since 1999, he has been with the Software Engineering Department at Citrix Systems, Inc. as a software engineer. His research interests are fuzzy logic, image processing and server-client computing.

About the Author—YING SUN received the Bachelor of Science degree in computer software in 1989 from the Department of Computer and System Science, Nankai University, Tianjin, China, and the Master of Science degree in computer science from Utah State University, Logan, UT, in 1999.

From 1991 to 1995, she was a software engineer in the Computing Center of Tianjin Foreign Trade and Economic Relations Committee, Tianjin, China, and participated in the design, development, and implementation of several financial management information systems for the state-run foreign trade companies in Tianjin, China. From 1996 to 1997, she joined Motorola (China) Electronics Ltd. and worked in Management Information System Department, Cellular Subscriber Sector, as an application analyst. She conducted software project design, plan, program, and trouble-shooting. Since 1999, she has been with the information Technology

Group, Personal Communication Systems, Motorola Inc. in Libertyville, ILL, as a software engineer. Her research interests include computer algorithms and image processing methodology.

About the Author—JINGLI WANG received the bachelor of science degree in computer science in 1990 from the Computer Science Department, Zhengzhou University, China. From 1990 to 1999, she was a software engineer in Chinese Academy of Sciences, participated in remote sensing instruments design, development, and implementation, worked on remote sensing information processing, such as real-time control and data processing, simulation, and calibration for the satellite and airborne altimeter and conducted software project designing, planning and programming. Since 1999, she has been a graduate student in Utah State University with a major in Computer Science. Her research interests include software engineering and image processing methodology.