Colour Image Segmentation
– A Survey –

Władysław Skarbek

Institute of Computer Science

Polish Academy of Sciences

and

Andreas Koschan

Institute for Technical Informatics

Technical University of Berlin

October 1994

Kurzfassung

In den letzten drei Jahrzehnten haben sich eine Reihe bedeutener Forschungsaktivitäten mit dem Problem der Segmentierung von Bildern, d.h. mit der Identifizierung homogener Regionen im Bild befaßt. Viele Algorithmen wurden hierbei für die Segmentierung von Grauwertbildern entwickelt. Der Bereich der Segmentierung von Farbbildern hat demgegenüber weit weniger Beachtung in der Wissenschaft erfahren, obwohl Farbbilder wesentlich mehr Informationen über die Objekte einer Szene enthalten als Grauwertbilder. Während einige Übersichtsartikel über Verfahren zur Segmentierung monochromer Bilder veröffentlicht wurden, existieren ähnliche umfaßende Übersichten für Farbbilder nach unserem Kenntnisstand nicht. Dieser Bericht enthält: eine ausführliche Übersicht über Algorithmen für die Segmentierung von Farbbildern, eine Kategorisierung der Verfahren hinsichtlich einer Liste von Attributen, Anregungen für ihre Verbesserung und Beschreibungen von einigen neuen Ansätzen.

Abstract

Image segmentation, i.e., identification of homogeneous regions in the image, has been the subject of considerable research activity over the last three decades. Many algorithms have been elaborated for gray scale images. However, the problem of segmentation for colour images, which convey much more information about objects in scenes, has received much less attention of scientific community. While several surveys of monochrome image segmentation techniques were published, similar comprehensive surveys for colour images, to our knowledge, did not emerge.

This report contains: an extensive survey of algorithms for colour image segmentation, a categorization of them according well defined list of attributes, suggestions for their improvements, and descriptions of few novel approaches.

Contents

1 Introduction

Image segmentation, i.e. identification of homogeneous regions in the image, has been the subject of considerable research activity over the last three decades. Many algorithms have been elaborated for gray scale images. However, the problem of segmentation for colour images, which convey much more information about objects in scenes, has received much less attention of scientific community. While several surveys of monochrome image segmentation techniques were published, similar surveys for colour images did not emerge.

This report contains: an extensive survey of algorithms for colour image segmentation, a categorization of them according well defined list of attributes, suggestions for their improvements, and descriptions of few novel approaches.

In categorization of colour image segmentation algorithms the following definition will be useful:

Colour image segmentation is a process of extracting from the image domain one or more connected regions satisfying uniformity (homogeneity) criterion which is based on feature(s) derived from spectral components. These components are defined in a chosen colour space model. The segmentation process could be augmented by some additional knowledge about the objects in the scene such as geometric and optical properties.

Perhaps the most important feature of the given segmentation method is *region definition*. We can identify four broad types of such definitions:

1. *Region* is a connected component of a pixel set specified by a class membership function defined in a colour space; For instance:

   (a) pixel colour is in the given half space defined by a plane;

   (b) pixel colour is in the given polyhedra;

   (c) pixel colour is in the Voronoi cell with the given representative;

   (d) fuzzy membership function for a class defuzzified by α-cutoff operation;

   Observe that generally the membership functions in colour spaces define multi component sets of pixels. However, if the volume in colour space specifying the membership is small, then intuitively, the region is *uniform*.

2. *Region* is a (maximal) connected set of pixels for which uniformity condition is satisfied. For instance:

   (a) uniform region derived by growing from a seed block by joining other pixels or blocks of pixels;

   (b) uniform region obtained by splitting a larger region, which is not uniform.

3. *Region* is a connected set of pixels bounded by edge pixels creating a colour contour. For instance edge detection by vector median or Hueckel operator

with model guided filling of gaps in contours. Observe also that in a sense regions here are *uniform* too, because they are in the complementary set of *nonuniform* set which edge pixels create.

4. *Region* corresponds to a surface or an object of homogeneous material, i.e., regions represent connected material in the scene. For instance shading, shadow, and highlight have no influence on the result of image segmentation although they change the colour values in the image.

Logically, region definition types 1 and 2 use some sort of *uniformity predicate* which in case 1 is single pixel based while in the case 2 it is area based. Definitions of type 3 use a *nonuniformity predicate*. Thus, this classification is consistent and complete. Region definition of type 4 represents a supplementary sub-class of regions which could also be treated as a special case of regions already defined by one of the other three types. The reason for introducing this additional region definition is that the objective of image segmentation and the assumptions made on the objects in the scene are different for region type 4 as compared to the other three types. None of the segmentation techniques using region type 4 is an extension of an intensity based approach. Contrarily, these methods can be exclusively applied to colour images.

Furthermore, segmentation techniques employing region definition of type 4 belong to a new class of computer vision methods that are categorised during the past few years as *physics based vision methods*. The reader of this survey can easily find these techniques due to this supplementary region definition.

Shortly, we call type 1 as *pixel based* region definitions, type 2 as *area based* region definitions, type 3 as *edge based* definitions, and type 4 as *physics based* definitions.

While categories 2 and 3 are very well known from the literature (see for instance the survey of Pal and Pal [54]), the category 1 embraces such techniques as histogramming and clustering. The category 4 uses reflection models based on the properties of the material in the scene. The last category has not yet been paid proper attention in the existing surveys on image segmentation.

Other important aspect of the segmentation method is the *colour space* from which *colour features* are computed (for instance RGB space with Euclidean colour distance).

Each segmentation technique is usually based on some mathematical model (theory) and/or algorithmic approach (for instance fuzzy clustering, Markov random field, recursive procedure, bottom–up algorithm etc.). Further we specify it under common heading *basic mathematical method(s)*.

Most of segmentation techniques assume something about the scene which is seen in the image (for instance objects are polyhedra made of dielectric materials). This is an *additional knowledge* attribute of the given segmentation method.

Frequently, but not always, the specified segmentation method was designed for some real life *applications*. This the last aspect we like to discuss while characterizing a particular technique.

In summary, the following attributes will be evaluated for reviewed segmentation techniques: **region definition, basic mathematical method, colour space, additional knowledge, applications**. It is also intended to identify chromatic features which could be of use in computational stereo vision such as colour edges (see for instance Jordan and Bovik[39]).

For clear comparison of research contributions and for further referencing it was decided to present papers through *metrics* in which above listed attributes (extended by **year, author**) are evaluated and then followed by a **description** with more details. The **title** of the paper is put into heading of the subsection. Own solutions are marked by $* * **$ in the **year** field.

2 Colour spaces

Many colour spaces are in use today. For pictures acquired by digital cameras the most popular is **RGB model**. According the tristimulus theory (Wyszecki and Stiles [79]) colour can be represented by three components, resulting from three separate colour filters S_X, $X = R, G, B$, on light radiance $E(\lambda)$ according the equations:

$$R = \int_\lambda E(\lambda) S_R(\lambda) d\lambda, \ G = \int_\lambda E(\lambda) S_G(\lambda) d\lambda, \ B = \int_\lambda E(\lambda) S_B(\lambda) d\lambda \ . \qquad (1)$$

In segmentation, reducing of dependence on changing in space lighting intensities is a desirable goal. If variations of intensities are uniform across the spectrum then **normalised RGB space** is of value:

$$r = \frac{R}{R+G+B}, \ g = \frac{G}{R+G+B}, \ b = \frac{B}{R+G+B} \ . \qquad (2)$$

Note that $r + g + b = 1$, so some authors (for instance Nevatia [50, 51]) use rg components only, adding independent one such as a sort of luminance

$$y = c_1 R + c_2 G + c_3 B.$$
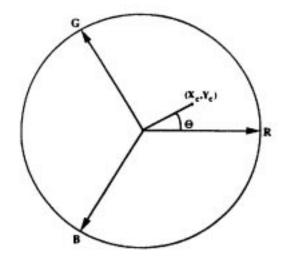
There are many colour models based on human colour perception. Such models refer to **hue, saturation**, and **intensity** components. Most of them follow Munsell colour system (Wyszecki [79]). For instance **HSI model** is transformed from RGB model using the following formulas:

$$
\begin{aligned}
I &= (R+G+B)/3 \ , & (3) \\
S &= 1 - \min(R, G, B)/I \ , & (4) \\
H &= \arctan'(\sqrt{3}(G-B), 2R-G-B) & (5)
\end{aligned}
$$

where $\arctan'(y, x)$ gives the angle between horizontal axis and the line $(0,0)-(x,y)$. While I and S are somewhat arbitrarily modelled, the formula for H can be easily derived if we make a projection of RGB axis onto plane $R+G+B = 0$ as shown in the following figure:
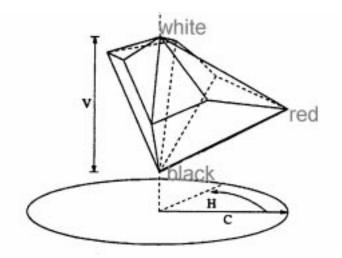
If hue absolute differences are to be calculated then special modular arithmetic must be used:

$$h_1 \ominus h_2 \doteq \begin{cases} |h_1 - h_2| & \text{if } |h_1 - h_2| \leq \pi, \\ \pi - |h_1 - h_2| & \text{otherwise.} \end{cases} \tag{6}$$

Since at small intensity and at small saturation, the hue is very imprecisely determined (see the proof below), then any colour comparisons should avoid hue subtractions in this situation. A practical solution is using then, the absolute differences of the intensity.

Another less known, *human oriented* colour space is TekHVCT model, which uses (V, C, H) coordinates modelling intensity, saturation, and hue, respectively:

If we map (intensity, saturation, hue) (say (i, s, h)) coordinates onto Munsell colour solid volume then Euclidean distance between colours (i_1, s_1, h_1) and (i_2, s_2, h_2) could be attempted for intensities which are greater than 10% of its range (25 at

[0,255] range):

$$\Delta C \doteq \sqrt{(i_1 - i_2)^2 + s_1^2 + s_2^2 - 2c_1 c_2 \cos(h_1 - h_2)} \; . \tag{7}$$

RGB tristimulus coordinates exclude some visible colours. They also depend on physical sensors. For this reason an international body, the CIE committee fixed **XYZ tristimulus** coordinates. They can be produced from RGB tristimulus coordinates by a linear transformation. However, the transformation matrix must be determined empirically. For instance, the matrix for the NTSC receiver primary system is:

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

If the XYZ coordinates are known, different CIE spaces can be constructed. It was verified that the **CIE(Lab) space** is perceptually uniform and gives good results for segmentation of colour pictures:

$$L = 25(100Y/Y_0)^{1/3} - 16 \; , \tag{8}$$

$$a = 500 \left[(X/X_0)^{1/3} - (Y/Y_0)^{1/3} \right] \; , \tag{9}$$

$$b = 200 \left[(Y/Y_0)^{1/3} - (Z/Z_0)^{1/3} \right] \tag{10}$$

where (X_0, Y_0, Z_0) are XYZ values for the reference white.

Another CIE colour space used in the field is **CIE(Luv) space**:

$$u = 13W(4X/(X + 15Y + 3Z) - 0.199 \tag{11}$$

$$v = 13W(6Y/(X + 15Y + 3Z) - 0.308 \tag{12}$$

Few authors use TV colour systems, **YIQ** of American system

$$\begin{bmatrix} Y \\ I \\ Q \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ 0.596 & -0.275 & -0.321 \\ 0.212 & -0.528 & 0.311 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{13}$$

and **YUV** of European system:

$$\begin{bmatrix} Y \\ U \\ V \end{bmatrix} = \begin{bmatrix} 0.299 & 0.587 & 0.114 \\ -0.147 & -0.289 & 0.437 \\ 0.615 & -0.515 & -0.100 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \tag{14}$$

¿From CIE coordinates there is a simple way to get hue and saturation coordinates:

$$\mathrm{sat} = \sqrt{a^2 + b^2} \; , \tag{15}$$

$$\mathrm{hue} = \arctan'(b, a) \; . \tag{16}$$

For CIE models as the **measure of two colours distance** the Euclidean distance is used. However, this is experimentally justified only for CIE(Luv) and CIE(Lab) spaces.

3 Properties of chromatic features

Perez and Koch have shown ([56]) advantages and some disadvantages of hue and normalised RGB components. Their conclusions are based on simplified models of light propagation used in computer graphics. However, experiments validate these assumptions with good degree of accuracy.

3.1 Uniform scaling and shifting invariance

By simple algebraic manipulations it is easy to prove the following four facts:

1. Hue is invariant at uniform scaling of RGB components:

$$H(\alpha R, \alpha G, \alpha B) = H(R, G, B) .$$

2. Normalised rgb components are invariant at uniform scaling of RGB components:

$$
\begin{aligned}
r(\alpha R, \alpha G, \alpha B) &= r(R, G, B) , \\
g(\alpha R, \alpha G, \alpha B) &= g(R, G, B) , \\
b(\alpha R, \alpha G, \alpha B) &= b(R, G, B) .
\end{aligned}
$$

3. Hue is invariant at uniform shifting of RGB components:

$$H(R + \beta, G + \beta, B + \beta) = H(R, G, B) .$$

4. Normalised rgb component are not invariant at uniform shifting of RGB components:

$$
\begin{aligned}
r(R + \beta, G + \beta, B + \beta) &\neq r(R, G, B) , \\
g(R + \beta, G + \beta, B + \beta) &\neq g(R, G, B) , \\
b(R + \beta, G + \beta, B + \beta) &\neq b(R, G, B) .
\end{aligned}
$$

3.2 Hue singularity near RGB zero

When $R = G = B = 0$, we see from the formula 5 that hue is undefined.

Moreover, a deviation of the colour point from $(0, 0, 0)$ to $((1 - \alpha)\epsilon, \alpha\epsilon, 0)$ gives at fixed weighting coefficient α and any ϵ the value of hue equal:

$$H = \arctan'(\sqrt{3}\alpha, 2 - 3\alpha) .$$

It implies that for even small ϵ the hue can vary from zero (for $\alpha = 0$) to $2\pi/3$ (for $\alpha = 1$).

In discrete computer representation (R, G, B), the least change occurs in unit steps on each component. Therefore more convincing proof of hue singularity near zero is taking $H(1, 0, 0) = 0$ and $H(0, 1, 0) = 2\pi/3$.

The change of colour from $(x, x, 0)$ to $(x+1, x, 0)$ shifts hue H by angle δ satisfying the following relation:
$$\cos(\delta) = \frac{2x + 1}{2\sqrt{x^2 + x + 1}} \ .$$

Hence for $x \geq 25$ the hue error is very small. The biggest error occurs for $x = 1$ reaching $\delta = \pi/6$.

In conclusion: calculation of hue for RGB colour with low intensity bears significant numerical errors and therefore it is not recommended.

3.3   Discounting transparency

On the distance d the light of wavelength λ is attenuated by $\tau(\lambda)^d$. If the transparent medium is uniform then τ is constant and the red components has the form:
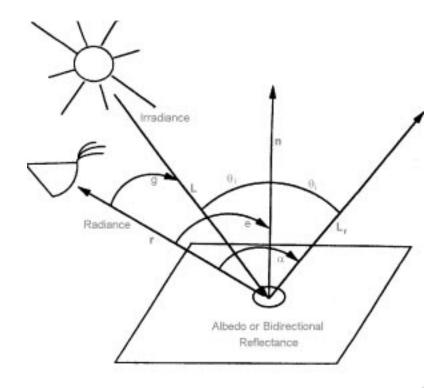
$$R = \int_\lambda E(\lambda) S_R(\lambda) d\lambda, \quad \hat{R}(d) = \int_\lambda E(\lambda) S_R(\lambda) \tau(\lambda)^d d\lambda = \tau_0^d R \ .$$

Hence at passing the uniformly transparent medium the RGB components are uniformly scaled. Therefore both *hue and normalised RGB are invariant at passing of light through uniformly transparent media.*

3.4   Discounting highlights

Some materials can interact (locally or globally) with light in a specular way (for instance glass). Then at some light falling angles, highlights may occur. Patches of highlights in the image create a serious problem for segmentation algorithms. Applying Phong's model for image formation Perez and Koch have shown that using hue can discount this undesirable effect.

The geometry of image formation in Phong's ([57]) model is shown in the following figure:

Irradiance

Radiance

Albedo or Bidirectional
Reflectance

In Phong's shading equations for the tristimulus of the radiance coming to the observer from the given surface small patch consists of the ambient, diffusion, and specular part:

$$I_R = k_{aR}I_{aR} + \frac{I_{lR}}{r+k}[k_{dR}\cos(\theta) + k_s\cos^n(\alpha)] \tag{17}$$

$$I_G = k_{aG}I_{aG} + \frac{I_{lG}}{r+k}[k_{dG}\cos(\theta) + k_s\cos^n(\alpha)] \tag{18}$$

$$I_B = k_{aB}I_{aB} + \frac{I_{lB}}{r+k}[k_{dB}\cos(\theta) + k_s\cos^n(\alpha)] \tag{19}$$

If we assume a white light source then the specular part for each RGB component is identical and equal to:

$$\beta = \frac{I_l}{r+k}k_s\cos^n(\alpha) \ .$$

It means that non highlighting component is uniformly shifted. It implies that *hue is invariant for highlights at white light sources* while normalised RGB is not invariant.

If we generalize the Phong equation to arbitrary wavelength we get:

$$I(\lambda) = I_a(\lambda)k_a(\lambda) + \frac{I_l(\lambda)}{r+k}[k_d(\lambda)\cos(\theta) + k_s\cos^n(\alpha)] \ .$$

According the tristimulus equations 1 we decompose red highlight component as follows:

$$R_H = R_{NH} + \frac{k_s\cos^N(\alpha)}{r+k}\int_\lambda I_l(\lambda)S_R(\lambda)d\lambda \ .$$

To get uniform shift for RGB we have to impose the following constraint (called the *integrated white condition*):

$$\int_\lambda I_l(\lambda)S_R(\lambda)d\lambda = \int_\lambda I_l(\lambda)S_G(\lambda)d\lambda = \int_\lambda I_l(\lambda)S_B(\lambda)d\lambda \ . \tag{20}$$

If the light source is white, the integrated white condition has the form:

$$\int_\lambda S_R(\lambda)d\lambda = \int_\lambda S_G(\lambda)d\lambda = \int_\lambda S_B(\lambda)d\lambda \ . \tag{21}$$

In artificial systems this condition may be implemented by carefully designed spectral filters.

Finally we conclude that within generalized Phong's shading model the *hue discounts highlights in the scene if the integrated white condition is true.*

3.5   Discounting shading and shadowing

Change of surface orienation is also a source of confusion in achromatic vision systems.

Assuming Phong's shading model and nonspecular object material (matte surfaces) we can eliminate the RGB dependence on surface orientation in some well defined circumstances.

Let two close surface points are illuminated at angles θ_1 and θ_2, respectively.

If we can ignore ambient light, then from tristimulus equations we get relations for RGB values (R_1, G_1, B_1) (R_2, G_2, B_2) viewed at these two points:

$$\begin{aligned}
R_2 &= R_1 \frac{\cos\theta_2}{\cos\theta_1}\frac{r+k}{r+\Delta r+k} \ , \\
G_2 &= G_1 \frac{\cos\theta_2}{\cos\theta_1}\frac{r+k}{r+\Delta r+k} \ , \\
B_2 &= B_1 \frac{\cos\theta_2}{\cos\theta_1}\frac{r+k}{r+\Delta r+k} \ .
\end{aligned}$$

Hence colours for two points are related by scaling operation. We conclude that *for matte surfaces while ignoring ambient light, both normalised RGB and hue are invariant to changes of surface orientation relatively to the light source.*

If we cannot ignore ambient light, then the colour relations have the form:

$$\begin{aligned}
R_2 &= R_1 + \int_\lambda k_d(\lambda)I_l(\lambda)S_R(\lambda)d\lambda \left(\frac{\cos\theta_1}{r+k} - \frac{\cos\theta_2}{r+\Delta r+k}\right) \ , \\
G_2 &= G_1 + \int_\lambda k_d(\lambda)I_l(\lambda)S_G(\lambda)d\lambda \left(\frac{\cos\theta_1}{r+k} - \frac{\cos\theta_2}{r+\Delta r+k}\right) \ , \\
B_2 &= B_1 + \int_\lambda k_d(\lambda)I_l(\lambda)S_B(\lambda)d\lambda \left(\frac{\cos\theta_1}{r+k} - \frac{\cos\theta_2)}{r+\Delta r+k}\right) \ .
\end{aligned}$$

We get uniform shifting of RGB colour if a sort of *integrated white condition* is true:

$$\int_\lambda k_d(\lambda) I_l(\lambda) S_R(\lambda) d\lambda = \int_\lambda k_d(\lambda) I_l(\lambda) S_G(\lambda) d\lambda = \int_\lambda k_d(\lambda) I_l(\lambda) S_B(\lambda) d\lambda \ . \qquad (22)$$

If diffusion properties of the material are spectrally uniform then this condition is reducing to the previous *integrated white condition.*

Since we proved that colours are uniformly shifted at the change of surface orientation, then it gives only *hue invariance to shading if ambient light is significant.* Normalised RGB colours are not invariant in these circumstances.

An analysis for shadowing is similar to the analysis for shading. If we assume that the colour (R_1, B_1, G_1) comes from the shadowed point of a surface and the colour (R_2, B_2, G_2) from nearby point which is not in a shadow, then for matte surfaces we get:

$$
\begin{aligned}
R_2 &= R_1 + \int_\lambda k_d(\lambda) I_l(\lambda) S_R(\lambda) d\lambda \ \frac{\cos\theta_2}{r + \Delta r + k} \ , \\
G_2 &= G_1 + \int_\lambda k_d(\lambda) I_l(\lambda) S_G(\lambda) d\lambda \ \frac{\cos\theta_2}{r + \Delta r + k} \ , \\
B_2 &= B_1 + \int_\lambda k_d(\lambda) I_l(\lambda) S_B(\lambda) d\lambda \ \frac{\cos\theta_2}{r + \Delta r + k} \ .
\end{aligned}
$$

We get the uniform shifting of colour when the same condition 22 is satisfied. Therefore *hue is invariant to shadows if the material is matte, generalized integrated white condition is valid, and Phong's image formation model can be accepted.* As for shading if material is spectrally uniform in light diffusion, then integrated white condition is enough.

Experiments ([56]) show that for typical scenes the change of hue on the shadow edge is small, i.e. about 3–5%.

4 Pixel based segmentation

In this section we discuss segmentation techniques which operate in colour space. Available papers can be broadly divided into three groups:

1. **histogram based techniques**: one or more peaks are identified, surrounding intervals are next utilized in pixel classification process;

2. **segmentation by clustering data in colour space**: pixel values are collected into groups with one or more representatives which next are used in pixel classification process;

3. **segmentation by fuzzy clustering**: fuzzy membership functions is evaluated for all pixels and for all fuzzy clusters defined; hard clusters of pixels are obtained by defuzzification process and next subdivided into maximal connected regions.

4.1 Thresholding histograms

4.1.1 Stable segmentation using colour information

Author(s): L. Bonsiepen and W. Coy ([10]).

Year: 1991.

Region definition(s): Pixels with a colour scalar feature below the threshold belong to the background, others belong to the object.

Basic mathematical method(s): Histogramming and single threshold computing in bimodal histograms.

Colour space(s): Normalised RGB from where a scalar feature F is extracted:

$$F = (0.25r + 4g - b)/5 \ .$$

Additional knowledge: Only lichens (creating a connected region in the image) and the background are in the image.

Application(s): Colour and speed of growth of liechens population (measured by the area of the region occupied by them) are used in environmental research.

Description: The feature F was found on the basis of 300 pictures by considering about 1000 colour features. The authors do not describe the method which was used to extract F, but for instance a neural network with single a perceptron would solve this problem. The threshold was found in minimum point of very well seperated bimodal histogram built for F.

4.1.2 Picture segmentation using a recursive region splitting method

Author(s): R. Ohlander, K. Price, and D.R. Reddy ([52]).

Year: 1978.

Region definition(s): The peak of the histogram in one of nine colour features determines the interval. Pixels falling in this interval create one region and pixels falling out of it another one. Both are further splitted recursively. The connected region with no clear peak is output.

Basic mathematical method(s): Histogramming of many colour features. Best peak selection. Decomposing of the region into connected components.

Colour space(s): Nine features collected from colour systems RGB, HSI, and YIQ.

Additional knowledge: None.

Application(s): General purpose.

Description: The peak selection is driven by a priority list of seven, rather arbitrary fixed conditions. Firstly all peaks in the set of histograms are located. The list of peaks of the lowest priority is built. The best peak on this list is determined and threshold values are chosen on each side of this peak.

The general segmentation scheme consists of steps:

0. Put the image domain into initially empty region list;
1. Get next region to segment. If no more then STOP;
2. Select best peak. If none then output this uniform region and Go to step 1.
3. Apply threshold;
4. Select connected regions and add them to the list of regions;
5. Go to step 1;

The authors suggested several improvements for the basic scheme: removing of small regions, work on reduced image version, adding textural features (e.g.: density of edges).

4.1.3 Colour information for region segmentation

Author(s): Y.I. Ohta, T. Kanade, and T. Sakai ([53]).

Year: 1980.

Region definition(s): As in Ohlander's paper, but with three features I_1, I_2, I_3 only.

Basic mathematical method(s): KL transform, i.e. principal component analysis in order to extract most important colour features. Other as in Ohlander's paper.

Colour space(s): Transformed RGB to $I_1 I_2 I_3$:

$$I_1 = (R + G + B)/3, \ I_2 = R - B, \ I_3 = (2G - R - B)/2 \ .$$

Additional knowledge: None.

Application(s): General purpose.

Description: I_1, I_2, I_3 model was obtained by an experimental evidence that it is good approximation for KL transform.

Essentially the algorithm is very similar to Ohlander's one, but its presentation is more clear and suggests data structures for more efficient implementation in both time and space aspects:

0. A mask corresponding to the whole image is placed at the bottom of the stack;

1. If stack is empty then STOP. Otherwise one mask is taken from the top of the stack. Let S denotes the region reprented by this mask;

2. Histograms of colour features in the region S are computed;

3. If any of the histograms show conspicuous peaks, a pair of cutoff values which seperate the peak in the histogram are determined at the positions of valleys, and the image of the colour feature corresponding to that histogram is thresholded using the cutoff values; thus the region S is partitioned. Otherwise output S and go to step 1;

4. Connected regions are extracted. For each connected region, a region mask is generated and it is pushed on the stack; Go to step 1.

4.1.4 Opponent colours as a 2-dimensional feature within a model of the first stages of the human visual system

Author(s): K. Holla ([32]).

Year: 1982.

Region definition(s): The mountains in the 2-d histogram of the opponent colour pairs determine areas. Pixels falling into one of these areas create one region. Pixels falling into another area create another region.

Basic mathematical method(s): Finding peaks and basis in the 2-d histogram of the opponent colour pairs.

Colour space(s): Opponent colour space separating the signal into luminance and chrominance. The chrominance plane is given by the opponent colour pairs red-green and yellow-blue.

Additional knowledge: None.

Application(s): General purpose.

Description: RGB values are transformed to the opponent colour pairs red-green (RG), yellow-blue (YB), and the intensity function (I). The three channels are smoothed by applying band-pass filters where the centre frequencies of the filters dispose of the proportion I : RG : YB = 4 : 2: 1. Then peaks and basis in the two-dimensional RG-YB-histogram are searched for. Peaks and basis points determine areas in the RG-YB-plane. Pixels falling into one of these areas create one region. Pixels falling into another area create another region. Due to this definition there remain some non-attachable parts in the image. Holla suggests to include additional features as luminance or the local connection of pixels into the segmentation process to enhance the result.

4.1.5 Segmentation of colour pictures with the aid of colour information and spatial neighbourhoods

Author(s): : H.-D. vom Stein and W. Reimers ([68]).

Year: 1983.

Region definition(s): The mountains in the 2-d histogram of the opponent colour pairs determine areas. Pixels falling into one of these areas create one region. Pixels falling into another area create another region.

Basic mathematical method(s): Finding peaks and basis in the 2-d histogram of the opponent colour pairs and merging of pixels due to spatial neighbourhood relations.

Colour space(s): Opponent colour pairs red-green and yellow-blue.

Additional knowledge: None.

Application(s): General purpose.

Description: The segmentation algorithm is a modification of the approach of Holla ([32]) which has been described in the previous subsection. The modification consists in an additional refinement process that is employed to the segmentation results obtained with Holla's technique. If one or more pixels in the 3x3 neighbourhood of a non- assigned pixel are assigned to the same region, the pixel is marked for assignment to this region. No decision is made if none of the pixels in the 3x3 neighbourhood is assigned or if several pixels

19

in the 3x3 neighbourhood belong to different regions. After the entire image is scanned the marked pixels are assigned to the corresponding regions. This procedure is applied five-times to the intermediate results. While 30% to 80% of the pixels are assigned to regions when employing Holla's approach, less than 10% are not assigned to regions when using this modification.

4.1.6 Colour image segmentation using modified HSI system for road following

Author(s): X. Lin and S. Chen ([47])

Year: 1991

Region definition(s): Two class membership function based on threshold value of one scalar feature V.

Basic mathematical method(s): Histogram based segmentation.

Colour space(s): Modified HSI, RGB.

Additional knowledge: Two objects (road and non-road).

Application(s): Robot vehicles.

Description: Lin and Chen selected the HSI space for road following and compared the results with those computed in the RGB space. The objective of colour segmentation for road following is to divide an outdoor image into road and non-road regions. Assuming that roads appear bright and with low saturation and non-road areas correspond with low intensity and high saturation, Lin and Chen reduce the segmentation process to a one-dimensional search problem. For every pixel in the image they calculate the value $V = (S - S_{mean}) - (I - I_{mean})$, where S denotes the saturation of the pixel, I denotes the intensity of the pixel, and S_{mean} and I_{mean} are the mean values of saturation and intensity, respectively, of the whole image. Pixels are classified into road and non-road areas by a threshold that is determined by the peaks in the histogram of all V-values. In this investigation, Lin and Chen found their algorithm to be more stable using the S and I components than when using the RGB space with the value V defined as $V = (2R - G - B)/(G - B)$. This holds especially when the image contains shadowed parts of the road.

4.1.7 Colour image segmentation using three perceptual attributes

Author(s): S. Tominaga ([70]).

Year: 1986.

Region definition(s): Peaks of the three histograms in the H, V, and C component of the Munsell space determine intervals. Pixels falling into one of these intervals create one region. Pixels falling into another area create another region.

Basic mathematical method(s): Finding peaks and basis in the 1-d histograms of the three components of the Munsell space.

Colour space(s): Munsell space Hue, Value, and Chroma.

Additional knowledge: None.

Application(s): General purpose.

Description: Since no analytical formula exists for the conversion between the CIE standard system and the Munsell system, conversion is based on a table ([79]). The algorithm for segmentation consists of the following steps.

1. First, the entire image is regarded as one region, and histograms are computed for each attribute of H, V, and C. In other cases, this computation is done for the specified regions to be segmented. The histograms are smoothed by an average operator.

2. The most significant peak is found in a set of three histograms. The peak selection is based on the shape analysis of each peak on the histogram. First, some clear peaks are selected. Next, the following criterion function is calculated for each candidate peak:

$$f = \frac{S_p}{T_a} \frac{100}{F_p},$$

   where S_p denotes a peak area between two valleys V_1 and V_2, F_p is the full-width at half-maximum of the peak, and T_a denotes the overall area of the histogram, that is the total number of pixels in the specified image region.

3. Thresholding of a colour image is executed using two thresholds derived from the lower bound V_1 and the upper one V_2 for the most significant peak in the set of three histograms. Herewith, an image region is partitioned into two sets of subregions; one consists of subregions corresponding to the colour attributes within the threshold limits, and the other is a set of subregions with the remaining attribute values.

4. The thresholding process is repeated for the extracted subregions. If all the histograms become monomodal, the cluster detection is to be finished and a suitable label is assigned to the latest extracted subregions. The 8-connection property is used to smooth the labelled image and to eliminate small regions and short lines.

21

5. Steps 1 through 4 are repeated for the remaining regions. The segmentation process is terminated when an area of the regions is sufficiently small in comparison to the original image size or no histogram has significant peaks. The remaining pixels without labelling are merged into the neighbouring labelled regions of similar colour.

4.1.8 A colour classification method for colour images using a uniform colour space

Author(s): S. Tominaga ([71]).

Year: 1990.

Region definition(s): Uniform regions are defined by thresholding histograms and due to minimal colour distances to cluster centres.

Basic mathematical method(s): Three 1-d histograms are analysed to find significant peaks and valleys. Then a reclassification process is employed classifying the representative colours for the extracted clusters on a colour distance.

Colour space(s): CIE(Lab).

Additional knowledge: None

Application(s): General purpose.

Description: The approach consists of two steps. The first step is a modification of the algorithm presented by Tominaga in 1986 ([70]) which has been described in the previous subsection (4.1.7). The modification is employed to overcome the problem of handling overlapping clusters. It consists in computing the principal components axes in the Lab-space for every region to be segmented. Next, peaks and valleys are searched for in the three one-dimensional histograms of every coordinate axis (as in the algorithm mentioned in the previous subsection). A second step is supplemented to the algorithm for the reclassification of the pixels on a colour distance. If a set of K representative colours $\{m_1, m_2, \ldots, m_K\}$ is extracted from the image, i.e., the image is labelled to K regions, then the reclassification is applied due to the following scheme. m_1 is chosen as the first cluster centre a_1 in the colour space $a_1 = m_1$. Next, the colour difference from m_2 to a_1 is computed. If this difference exceeds a given threshold T, a new cluster centre a_2 is created as $a_2 = m_2$. Otherwise m_2 is assigned to the domain of the class a_1. In a similar fashion, the colour difference from each representative colour $(m_3, m_4, \ldots)$ to every established cluster centre is computed and thresholded. A new cluster is created if all of these distances exceed T, otherwise the colour is assigned to the class to which it is closest. Unfortunately, the author does not mention the colour measure he uses but it seems as any colour measure should be convenient.

4.2 Clustering in colour space

4.2.1 Colour image segmentation and labeling through multiediting and condensing

Author(s): F. Ferri and E. Vidal ([18]).

Year: 1992

Region definition(s): Pixels are assigned to the given region using NN rule. Prototypes for the region are derived by multiediting and next condensing technique.

Basic mathematical method(s): Clustering with NN algorithm. Number of prototypes reduction using multiediting and condensing.

Colour space(s): YUV.

Additional knowledge: The scene consists of leaves, fruits, and sky.

Application(s): Robotic citric harvester.

Description: Supervised clustering is performed in 10–dimensional feature space consisting of feature vectors F based on chromatic components of YUV colour system (14). Namely with each pixel (i, j) we join the feature vector:

$$F_{i,j} = (U_{i,j}, V_{i,j}, U_{i+h,j}, V_{i+h,j}, U_{i-h,j}, V_{i-h,j}, U_{i,j+h}, V_{i,j+h}, U_{i,j-h}, V_{i,j-h}) \ .$$

The step h is chosen in accordance with the expected size of image fruit pieces, given the known optics and camera-to-scene distance.

In the experiment two representative training images of resolution 128×128 were manually segmented and each region is randomly subsampled with bias on smaller objects, i.e. fruits. In this way 1513 element training set was obtained (493 for fruits, 644 for leaves, and 376 for sky). In the multiediting step total number of prototypes was 1145 (358 for fruits, 428 for leaves, and 359 for sky). The condensing step gave dramatic data reduction to 22 prototypes (9 for fruits, 12 for leaves, and 1 for sky). In four testing images NN performance was about 90% correct pixel classification.

Multiedit Algorithm
Input: initial reference set R, number of no change iterations $I \in N$, number of blocks in each partition $m > 2$.
Output: multiedited reference set.
Method

1. Randomly divide R into m subsets $R(i)$, $i = 1, \ldots, m$.

2. For $i = 1, \ldots, m$ classify references in $R(i)$ with the NN rule using $R(i \bmod m + 1)$ as training set.

3. Discard those references that were misclassified in Step 2 and with remaining ones form a new set R.

4. If the last I iterations elapsed with no editing, then STOP, else Go to step 1.

Condensing Algorithm
Input: multiedited reference set R.
Output: condensed reference set S (initialized to empty set).
Method

1. Select a reference point from R and join it to S.

2. For each point $p \in R$ classify p with S as the training set, if p is misclassified then transfer p from R to S.

3. If no transfers occurred during Step 2 or R is empty then STOP, else go to Step 2.

4.2.2 Automatic colour segmentation algorithms with application to skin tumor feature identification

Author(s): S.E. Umbaugh, R.H. Moss, W.V. Stoecker, and G.A. Hance ([75]).

Year: 1993.

Region definition(s): Pixels are classified by minimum distance to single representatives (prototypes) of classes.

Basic mathematical method(s): KL transform in colour space. Median subdivision along maximum range axis.

Colour space(s): RGB, HSI, Luv, XYZ, and a sort of rgb.

Additional knowledge: The number of segmented regions is known.

Application(s): Skin tumor feature identification.

Description: Representatives are obtained by median splitting process in KL transform coordinates in colour space. Namely, at each subdivision step the most occupied box is chosen and the axis with the maximum range is taken and splitted in median point on that axis. The subdivision is continued till the specified number of boxes is obtained. Then representatives are taken as gravity centres of points in each box.

The authors report the best classification results for chromatic coordinates, but from the text we can only guess that a sort of rgb (normalised by intensity a linear combination of RGB coordinates) was used.

It is interesting that the number of target regions in the segmentation process is not given by the operator but is offered from AI induction engine which has inductive rules based on the first and the second order RGB statistics. The knowledge base was defined by dermatologists.

4.2.3 Colour image segmentation using recursive principal component analysis and vector median splitting

Author(s): W. Skarbek

Year: $* * **$

Region definition(s): Pixels in the given region are in the list assigned to a leaf in the median tree. This tree is created by splitting of the space by the plane perpendicular to the direction of biggest variance going through the median point. Splitting is not continued if uniformity condition is true.

Basic mathematical method(s): Principal component analysis. Jacobi method for eigenvectors. Binary tree manipulations.

Colour space(s): rgb.

Additional knowledge: None.

Application(s): General purpose.

Description: Before the algorithm specification, we give the formula for the direction e with the maximal variance for the given set of colours taken from pixel list L. By the variance $var(L, e)$ we mean the variance of projections onto the direction e passing through the gravity centre of colours in L:

$$var(L, e) \doteq \frac{1}{|L|} \sum_{s \in L} ((s.v - q)^T e)^2 \ ,$$

where

$$q \doteq \frac{1}{|L|} \sum_{s \in L} s.v \ .$$

Further we need the following statistics:

$$N_L \doteq |L| \ ,$$

$$R_L \doteq \sum_{s \in L} s.v \cdot s.v^T \ ,$$

$$m_L \doteq \sum_{s \in L} s.v \ , \tag{23}$$

$$q_L \doteq \frac{m_L}{N_L} \ ,$$

$$\tilde{R}_L \doteq \sum_{s \in L}(s.v - q_L) \cdot (s.v - q_L)^T = R_L - \frac{1}{N_L} m_L m_L^T \ .$$

It can be proved that:

1. Matrix $\tilde{R}_L$ is symmetric.

2. Direction e_L of maximal variance $var(L, e)$ is the eigenvector e_L of matrix $\tilde{R}_L$ corresponding to maximal eigenvalue λ_L.

3. $var(L, e_L) = \lambda_L / N_L$.

It follows, that for space subdivision such list L should be chosen for which λ_L is the biggest. The partition of the list L into two lists L_1 and L_2 is performed according the membership to halfspace determined by the plane $e_L^T(x - Q_L) = 0$, where Q_L is the vector median (not centroid !) of colours in L:

$$L_1 \doteq \left\{ s \in L : e_L^T s.v \leq e_L^T Q_L \right\} \ ,$$

$$L_2 \doteq \left\{ s \in L : e_L^T s.v > e_L^T Q_L \right\} \ . \tag{24}$$

Let L_1 and L_2 be subdivision of L. Then we say that the list of pixels L, L_1, L_2 is *colour consistent* iff

$$\frac{\min(f(L), f(L_1), f(L_2))}{\max(f(L), f(L_1), f(L_2))} \geq 0.9$$

for all colour features f (here $f = r, g, b$), where $f(L)$ denotes the average of feature f in the set L.

Construction of median tree

1. Init the root of tree T and initial list:

$$root.L := \text{all pixels in the image} \ ;$$

2. Determine for $root.L$ statistics R, m, N, q;

3. Do until all leaves are marked UNIFORM:

(a) Find in T the leave n not marked UNIFORM such that $\lambda_{n.L}$ is the biggest;

(b) Subdivide temporarily $n.L$ into two lists L_1 oraz L_2 according the criterion 24. If sets of pixels L, L_1, L_2 are *colour consistent* then cancel this subdivision, mark n as UNIFORM and continue from step 3, otherwise assign them to two children of node n:

$$n_1.L := L_1; \; n_2.L := L_2;$$

(c) Determine for $n_1.L$ oraz dla $n_2.L$ statistics R, m, N, q.

4. For each leave n in T find all connected components in the set $n.L$.

4.2.4 Segmentation of coloured topographic maps in perception based colour spaces

Author(s): B. Lauterbach and W. Anheier ([45]).

Year: 1993.

Region definition(s): Pixels are classified by their minimum distance to additive colour matching lines defined between each two cluster centres which are peaks in the two-dimensional uv-histogram.

Basic mathematical method(s): Peak selection in 2-d uv-histogram. Construction of additive colour matching lines and computation of geometric distance of pixels to those lines.

Colour space(s): CIE(Luv).

Additional knowledge: Maps are printed on white paper with black letters.

Application(s): Segmentation of coloured topographic maps.

Description: Maxima in the cumulative uv-histogram of the Luv colour space are searched for to define cluster centres for segmentation. The maxima are detected by computing the difference between the value of the cumulative histogram and the mean value of a surrounding window. Additive colour matching lines, *acl*, are defined by straight lines between two cluster centres. Instead of classifying pixels by the colour distance to the cluster centre, pixels are classified to a pair of clusters by the Euclidean distance of the pixel value to the *acl* between those two cluster centres in the uv-space. Furthermore, a circle is defined around every cluster centre to avoid misclassification of a pixel that is near to the cluster centre. The distance of the pixel to the *acl* has to be less than the distance between the two centres for classification. After every pixels has been assigned to a pair of clusters, the final classification is made due to

27

the minimum distance of the pixel to one of the two cluster centres. This segmentation does not take luminance into account for segmentation. Therefore, the monochrome cluster is partitioned into further clusters by peak searching in the one- dimensional L-histogram. Due to the procedure mentioned last, this algorithm is not a general approach to colour image segmentation.

4.2.5   A recursive clustering technique for colour picture segmentation

Author(s): M. Celenk ([13]).

Year: 1988.

Region definition(s): 3-D volumes in the CIE(Lab) space determine the colour distribution of clusters. Pixels with colour vectors located inside this volume belong to the same region.

Basic mathematical method(s): Detect clusters by fitting to them some circular-cylindrical decision volumes in the CIE(Lab) space. Determine boundaries of the clusters by finding peaks and valleys in the 1-d histograms of L, a, and b. Then, project the estimated colour distributions onto the *Fisher* linear function for 1-D thresholding.

Colour space(s): CIE(Lab), $L^*H^0C^*$.

Additional knowledge: None.

Application(s): General purpose.

Description: Colour image segmentation is considered as a recursive cluster detection problem. The method operates in the CIE(Lab) colour space and detects image clusters by fitting them some circular-cylindrical decision volumes. This estimates the clusters' distribution in the uniform colour space without imposing any constraints on their forms. Boundaries of the decision elements consist of two planes of constant lightness, two cylinders of constant chroma, and two planes of constant hue. They are determined using 1-D histograms of the $L^*H^0C^*$ coordinates of the space. The psychometric colour space $L^*H^0C^*$ or LHC (in short notation) is the cylindrical representation of the Lab-space. It is given by $L = L, H = arctan(b/a)$,and $C = (a^2 + b^2)^{1/2}$ (compare eqs. (15) and (16)). The detected clusters are then isolated from their neighbours by projecting their estimated colour distributions onto the *Fisher* linear discriminant function. For two colour clusters w_1 and w_2 the *Fisher line* W is given by $W = (K_1 + K_2)^{-1}(M_1 - M_2)$, where (K_1, K_2) and (M_1, M_2) are the covariance matrices and the mean vectors, respectively, of the two clusters. The colour vectors of the image points, which are the elements of w_1 and w_2, are then projected onto this line given by the above equation using $d(C) = W^T C$. Here

d is the linear discriminant function and C is a colour vector in one of the clusters. A one-dimensional histogram is calculated for the projected data points and thresholds are determined by the peak and valleys in this histogram.

4.2.6 Initial segmentation for knowledge indexing

Author(s): M. Hild, Y. Shirai, and M. Asada ([31]).

Year: 1992.

Region definition(s): Regions are defined by 19 given hue clusters.

Basic mathematical method(s): Backproject image pixels into 19 predefined hue clusters and find hue groups due to the pixel densities in the clusters.

Colour space(s): HSI.

Additional knowledge: Dichromatic reflection model.

Application(s): Generation of indices into a knowledge base of object models.

Description: Hild, Shirai, and Asada define knowledge indexing as finding relevant pieces of knowledge in a knowledge base with the help of a set of descriptive properties, called index. There is an affinity between indexing and matching. While matching refers to establishing correspondences between a set of properties of a stored model, the role of indexing is to establish connections to models or parts of models without verifying the appropriateness of these connections. Thus, an index can be used as a hypothesis of the object depicted in the image. The segmentation process is used to find relevant features that can be used as indices into a knowledge base. Obviously, colour is not the only possible feature in a knowledge base, but it is an important feature. Since histograms of hue, saturation, and intensity are often ambiguous there are not suitable for a sufficient discrimination of objects to be segmented in similar images. The authors propose to divide the hue space into 19 different partitions. Then the images are backprojected into these partitions and the pixel density is calculated in each of the partitions. Hue clusters are grouped due to these densities to define chromatic features that can be used as indices into a knowledge base.

4.3 Fuzzy clustering in colour space

4.3.1 On the color image segmenatation algorithm based on the thresholding and the fuzzy c−means techniques

Author(s): Y.W. Lim and S.U. Lee ([46]).

Year: 1990.

Region definition(s): Pixel is allocating to the region using maximum of the fuzzy membership function. However, from coarse segmentation most of pixels is already assigned to regions using hexahedra bounds.

Basic mathematical method(s): Histogram scale space analysis. Fuzzy c–means clustering.

Colour space(s): RGB, XYZ, YIQ, Luv, $I_1 I_2 I_3$ (defined in Ohta's paper [53]).

Additional knowledge: None.

Application(s): General purpose.

Description: Since the number of regions is not known, a coarse segmentation is made by the 1D histograms scale space analysis. This analysis enables reliable detection of meanigful peaks in the given histogram and in the same time it determines significant intervals around those peaks. The bounds of the intervals are found as zero–crossings of the second derivative for τ-scaled version of the histogram. The proper scaling parameters follow from Witkin's interval tree associated with the *fingerprint diagram*.

The τ-scaling of the histogram $h(x)$ is defined by the convolution of h with Gaussian function which has the mean zero and the standard devaition equal to τ:

$$H(x, \tau) \doteq h(x) * g(x, \tau) = \int_{-\infty}^{+\infty} h(u) \frac{1}{\sqrt{2\pi}\tau} \exp\left[-\frac{(x - u)^2}{2\tau^2} \right] du \ .$$

The second derivative of scaled function can be computed by the convolution with the second derivative of Gaussian function:

$$H_{xx}(x, \tau) = h(x) * g_{xx}(x, \tau) \ .$$

The coarse segmentation produces a number of hexahedra obtained as cartesian products of peak intervals found for each colour component seperately. Pixels not entering into hexahedras, enter into ambiguous regions. Their fine segmentation is made with fuzzy c–means clustering. Gravity centre of pixels in hexahedras are used initially as representatives of fuzzy classes.

Coarse segmentation algorithm

1. Compute histograms for each colour component of the image;

2. Apply scale space analysis to the histograms;

3. Define valid classes (hexahedra) $V_1, \ldots, V_c$;

4. For each pixel p if it belongs to one of hexahedra V_i then p is classified and labelled by i, otherwise it is unclassified;

5. Calculate the gravity centre v_i for each class V_i.

For any unclassified pixel (x, y) with value $P(x, y)$ the fuzzy membership function U is computed. It evaluates the degree of memebership of the given pixel to the given fuzzy class V_i, $i = 1, \ldots, c$:

$$U_{i,(x,y)} \doteq \frac{1}{\sum_{j=1}^{c} \left(\frac{\|P(x,y) - v_i\|}{\|P(x,y) - v_j\|} \right)^{2/(m-1)}}$$

The constant m controls the clustering process. Typically $m = 2$.

The pixel (x, y) is assigned to the class V_k such that its membership value is maximum:

$$k(x, y) = \arg \max_{1 \le i \le c} U_{i,(x,y)} \ .$$

4.3.2 Iterative fuzzy image segmentation

Author(s): T.L. Huntsberger, C.L. Jacobs, and R.L. Cannon ([37]).

Year: 1985.

Region definition(s): Regions are created by pixels which has their fuzzy membership value above α-cutoff.

Basic mathematical method(s): Fuzzy c–means algorithm.

Colour space(s): RGB, $I_1 I_2 I_3$.

Additional knowledge: None.

Application(s): General purpose.

Description: This approach is fully based on original Bezdek's algorithm ([9]) for fuzzy clustering. In the consecutive stage the clustering procedure is applied for remaining unclassified pixels. For each pixel it elaborates the fuzzy membership function and once its maximum value satisfies the α-cutoff, the pixel is classified.

Fuzzy c–means algorithm
For the given data points $x_1, \ldots, x_n \in R^p$:

1. Fix c, $2 \le c < n$; choose any inner product norm metric for R^p; and fix m, $1 \le m < \infty$. Initialize (e.g.:uniformly) the membership function $U^0 : \{1, \ldots, m\} \times \{1, \ldots, c\} \to [0, 1]$. Then at step $l = 0, 1, 2, \ldots$:

2. Calculate c fuzzy cluster centre $\{v_i^{(l)}\}$:

$$v_i = \sum_{k=1}^{n} (u_{ik})^m x_k \Big/ \sum_{k=1}^{n} (u_{ik})^m \ , \ i = 1, \ldots, c \ .$$

31

3. Update $U^{(l)}$: Let $d_{ik} = \|x_k - v_i\|$:

$$\text{If } d_{ik} \neq 0 \text{ then } \quad u_{ik} = 1 \left/ \left[\sum_{j=1}^{c} \left(\frac{d_{ik}}{d_{jk}} \right)^{2/(m-1)} \right] \right. \quad \text{else } u_{ik} = 0 \; .$$

4. Compare $U^{(l)}$ to $U^{(l+1)}$: if $\|U^{(l+1)} - U^{(l)}\| \leq \epsilon$ then STOP, otherwise Go to Step 2.

The authors adopted the above general algorithm by setting $c = 4$, $m = 2$. In each stage only 2400 pixels are randomly chosen from the remaining unlabelled pixels. After getting stabilized clusters on these 2400 data points, for not included pixels the fuzzy membership function is calculated and defuzzification, i.e. α-cutoff is performed. The remaining pixels are transfered to the next stage.

It may happen that two cluster centre are very close. Then it is better to join them into one cluster. The authors used the following equivalence relation for cluster centres:

$$\text{for each colour feature } f \quad \frac{2|f_1 - f_2|}{f_1 + f_2} < 0.075 \; .$$

4.3.3 Low-level segmentation of aerial images with fuzzy clustering

Author(s): M. Trivedi and J.C. Bezdek ([73]).

Year: 1986.

Region definition(s): Regions are created by pixels that have their fuzzy membership value above α-cutoff.

Basic mathematical method(s): Fuzzy c-means algorithm.

Colour space(s): Multispectral space (e.g.: the channels of LANDSAT images).

Additional knowledge: None.

Application(s): Segmentation of multispectral images.

Description: This approach is based on Bezdek's algorithm ([9]) for fuzzy clustering. Thus, the algorithm is identical to the fuzzy c-means algorithm employed by Huntsberger, Jacobs, and Cannon ([37]) which has been described in the previous subsection. The difference between both approaches is that Trivedi and Bezdek implemented their algorithm on a pyramid data structure. Four values are replaced by their mean value to construct a higher level in the pyramid. Starting from the highest level, regions are created by pixels that have their fuzzy membership value above α-cutoff. For those areas in the image that are not homogeneous due to the definition mentioned above, the fuzzy membership is checked in the next lower level of the pyramid.

5 Area based segmentation

In this section segmentation algorithms use uniformity criteria calculated in regions of image domain. We divide available techniques into two groups:

1. **Region growing**: in this class of algorithms a number of basic uniform regions (seeds) are given and different strategies are applied to join surrounding neighbourhoods; to distingush this group from the next one it is important that seeds here are not resulting from splitting or subdivision processes of nonuniform regions;

2. **Split and merge**: here algorithms start from nonuniform regions, subdivide them until uniform ones are obtained, and then apply some merging heuristics to fit them to maximal possible uniform area.

5.1 Splitting versus merging

In this point we present two general schemes of region based segmentation algorithms, i.e. splitting and merging techniques. They use explicitly or implicitly quadtree for block subdivision structure. Recursive formulas for the variance σ^2 as the non uniformity measure and the complexity analysis of both methods is also given.

All material of this subsection is based on the point 1.2.2 of Skarbek's monography ([67]).

Segmentation of image I by splitting
(1) Subdivide the image I into blocks B of size $N \times N$, $N = 2^n$,
 where n is the *rank* of block.
(2) For each block B:
  (3) If NONUNIFORMITY$(B) > T$ and $k = \text{rank}(B) > 0$, then
    - divide the block B into four equal blocks B_1, B_2, B_3, B_4;
    - repeat step (3) in turn for B_1, B_2, B_3, B_4;
   otherwise output B as the uniform block.

The above algorithm works recursively in a *top-down* way. The next algorithm performs the segmentation by merging of image blocks in a *bottom-up* way. The image scan is controlled here by the Hilbert curve, which is one of space filling curves. This curve implements a scan in consecutive quadrants of lower rank included in the block of higher rank. This property enables to merge smaller quadrants into bigger ones, effectively segmenting the image by the bottom-up technique. Note, that any other scan satisfying this property may replace the Hilbert scan in this algorithm (for instance Morton sequence).

Segmentation of image I by merging

Subdivide the image I into blocks B of size $N \times N,\ N = 2^n$. Init empty stack of block records, controlled by the variable top. Then:

$SetScan(B, HILBERT);\ top := 0;\ k := 0$

scan B **with** p **do**

    $Push(p);\ k := k + 1$

    if $k = 4$ **then** $\{\ Reduce();\ k := 0\}$

endscan

Let B_1, B_2, B_3, B_4 be four blocks from the top of the stack and $B' = B_1 \cup B_2 \cup B_3 \cup B_4$. Then the recursive function $Reduce()$ performs the following actions:

1. **if** NONUNIFORMITY$(B') > T$ **then** output top four blocks as uniform ones; **return**;

2. Pop B_1, B_2, B_3, B_4;

3. Push B';

4. **if** $EqualRankTopFour()$ **then** $Reduce()$;

5. **return**.

Computational savings of merging techniques follow from the fact that we can compute (non) uniformity measure (for instance colour variance, correlation of histograms for a colour feature, moments, regression analysis coefficients, etc.) for the given block using corresponding statistics for its subblocks. We show here this possibility only for the variance of scalar feature. However, analogous facts could be obtained for other mentioned above (non) uniformity measures.

Recursive variance computing for block B

Let $sum(B)$ be the sum of pixel values in B, $ssum(B)$ is the sum of their squares, $M(B)$ is their average, and $\sigma^2(B)$ is their variance. Then

1. If B is single pixel p, then $sum(B) = p.v,\ \ ssum(B) = p.v^2$;

2. If B is the union of equal size, disjoint blocks B_1, B_2, B_3, B_4, then

    (a) statistics sum and $ssum$ are computed recursively:

$$sum(B) = \sum_{i=1}^{4} sum(B_i), \quad ssum(B) = \sum_{i=1}^{4} ssum(B_i),$$

$$M(B) = \frac{sum(B)}{|B|}, \quad \sigma^2(B) = \frac{ssum(B)}{|B|} - M^2(B);$$

(b) statistics M, σ^2 can be computed recursively, too:

$$M(B) = \tfrac{1}{4} \sum_{i=1}^{4} M(B_i),$$

$$\sigma^2(B) = \tfrac{1}{4} \left(\sum_{i=1}^{4} M^2(B_i) + \sum_{i=1}^{4} \sigma^2(B_i) \right) - M^2(B) \ .$$

Assuming full precision of real arithmetic, it can be proved that both methods, i.e. splitting and merging produce the same collection of uniform blocks.

Let the segmented image block is of rank n. Denote by $\mathcal{S}_i$ the collection of its uniform subblocks of rank i of cardinality s_i, $i = 0, \ldots, n$. Let also

$$\mathcal{N}_s(n, s_0, \ldots, s_n) \text{ and } \mathcal{N}_m(n, s_0, \ldots, s_n)$$

be the number of operations $+, -, *, /$ in splitting and merging methods, respectively.

Comparison of computational complexity of splitting and merging

1.

$$3\left(-s_0 + \sum_{i=0}^{n} 4^i(n+1-i)s_i\right) \ < \ \mathcal{N}_s(n, s_0, \ldots, s_n)$$

$$\mathcal{N}_s(n, s_0, \ldots, s_n) \ \leq \ 3.5\left(-s_0 + \sum_{i=0}^{n} 4^i(n+1-i)s_i\right);$$

2.

$$\mathcal{N}_m(n, s_0, \ldots, s_n) = \frac{10}{3} \sum_{i=1}^{n} (4^i - 1)s_i + \frac{10}{4} \sum_{i=0}^{n-1} s_i \ .$$

3. If $s_n \neq 0$, then

$$\mathcal{N}_s - \mathcal{N}_m \ > \ 2.5 \sum_{i=0}^{n-1} 4^i s_i \ .$$

4. Acting on the block of size $2^n \times 2^n$, the merging method requires more operations than the splitting method if and only if the result of segmentation is $\mathcal{S}_0 = \ldots = \mathcal{S}_{n-1} = \emptyset, \mathcal{S}_n = \{B\}$, i.e. if the image block B is uniform.

5.2 Region growing

5.2.1 Colour image segmentation using boundary relaxation

Author(s): R.I. Taylor and P.H. Lewis ([69]).

Year: 1992.

Region definition(s): Region R is uniform if average distance $D(R)$ of its colour points to its centroid is less than threshold D_{max} (experimentally set to 10% of distance between black and white points in the colour space).

Basic mathematical method(s): Boundary relaxation technique.

Colour space(s): TekHVCT.

Additional knowledge: None.

Application(s): General purpose.

Description: Two algorithms are given. The first one, called *boundary relaxation method* is used as a correction technique for already coarsely segmented image. The second one (*alternative segmentation technique*) is true region growing technique starting from seeds and extending regions by one pixel in each step.

Boundary relaxation method
In one iteration perform for each pixel p the following actions:
If $p \in R_a$ and p is adjacent to R_b and moving p from R_a to R_b reduces $D(R_a) + D(R_b)$ and maintains $D(R) < D_{max}$ then move p from R_a to R_b. If there is more than one possible move then the one giving the lowest $\sum D$ across all regions involved, is performed.

Alternative segmentation technique
It starts from a seed in the form of 2×2 image blocks. It grows up the region by adding ajacent pixel which has the nearest colour to the present colour centroid of the region. The pixel is only added if its colour is distant from the centroid less than $2D_{max}$. At the end small noisy regions are joined to adjacent or surrounding regions.

5.2.2 Low–level segmentation of multispectral images via agglomerative clustering of uniform neighbourhoods

Author(s): M. Amadasun and R.A. King ([3])

Year: 1988.

Region definition(s): Region is a maximal area satisfying uniformity condition and including the given seed.

Basic mathematical method(s): Merging regions and recursive computation of statistics.

Colour space(s): Multispectral images (e.g.: LANDSAT images).

Additional knowledge: Known number c of classes.

Application(s): Segmentation of LANDSAT images.

Description: Firstly, the algorithm looks for uniform seeds among blocks creating uniform partition of image domain (typically blocks are 12×12 or 16×16). The uniformity criterion has the form:

$$\frac{\min(\bar{f}_{block}, \bar{f}_{quarter})}{\max(\bar{f}_{block}, \bar{f}_{quarter})} \geq \alpha$$

which should be true for all features f and for all quarters of the block. The algorithm starts with $\alpha = 0.95$. If the number of seeds is less than c then this condition is relaxed by decreasing of α.

In the merging phase, not necessary adjacent regions with the nearest centroids are joined together to create one region. This merging operation is repeated while the number of regions is greater than desirable number of classes.

The merging procedure uses recursive formulas for means and variances of regions. Namely if i-th region has N_i pixels, the mean of k-th feature equal to $\bar{x}_{ik}$, and the variance s_{ik}^2, then for the region R_{ij} which is merged from R_i and R_j, we have the following formulas:

$$\bar{x}_{ijk} = \frac{N_i \bar{x}_{ik} + N_j \bar{x}_{jk}}{N_i + N_j}, \tag{25}$$

$$s_{ijk}^2 = \frac{N_i (\bar{x}_{ik}^2 + s_{ik}^2) + N_j (\bar{x}_{jk}^2 + s_{jk}^2)}{N_i + N_j} - \bar{x}_{ijk}^2 . \tag{26}$$

5.2.3 Colour image segmentation using regression analysis in RGB space

Author(s): Ismaili I.A. and Gillies D.F. ([38]).

Year: 1994.

Region definition(s): Uniform hue seeds are found by bottom–up process. Their secondary classification is made by slope and correlation check for the regression line in RG plane. Further growing is made using neighbourhood seeds in a quadtree of block subdivision.

Basic mathematical method(s): Area merging process. Analysis of regression. Recursive formulas for slope and intercept in analysis of regression.

Colour space(s): Hue and RGB.

Additional knowledge: Endoscopic images of colon.

Application(s): Automatic removing of fluid during endoscopy.

Description: Starting from small blocks, merging of uniform blocks is made. It is driven by a quadtree structure. Uniformity is defined by exceeding a threshold

for the dominant feature. If the uniformity criterion is satisfied then the classi-
fication according the correlation r and regression line slope β_1 on RG plane is
performed. Histograms and regression parameters are computed recursively:
The maximal blocks obtained from this process become seeds. Region growing
is attempted by merging seeds, first from the same level of quadtree, and next
from the lower ones.

Experimentally was found that the high correlation and the slope of the re-
gression line, between 0.3 and 1.35, discriminate the fluid from other similar
in hue feature areas.

5.2.4   Image segmentation by colour feature flow

Author(s): W. Skarbek

Year: $****$

Region definition(s): Regions are expanded by flow of colour through enabled
edges. Colour relaxation at region borders.

Basic mathematical method(s): Colour distribution functions for objects. Wave
propagation. Segmenting by relaxation.

Colour space(s): HSI with modified metric for small intensities.

Additional knowledge: Colour models for all objects in the scene are known.

Application(s): General purpose.

Description: It is assumed that objects in the scene have known hue characteristic
curves and two different materials in the scene have significantly different hue
curves.

Algorithm has two stages: preprocessing stage and feature flow stage.

Preprocessing stage: Image is divided into 4×4 blocks. Each edge E of
such block is labelled by the material id which is assigned by hue histogram
type of two blocks union: 4×4 and 8×8 block put on, externally, on the edge
E. The local recognition of the material is made by correlation of hue curves.
The most correlated 4×4 blocks are put on the *seed list*.

Feature flow stage: All seeds are filled with colours assigned to their materi-
als. In one pass the colours are dispersed through edges which are compatible
(have the same material id) with colours inside the blocks. A necessary number
of passes is made to make the flow of colours to all possible blocks. If no fur-
ther flow is possible, then uncoloured blocks are coloured according the nearest
colour centroid of adjacent regions. Since the flow is made in 4×4 blocks then
fine corrections on borders is made by the pixel relaxation technique.

5.2.5 Real-time colour image segmentation

Author(s): M.N. Fesharaki and G.R. Hellestrand ([19]).

Year: 1993.

Region definition(s): Regions are homogeneous sets of pixels. Homogeneity is verified based on the *Kolmogorov-Smirnov* statistics on the cumulative distribution functions of four 5 by 5 windows representing different orientations.

Basic mathematical method(s): Region growing.
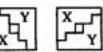
Colour space(s): RGB.

Additional knowledge: None

Application(s): General purpose.

Description: The algorithm is based on testing the homogeneity of pixels around a centre pixel by using statistical inference techniques. A 5 by 5 window around each pixel is partitioned into two sub-samples in four different orientations as shown in the following figure:

In each window, the number of pixels in the sub-regions X and Y are m and n, respectively. That is N = m + n observations denoted as $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$. Two hypotheses are formulated about pixels in the neighbourhood: H_0 : All pixels in the given neighbourhood belong to one homogeneous region, or H_1 : they belong to at least two different homogeneous regions. Assuming that X and Y are mutually independent and come from two populations, the null hypothesis may be stated as $H_0 : P(X \leq a) = P(Y \leq a)$ for all a. To verify this hypothesis, the cumulative distribution functions, cdf, of two sub-samples are compared with each other. The *Kolmogorov-Smirnov* statistics considers the base for the verification of the homogeneity of the two sub-samples. To achieve this, the N observations from the two samples are ordered to form the set $Z_{(i)} : i = 1, \ldots, N$, where $Z_{(1)} \leq Z_{(2)} \leq \ldots \leq Z_{(N)}$. The statistics J is defined as

$$J = \frac{mn}{d} \max_{i=1,\ldots,N} (|F_m(Z_i) - G_n(Z_i)|)$$

where m and n are as before, d is the greatest common divisor of m and n, and

$$F_m(a) = \frac{\text{number of } X\text{'s} \leq a}{m} \text{ and } G_n(a) = \frac{\text{number of } Y\text{'s} \leq a}{n}$$

are considered to be the empirical cumulative distribution functions for two random samples of sizes m and n from two distributions with cdf $F(x)$ and $G(y)$. All J's are calculated for the four different orientations and ordered to form a set $\{(J_{(i)} : i = 1, ..., 4)\}$, where $J_{(1)} \leq ... \leq J_{(4)}$. If $J_4 \leq 5$, the centre pixel of the window is considered to be a part of a homogeneous region. Otherwise, each pixel is compared with the pixels to its left and above. If the absolute difference in the comparison metric is less than a threshold T (fixed to $T = 12$ in the experiments presented in ([19])) the pixels are considered homogeneous. The homogeneity test is applied to every colour channel producing three Boolean lists for each channel. The values are set to one, if the pixel verifies one of the corresponding homogeneity criteria mentioned above (homogeneous, left merge, or above merge). The results obtained in the three colour channels are intersected to create a homogeneity criterion in the colour space, i.e., the Boolean values are "ANDed".

Region growing is started after this preprocessing. For this purpose, each pixel is considered as a region. A 2x2 window containing the left and above pixel is moved across the image. Neighbouring pixels are candidates for merging if the related signals are set. For example, each pixel is allowed to merge with its adjacent left pixel, if either its left merge signal or the homogeneous signal computed from its left pixel is set. The same procedure is carried out for merging the current pixel with the pixel above. If it is not allowed to merge with its adjacent pixels, a new label is assigned. If it is allowed to merge with both its left and above pixels, the minimum label of those two pixels is assigned to that pixel. A block diagram and a floor plan are presented to show how this algorithm can be mapped to a pipelined real-time hardware architecture designed in VLSI technique.

5.2.6 Colour image segmentation (using the topographic watershed algorithm)

Author(s): F. Meyer ([49]).

Year: 1992.

Region definition(s): Regions are homogeneous sets of pixels. Homogeneity is verified by the minimum colour difference between pixels.

Basic mathematical method(s): Region growing using the topographic watershed algorithm.

Colour space(s): RGB, HSL.

Additional knowledge: The relevance of seed points has to be known in advance.

Application(s): General purpose.

Description: An intensity image becomes a topographic surface if a grey-tone is considered as an altitude. The basic idea of the watershed algorithm may then be understood as a flooding process of this topographic surface. At each minimum of the relief a hole is bored and each hole acts as a water source. The relief converts to a lake district when pumping water into each hole. During this procedure, the level of each lake increases with uniform speed. Holes or seed points for region growing, respectively, might be found by calculating the minima of the gradient image. Starting from these seed points a hierarchy for region growing is established based on the differences between the intensity value of the seed point and the intensity values of the neighbouring pixels. First, all neighbouring pixels of a seed point get the same label as the seed point itself if their intensity difference is zero. Afterwards, all neighbouring pixels of a seed point get the same label as the seed point if their intensity difference is 1. Then the difference level is further increased and the procedure is continued until all pixels in the image are labelled. When this scheme is applied to colour image segmentation the seed points in the colour image might be found by calculating the minima of the gradients in all three colour channels. Furthermore, the difference between the intensity values is replaced by a colour difference between two pixels. Due to computational efficiency the author employed the maximum norm to measure the colour difference between two pixels, but any colour norm is suitable for this purpose. Some additional information is needed to avoid over-segmentation. It has to be known in advance which minima are relevant and which are not. The author assumes that this problem is solved but he gives no information on how to solve it.

5.2.7 A fast hybrid colour segmentation method

Author(s): L. Priese and V. Rehrmann ([59, 60, 61]).

Year: 1992, 1993.

Region definition(s): A region is uniform if a seven-dimensional predicate D holds in several levels of hierarchy.

Basic mathematical method(s): Bottom-up region growing and top-down separation technique.

Colour space(s): RGB, HSV.

Additional knowledge: None.

Application(s): General purpose (with emphasis on real-time segmentation).

Description: In a first processing step, colour values are smoothed by replacing them by the mean value of the adjacent pixels in a hexagonal structure. The segmentation is based on a hierarchical organisation of hexagonal structures.

41

At the lowest level, local region growing is applied and connected regions of small size are obtained. These connected regions are treated as pixels at higher levels. The distance between the colour content $c(r_1)$ and $c(r_2)$ of r_1 and r_2 is measured to decide whether two neighboured regions r_1 and r_2 become connected. The decision whether to combine or not is formulated as a predicate $D(c(r_1), c(r_2))$. As the colour models are 3-dimensional, D is a 6-dimensional predicate. Instead of using metrics, Priese and Rehrmann suggest to use a navigation in these six-dimensional spaces to compute whether D holds. As the level n in the hierarchy may influence D, D is treated as a 7-dimensional predicate $D(c(r_1), c(r_2), n)$. This predicate is only computed for those regions that are candidates for merging to avoid mismatches. One simple criterion for being candidates is that both regions have a common boundary. To create a second criterion, both regions r_1 and r_2 are linked by a chain $p_1, \ldots, p_k$ of pixels that are pairwise connected due to $D(c(p_i), c(p_{i+1}), 0)$ for $1 \leq$ i ¡ k, $p_1 \in r_1, p_2 \in r_2$. If two candidates r_1 and r_2 are not similar according to $D(c(r_1), c(r_2), n)$, a mismatch of their aggregated colour values $c(r_1)$ and $c(r_2)$ is detected despite a linking chain between r_1 and r_2.

5.2.8 A graph-theoretic approach to colour image segmentation and contour classification

Author(s): T. Vlachos and A.G. Constantinides ([76])

Year: 1992.

Region definition(s): Regions are homogeneous sets of pixels. Pixels are merged based on the minimum Euclidean colour distance in a four-connected neighbourhood.

Basic mathematical method(s): Region growing by iterative link removal in a pixel graph.

Colour space(s): RGB, CIE(Lab).

Additional knowledge: None.

Application(s): General purpose.

Description: The process of image segmentation is considered in a graph-theoretic context. The algorithm is based on region growing in the RGB and CIE(Lab) space using Euclidean distance to measure the difference between colours. The suppression of artificial contouring is formulated as a dual graph-theoretic problem. A hierarchical classification of contours is obtained which facilitates the elimination of the undesirable contours. Regions are represented by vertices in the graph and links between geometrically adjacent regions have weights that are proportional to the colour distance between the regions

they connect. The link with the smallest weight determines the regions to be merged. At the next iteration of the algorithm the weights of all the links that are connected to a new region are recomputed before the minimum-weight link is selected. The links chosen in this way define a spanning tree on the original graph and the order in which links are chosen defines a hierarchy of image representations. No clear perceptual advantage was gained by the authors when employing the CIE(Lab) colour space instead of the RGB space.

5.2.9 Colour segmentation by hierarchical connected components analysis with image enhancement by symmetric neighbourhood filters

Author(s): T. Westman, D. Harwood, T. Laitinen, and M. Pietikäinen ([77]).

Year: 1990.

Region definition(s): Regions are homogeneous sets of pixels. Homogeneity is verified based on the colour difference between two adjacent pixels and the average edge contrast between adjacent regions. Both values have to be less than corresponding thresholds to create a region.

Basic mathematical method(s): Image enhancement employing a Symmetric Nearest Neighbour operator and region growing based on hierarchical connected components analysis.

Colour space(s): RGB.

Additional knowledge: None.

Application(s): General purpose.

Description: The approach consists of two processing steps. First, a Symmetric Nearest Neighbour (SNN) operator is employed to enhance image values. In the original SNN operator the centre pixel value is re-substituted by averaging the values of four pixels, each a nearest neighbour of the centre pixel in one of the four symmetric pairs of a 3x3 neighbourhood. A nearest neighbour of a pair is the one closest in colour value to the centre value. In the modified version presented in ([77]), a tolerance for local directional contrast is used to detect when a centre pixel is part of a thin structure that outlies the symmetric pair lying in a background. If the centre is an outlier (ridge, valley) with respect to a pair, then the centre value is left fixed. Otherwise, the nearest neighbour of the pair contributes its value to the average. The re-substituted average of these four values for the four symmetric pairs of a 3x3 size operator is the mean of previous centre value and the average of the four values obtained by selection. After image enhancement, a region growing technique is applied based on hierarchical connected components analysis. In the first pass, all pixels are assigned labels by comparing them with the four adjacent labelled

pixels above and to the left (8-connectivity). Adjacent pixels are merged if their colour difference is lower than a predetermined threshold value. For each region the labels of adjacent regions, the average edge contrast between adjacent regions, and the lengths of the borders between adjacent regions are stored in a region adjacency graph. In the second stage, for each region starting from the one with the smallest label, the neighbouring regions with an average edge contrast smaller than a threshold epsilon are merged and the adjacency graph is updated. The region-merging step is iterated using gradually increased epsilon values if the multi-stage version of the algorithm is employed. The authors obtained about equally good results employing the maximum metric (maxRGB) as employing the Euclidean colour metric.

5.2.10 A short note on local region growing by pseudophysical simulation

Author(s): M. Brand ([11]).

Year: 1993.

Region definition(s): Regions are defined by a simulated mass moving through an irregular field of forces.

Basic mathematical method(s): Region growing by pseudophysical simulation in analogy to a mass moving through an irregular field of forces.

Colour space(s): Any tristimulus colour space.

Additional knowledge: Seed points are known in advance.

Application(s): Systems that are looking for specific simply connected shapes in local areas of interest, for example, a rectangle of unknown size, elongation, or orientation.

Description: A region grower is presented which expands to cover the largest shaped region of consistent colour values in the vicinity of its seed position. It is build on a pseudophysical analogy to a mass moving through an irregular field of forces. Local interactions between the mass and the forces will push, turn, squeeze, or deform it. The mass grows until all growth is cancelled out by the squeezing of local forces. The cumulative forces on the pseudomass are:

$$I \quad \sim \quad s^{-2} \oint_{Area} (((\mathbf{P} - \mathbf{C}/\|\mathbf{P} - \mathbf{C}/\|)(\mathbf{F} - (\mathbf{P} - \mathbf{C}))\xi(\mathbf{P}, \mathbf{C})d\mathbf{P}, \quad (27)$$

$$T \quad \sim \quad s^{-3} \oint_{Area} ((\mathbf{F} - (\overline{\mathbf{P} - \mathbf{C}})\xi(\mathbf{P}, \mathbf{C})d\mathbf{P}, \quad (28)$$

44

$$C_w \sim s - \left(s^2 - s^{-2} \oint_{Area} \mathbf{F} - \mathbf{W}\xi(\mathbf{P}, \mathbf{C}))d\mathbf{P} \right)^{1/2}, \tag{29}$$

$$C_h \sim s - \left(s^2 - s^{-2} \oint_{Area} \mathbf{F} - \mathbf{H}\xi(\mathbf{P}, \mathbf{C}))d\mathbf{P} \right)^{1/2}, \tag{30}$$

where I is the impulse vector, T is a torque, C_w and C_h represent compressive forces on the region, s is the length of a side, $\mathbf{P}$ is each point in the rectangular area, $\mathbf{C}$ is the centroid of the area, $\mathbf{F}$ is a vector normal to the edge nearest to $\mathbf{P}$ and proportional to the distance from that edge to the centroid, $\bar{\mathbf{F}}$ is a vector normal to $\mathbf{F}$ and pointing away from the centre at $\mathbf{P}$, and $\xi(\mathbf{P}, \mathbf{C})$ is a monotonic colour distance measure. $\mathbf{W}$ and $\mathbf{H}$ are not defined in the paper. The two page paper does not include enough details to be seriously evaluated.

5.3 Split and merge

5.3.1 A segmentation algorithm for colour images

Author(s): R. Schettini ([64]).

Year: 1993.

Region definition(s): Splitting is made by sequential histogramming for five colour features (as in Ohlander's paper [52], but with another peak detection technique). Merging of adjacent regions which are the most similar.

Basic mathematical method(s): Histogram analysis using fixed scaling. Fisher's distance of regions.

Colour space(s): CIE(Luv) and chromatic features as defined by formulas 16.

Additional knowledge: None.

Application(s): General purpose.

Description: At the fixed scale ($\tau = 2, 4, 6$) peak of histogram is chosen using Tominaga's criterion: proper peak maximises

$$f = \frac{S_p}{T} \frac{R}{F_p}$$

where S_p is the area of the peak counted at the half maximum of the peak; F_p is the distance between half maximum arguments; T is the overall area of the histogram; R is full range of the histogram.

The best peak is chosen for all five histograms. All pixels with colour between arguments giving half peak are taken to one region from which connected components are extracted. The procedure is repeated recursively for all remaining

pixels. At the end of splitting phase, the very small regions are merged to adjacent one with the closest colour.

Global dissimilarity measure S_{ij} of two adjacent regions is defined using Fisher's distance FD_{ij} and region adjacency measure C_{ij}:

$$
\begin{aligned}
S_{ij} &\doteq FD_{ij}C_{ij}/4 \;, \\
C_{ij} &\doteq \min(P_i, P_j)/(4P_{ij}) \;, \\
P_i &= \text{perimeter of } i\text{-th region} \;, \\
P_{ij} &= \text{perimeter of common boundary } i\text{-th and } j\text{-th regions} \;, \\
FD_{ij} &\doteq \max_{k=L,u,v} FD_{ijk}
\end{aligned}
\tag{31}
$$

The Fisher's distance of i-th and j-th regions at k-th feature is defined as follows:

$$
\text{if } (s_{ik}^2, s_{jk}^2) \neq (0,0) \text{ then } FD_{ijk} = \frac{\sqrt{N_i + N_j}|\bar{x}_{ik} - \bar{x}_{jk}|}{\sqrt{N_i s_{ik}^2 + N_j s_{jk}^2}} \;,
$$

$$
\text{if } s_{ik}^2 = s_{jk}^2 = 0 \text{ then } FD_{ijk} = |\bar{x}_{ik} - \bar{x}_{jk}|
$$

Two adjacent regions are *candidates for merging* if the following inequalities are true:

$$
FD_{ij} < 4, \; 0.5 \leq C_{ij} \leq 2, \; S_{ij} < 1 \;.
$$

In one step of the merging phase two adjacent regions are joined if their dissimilarity measure S_{ij} is minimum. The merging phase is terminated if there is no candidates for merging.

5.3.2 Colour images' segmentation using scale space filter and Markov random field

Author(s): C.L. Huang, T.Y. Cheng, and C.C. Chen ([34]).

Year: 1992.

Region definition(s): Coarsely segmented image produced by scale space filtering of histograms is next segmented using Markow random field with region labels.

Basic mathematical method(s): Histogram analysis using scale space filtering and fingerprints. Hidden Markov random fields.

Colour space(s): RGB.

Additional knowledge: None.

Application(s): General purpose.

Description: The paper does not include enough algorithmic details for construction energy function of random Markov field.

5.3.3 Unsupervised segmentation of textured colour images using Markov random field models

Author(s): D.K. Panjwani and G. Healey ([55]).

Year: 1993.

Region definition(s): Uniform blocks are obtained by splitting initial fixed size blocks. Next conservative merging is applied. Finally, agglomerative clustering using global functional expressed in Gaussian Markov random field.

Basic mathematical method(s): Gaussian Markov random field.

Colour space(s): RGB.

Additional knowledge: None.

Application(s): General purpose.

Description: Two page paper does not include clear idea and enough details for Markov field optimisation to be seriously evaluated.

5.3.4 Colour segmentation using perceptual attributes

Author(s): D.-C. Tseng and C.-H. Chang ([74]).

Year: 1992.

Region definition(s): Images are first split into chromatic and achromatic regions. Then they are further split by histogram thresholding and merged using an 8x8 operator.

Basic mathematical method(s): Region splitting using histogram thresholding and merging of regions employing region growing.

Colour space(s): HSI.

Additional knowledge: None.

Application(s): General purpose.

Description: The perceptual HSI space is employed for image segmentation. In analogy to human colour perception three problems occur in colour segmentation when using the hue attribute. First, hue is meaningless when the intensity is very low or very high. Second, hue is unstable when the saturation is very low, and, third, saturation is meaningless when the intensity is very low or very high. Tseng and Chang suggest to split up the colour image into chromatic and achromatic areas to determine effective ranges of hue and saturation for colour image segmentation. The criteria for achromatic areas were measured by experimental observation of human eyes and are defined as follows:

Case 1: (intensity > 95) or (intensity ≤ 25)

Case 2: ($81 <$ intensity ≤ 95) and (saturation < 18),

Case 3: ($61 <$ intensity ≤ 81) and (saturation < 20),

Case 4: ($51 <$ intensity ≤ 61) and (saturation < 30),

Case 5: ($41 <$ intensity ≤ 51) and (saturation < 40),

Case 6: ($25 <$ intensity ≤ 41) and (saturation < 60).

After an image has been segmented into chromatic and achromatic regions, chromatic regions are further segmented using hue histogram thresholding. Achromatic regions are segmented using intensity histogram thresholding. If the saturation histogram of a chromatic region has obvious variation distribution, the region will be split based on the saturation histogram. The significance of a peak is measured according to a sharpness function $S = T_p/W_p$, where T_p denotes the total number of pixels between two valleys and W_p denotes the distance between those valleys. A peak is chosen if its S value is greater than the pre-defined thresholds (in the experiments $S > 256$ for hue histograms and $S > 1536$ for intensity histograms). A further process is applied to avoid oversegmentation. An 8x8 mask is evenly divided into sixteen 2x2 sub-masks. If there is at least a chromatic and an achromatic pixel in the 2x2 sub-mask, then the sub-mask has a vote to the dispersion of the mask. A special label is assigned to the 8x8 region if the mask possesses more than seven votes. After convoluting the mask throughout the image, region growing is used to merge the labelled regions with the segmented regions or to form some new regions based on the idea of clustering in the HSI space.

6 Edge based segmentation

Among reviewed papers which are using colours for edge detection, only few utilized detected edge points for contour tracking along a unique colour region. So, only few papers are here related to image segmentation through edge points detection. However, since we are interested in use of colour information for the correspondence problem in machine stereo vision and edge points are important in solving this problem, we discuss also the remaining papers.

We can broadly classify edge detection techniques as **local** and **global techniques**. The local technique to determine an edge point needs only information in the neighbourhood of that point. The global technique, on the contrary, makes a sort of global optimisation, and therefore the given edge point could be identified after many optimisation steps involving changes in large areas. All global techniques in this section are based on different approaches to Markov random fields.

6.1 Local techniques

6.1.1 Application of colour information to visual perception

Author(s): M. Yachida and S. Tsuji ([80]).

Year: 1971.

Region definition(s): Region is a connected set of pixels having almost the same colour.

Basic mathematical method(s): Straight line fitting while contour tracking.

Colour space(s): rgb. The colour equivalence relation is used: (r_1, g_1, b_1) is *almost the same* as (r_2, g_2, b_2) iff $\max(|r_1 - r_2|, |g_1 - g_2|, |b_1 - b_2|) < T$.

Additional knowledge: The scene consists of commercial building blocks. Each block is a multiface volume of uniform colour. Blocks in the scene have different colours.

Application(s): Robot control for picking up objects of known models (i.e. with given uniform colour).

Description: In one stage of the segmentation process one region is extracted and replaced by the background colour. The stage consists of the following steps:

1. Start the raster scan from the top leftmost point of recently removed region. Visit only every S-th pixel in the row and in the column;

2. Stop scanning at the first pixel p with colour C_p different from the background colour C_B;

3. If p has at most three 8-neighbours with colours different from C_B then consider p as the noise pixel and continue scanning in step 2;

4. Go back in the same row to find the exact boundary pixel P.

5. Find in the vicinity of P, 16 not background pixels and calculate the average colour C_R as estimate of region average colour;

6. Track contour using C_R and *almost the same colour* criterion. Looking for the next boundary point start always from the direction $D - 3$ where D was the previous move direction ($D = 0, \ldots, 7$). Collect all boundary points in a list δR.

7. Fit straight lines to the boundary δR using LSM (Least Square Method): take the first two points from δR and create temporary line L. For any next point Q in δR if Q fits to the line L, i.e. $dist(Q, L) < T$, then include Q by modifying line L using LSM, otherwise Q is considered as a candidate for end of line point. If the next point fits L then Q is not considered any more, otherwise Q is the beginning of the new line.

The threshold T depends on the position n in the list δR of the current point Q:

$$T = \begin{cases} 2 - 0.1n & \text{if } 0 \leq n \leq 10 \ , \\ 0.9 & \text{if } n \geq 11 \ . \end{cases}$$

6.1.2 Colour edge detection

Author(s): T.L. Huntsberger and M.F. Descalzi ([36]).

Year: 1985.

Region definition(s): Edge points are output only.

Basic mathematical method(s): Fuzzy c−means clustering algorithm.

Colour space(s): RGB, CIE(XYZ) with a Riemanian distance.

Additional knowledge: None.

Application(s): General purpose.

Description: Colour edge detection is performed on the basis of fuzzy membership functions for regions obtained in a fuzzy segmentation algorithm. For instance we could take the algorithm described in the section 4.3.2. However, here we take only membership functions, ignoring final steps of defuzzification and connected components extraction.

For the pixel k which belongs to the region R_i in degree u_{ik} and to the region j in degree u_{jk}, edge measure is defined:

$$\text{EDGE}_{ij}(k) \doteq u_{ik} - u_{jk} \ .$$

For each pair of regions R_i and R_j we identify edges as zero crossings of edge measure EDGE_{ij}. Indeed, if we take adjacent pixels k_1 and k_2 belonging to *hard regions* R_i and R_j, respectively, then $\mathrm{EDGE}_{ij}(k_1) \geq 0$ and $\mathrm{EDGE}_{ij}(k_2) \leq 0$. Hence there is a zero crossing for the edge measure EDGE_{ij} when we pass the border between two hard regions.

Note that in practice we have at most 16 pairs of regions to be considered ij, since in the fuzzy segmentation algorithm from section 4.3.2 only four classes of pixels are considered in the same time.

The measure of edge sharpness was introduced as follows:

$$\mathrm{SHARPNESS}_k = \sum_l |\mathrm{EDGE}_k - \mathrm{EDGE}_l| \, ,$$

where l is changed over 8-neighbours of pixel k. Small sharpness measure means diffuse edge, while big one means really sharp edge. Detecting diffuse edges could be useful for detecting of shadows in the image since penumbra produce diffuse edges.

6.1.3 A colour edge detection and its use in scene segmentation

Author(s): R. Nevatia ([50]).

Year: 1977.

Region definition(s): Edges are linked and thus contours are created.

Basic mathematical method(s): Hueckel operator, i.e. Fourier expansion in radial basis functions.

Colour space(s): Yrg.

Additional knowledge: None.

Application(s): General purpose.

Description: In this approach the colour edge at the point P is locally modelled by: r – its distance from P, θ – the angle between the normal to the edge (led from P) and the x axis, (c_1, c_2, c_3) – the edge colour on the side of P, $(c_1 + h, c_2 + h, c_3 + h)$ – the edge colour on the opposite side of P.

All the above edge parameters, represented by the vector ξ, can be found by minimisation of the modelling error function calculated for a neighbourhood $K(P)$, between the image colour function $f(\cdot)$ and the ideal edge step function $E(\cdot; \xi)$:

$$\mathrm{Error} = \sum_{p \in K(P)} \|f(p) - E(p; \xi)\|^2 \, .$$

The author attacks these nonlinear optimisation problem using Hueckel approach ([35]) who applied radial Fourier transform to detect gray scale edges. It leads to a polynomial equation of high degree. Therefore Nevatia recommends applying Hueckel operator for each spectral component seperately and then averaging θ and r. We get the edge parameters as folows:

$$\xi = \left(c_1, c_2, c_3, h_1, h_2, h_3, \theta = \frac{\theta_1 + \theta_2 + \theta_3}{3}, r = \frac{r_1 + r_2 + r_3}{3} \right) \quad .$$

The linking of colour edge points into edge segments, is made only for neighbouring pixels which have: orientations within specified limits of each other and compatible colour characteristics, i.e. the angle difference between space and colour vectors are below a threshold. Further, only edge segments that contain at least a certain minimum of edge points are preserved.

6.1.4 Colour edge detection using 3D median

Author(s): W. Skarbek

Year: $****$

Region definition(s): Edge pixels are only detected.

Basic mathematical method(s): Vector median filtering.

Colour space(s): rgb with Euclidean distance.

Additional knowledge: None.

Application(s): General purpose.

Description: For each pixel p find the direction of the window of size $k \times k$ where vector median for upper half m_1 and lower half m_2 differ the most (say by $D = \|m_1 - m_2\|^2$). If $D > T$ for some threshold T, then memorize the pixel p, the direction n and the colour difference D.

Typically $k = 7, 9$ and the number of quantized space directions is $4, 6$.

6.1.5 Edge based colour segmentation in the CIE(Lab) space

Author(s): S. Lanser ([44]).

Year: 1993.

Region definition(s): Connected set of pixels bounded by edge pixels creating a colour contour.

Basic mathematical method(s): Complementing the edge detection result with additional region growing.

Colour space(s): CIE(Lab).

Additional knowledge: None.

Application(s): General purpose.

Description: The CIE(Lab) colour space is selected for image segmentation. First, the edges are detected separately in every channel due to the following scheme. The first derivative of the image function is computed in one channel applying the modified Deriche operator ([43]). Afterwards, a threshold function is applied to the results instead of computing the maxima of the first derivative. In addition, the zero crossings of the second derivative of the image function are computed to get a second set of edge elements. The contours are defined by the intersections of both sets of edge elements. The two pixels which are adjacent to the ends of an edge and which show the strongest amplitudes are added to the contours to fill gaps. The contours in the colour image are defined as the union of the intermediate contours extracted in every channel. Regions are defined as the complement of this contour image. Afterwards, a region growing technique is applied to fill gaps in the regions. Non-attached pixels are assigned to the adjacent region where the difference between the image signals is minimal.

6.1.6   A new chromatic edge detector used for colour image segmentation

Author(s): M. Chapron ([14]).

Year: 1992.

Region definition(s): Only edge pixels are detected.

Basic mathematical method(s): The maxima of a combination of elementary chromatic gradients are computed.

Colour space(s): RGB.

Additional knowledge: None.

Application(s): General purpose.

Description: The algorithm is a modification of an approach suggested by DiZenzo ([17]). DiZenzo proposed a combination of three chromatic gradients for getting a global gradient of the colour image. With DiZenzo's notations, x be any pixel, f the global luminance of the three image channels with $x = (x^1, x^2), f = (f^1, f^2, f^3)$, and $y = f(x) = (f^1(x), f^2(x), f^3(x)) = (y^1, y^2, y^3)$.

$\frac{\partial f^j}{\partial x^h}$ is supposed to be of rank 2 everywhere in the image, $h = 1,2$ and $f_h(x) = (\frac{\partial f^1}{\partial x^h}, \ldots, \frac{\partial f^3}{\partial x^h})$. $f_h(x)$ and its first derivatives are assumed continuous. Let X the scalar product in R^3. For $h, k = 1, 2 : g_{hk}(x) = f_h(x) \cdot f_k(x)$. The four numbers $g_{hk}(x)$ are the components of the symmetric tensor $g(x) = g(x^1, x^2)$ of rank 2. The term

$$df^2 = \sum_{h=1}^{2} \sum_{k=1}^{2} g_{hk} dx^h dx^k$$

has to be maximised under the constraint

$$\sum_{h=1}^{2} dx^h dx^k = 1.$$

This problem can be seen as the problem of determining the value Θ maximising the form

$$F(\Theta) = g_{11} cos^2(\Theta) + 2g_{12} cos(\Theta) sin(\Theta) + g_{22} sin^2(\Theta) .$$

The maximum is obtained when $\frac{dF}{d\Theta}$ is equal to 0. This corresponds to $\Theta = \frac{1}{2} arctan \frac{2g_{12}}{g_{11} - g_{22}}$. Θ and $\Theta \pm \pi/2$ are solutions and the one which maximises $F(\Theta)$ is chosen. This maximum gives the magnitude of the chromatic gradient and the direction of the gradient is Θ which corresponds to this maximum. g_{11}, g_{12} and g_{22} can be written for colour image processing as follows:

$$u = \frac{\partial R}{\partial x}\mathbf{r} + \frac{\partial G}{\partial x}\mathbf{g} + \frac{\partial B}{\partial x}\mathbf{b}, \tag{32}$$

$$v = \frac{\partial R}{\partial y}\mathbf{r} + \frac{\partial G}{\partial y}\mathbf{g} + \frac{\partial B}{\partial y}\mathbf{b}, \tag{33}$$

$$g_{11} = u \cdot u = \left|\frac{\partial R}{\partial x}\right|^2 + \left|(\frac{\partial G}{\partial x}\right|^2 + \left|(\frac{\partial B}{\partial x}\right|^2, \tag{34}$$

$$g_{22} = v \cdot v = \left|(\frac{\partial R}{\partial y}\right|^2 + \left|\frac{\partial G}{\partial y}\right|^2 + \left|(\frac{\partial B}{\partial y}\right|^2, \tag{35}$$

$$g_{12} = u \cdot v = \frac{\partial R}{\partial x}\frac{\partial R}{\partial y} + \frac{\partial G}{\partial x}\frac{\partial G}{\partial y} + \frac{\partial B}{\partial x}\frac{\partial B}{\partial y}. \tag{36}$$

The four numbers g_{11}, g_{22}, g_{12} and g_{21} are the components of the symmetrical tensor $g(x) = g(x^1, x^2)$ of rank 2. A hysteresis threshold is applied to the magnitude of the global gradient. The difference to the approach proposed by DiZenzo is that elementary gradients

$$\frac{\partial R}{\partial x}, \frac{\partial R}{\partial y}, \frac{\partial G}{\partial x}, \frac{\partial G}{\partial y}, \frac{\partial B}{\partial x}, \text{ and } \frac{\partial B}{\partial y}$$

are computed instead of Sobel gradients as DiZenzo does. Regions are defined as pixels being inside a contour.

6.1.7 Detection and segmentation of man-made objects in outdoor scences: concrete bridges

Author(s): D.C. Baker, S.S. Hwang, and J.K. Aggarwal ([7]).

Year: 1989.

Region definition(s): Connected set of pixels bounded by edge pixels creating a rectangle.

Basic mathematical method(s): Detection of edges, straight lines and rectangles in the image. Confidence test of the mean spectral values in the CIE(Lab) space regarding the training sample, for example, a concrete bridge.

Colour space(s): CIE(Lab).

Additional knowledge: Bridge and non-bridge. Geometry and spectral properties are known for the object to be recognized.

Application(s): Autonomous vehicles, automatic recognition of military targets, recognition of man-made objects in nonurban scenes.

Description: An edge map is obtained from the Laplacian of Gaussian employed to the intensity image. Rectangular areas in the image are detected using straight line fitting and a parallelism criterion. Then colour is used to reject rectangles occurring in natural, i.e., nonartificial, portions of the scene. For this pupose, the image is transformed to the Lab-space or LHC-space, respectively (compare Celenk ([13]) or subsection 4.2.5 for definition). The mean values of the L, H, and C components of all pixels inside a detected rectangle are computed. The confidence functions are obtained by fitting superquadrics to the mean data obtained by manual qualification of the rectangles in the training set.

6.1.8 Simulation of human retinal function with the Gaussian derivative model

Author(s): R. Young ([81]).

Year: 1986.

Region definition(s): Connected set of pixels bounded by edge pixels that are detected employing the DOOG operator.

Basic mathematical method(s): Apply the DOOG operator to each opponent colour pair and detect zero-crossings.

Colour space(s): Opponent colour pairs red-green and yellow-blue.

Additional knowledge: None.

Application(s): General purpose.

Description: Young ([81]) suggest to model shapes of receptive fields by difference-of-offset-Gaussians (DOOG) based on physiological findings at higher cortical levels. The DOOG operator or "Helmholtzian" of a Gaussian $H * G$, respectively, is defined by

$$O(x,y) = H * G * I = k_1 G * I(x,y) - k_2 \nabla^2 G * I(x,y) \ ,$$

where O is the output of the cell (or filter), I is the intensity distribution coming into the eye (or camera), ∇^2 is the Laplacian, $*$ denotes convolution, and k_1 and k_2 are constants whose relative balance can vary adaptively depending upon the particular cell (or filter). Zero-crossings in the output image O define edges. Colours are well segmented by the model, in such a way that colours which appeared to the human observer to be of the hues red, green, blue, or yellow in the original image are successfully segmented into different channels, when the colour offset difference signals are formed according to the DOOG model.

6.1.9 Comparing colour edge detection and segmentation methods

Author(s): J.T. Allen and T. Huntsberger ([2]).

Year: 1989.

Region definition(s): Connected set of pixels bounded by edge pixels creating a colour contour.

Basic mathematical method(s): Colour edge detection, histogram based splitting, and fuzzy c-means segmentation.

Colour space(s): RGB, Yrg, $I_1 I_2 I_3$ (defined in Ohta's paper ([53])), opponent colour pairs red-green and yellow-blue.

Additional knowledge: None.

Application(s): General purpose.

Description: Four proposed methods for colour image segmentation are compared. The first two techniques are the edge detection methods of Nevatia ([50]), described in 6.1.3, and Young ([81]), described in 6.1.8. The latter pair of methods define boundaries as consequence of segmenting the image into regions of similarity. They are the histogram based technique of Ohta, Kanade, and Sakai ([53]), described in 4.1.3, and the fuzzy c-means technique of Huntsberger, Jacobs and Cannon ([36]), described in 4.3.2, and Huntsberger and

Descalzi ([37]), described in 6.1.2. The criteria for evaluating the performance of the methods are accuracy, numerical stability, and robustness to noise. All four methods are applied to artificially generated and natural scene images. The synthetic images are systematically corrupted by random noise and the performance of the methods is studied dependent on the amount of noise. For this purpose an average figure of merit is derived from the measure proposed by Abou and Pratt ([1]). The measure F is based on the displacement of each detected edge pixel from its known correct position. A scaling factor rewards thin, complete but offset edges over edges both thick, smeared edges and broken sparse edges. It is defined by

$$F = \frac{1}{\max(I_C, I_A)} \sum_{i=1}^{I_C} \frac{1}{1 + \alpha d^2(i)},$$

where I_C and I_A are the numbers of the correct and actual edge points, respectively, $d(i)$ is the distance of the i-th edge element from the actual location of the edge, and α is a scaling constant. Abou and Pratt's scaling $\alpha = 1/9$ is used throughout. As a result for the synthetic images, the performance of the fuzzy c-means and Ohta's method is best up to 10% noise ($F > 0.8$) and decreases dramatically thereafter. On the other hand, the Young method and the Nevatia method behaved about the same across the noise spectrum ($F \approx 0.5$ for Nevatia and $F \approx 0.6$ for Young). The performance of the four methods applied to two natural scene images is assessed by visual inspection. The Nevatia operator yields considerable smearing of edges in both images. The multiple edge or ringing effect is very strong nearly everywhere in the images. The tests on the natural images suggest that the Young operator is sensitive only to colour contrast and is largely unperturbed by the kinds of local variations in intensity information. This is not an advantage since it generally means that only sparse edges are detected in the image. The authors mentioned the operator's failure to detect a number of features in the natural images. The Ohta method showed the best performance in this comparison. The natural images are nearly well segmented except in some small parts in one of the images. The fuzzy c-means method seems to be extremely sensitive to local variations corresponding to textured information in the image. Areas in the images without texture are well segmented. Opposed to this, the method failed to segment textured areas. In conclusion, Ohta's method behaved best in this comparison but more detailed investigations are necessary to evaluate the four techniques.

6.1.10 A comparison of three colour image segmentation algorithms in four colour spaces

Author(s): J. Gauch and C.-W. Hsia ([22]).

Year: 1992.

Region definition(s): Regions are homogeneous set of pixels. Due to the underlying algorithm homogeneity is measured based on the Euclidean distance in the colour space, the covariance matrix, or the location of a pixel inside or outside a bounding contour, respectively.

Basic mathematical method(s): Seed based region growing, edge based segmentation using a directional derivative operator, and a recursive split and merge technique.

Colour space(s): RGB, YIQ, HSL, CIE(Lab).

Additional knowledge: None.

Application(s): General purpose.

Description: The three considered segmentation algorithms are: 1) seed-based region growing, 2) directional derivative based edge detection, and 3) recursive split and merge. The colour region growing algorithm defines a region of interest by adding those adjacent pixels to an existing region that are "close" to the average colour of the region calculated so far. A normalised Euclidean distance metric is used to measure colour similarity. The colour edge detection algorithm locates edges by computing the zero-crossings of the directional derivative of the gradient magnitude function in the gradient direction. The Euclidean distance metric associated with each colour space is employed to estimate the partial derivatives necessary to calculate the colour gradients. For sake of simplicity only simple finite differences are used for the gradient estimation. The colour split and merge algorithm recursively subdivides regions whose colour variance is above one threshold and then merges adjacent regions whose combined colour variance is below a second threshold. All three colour segmentation algorithms were applied to two test images in the four colour spaces. The images are an outdoor road image and a synthetic image representing six different colour patches. The performance of the three methods varies with the colour space. Furthermore, the results depend on the image contents. While the region grower was more effective in the RGB and YIQ space than in the HSL and Lab space for the synthetic image, the opposite is true for the road image. The edge detection results in the synthetic image indicate good results for some patches and poorer irregular results for some other patches in all four colour spaces. For the road image, the outlines of the road are more irregular in the RGB and YIQ images than in the HSL and Lab images when applying the edge detection scheme. The split and merge results indicate that segmentation in YIQ and RGB performs better than in the Lab and HSL space for the synthetic image. Opposed to this, the road region has the best shape in the HSL image as compared to the other colour

spaces. In conclusion, the split and merge algorithm performed best in this comparison but no single colour space was best for all three segmentation algorithms. Nevertheless, a more detailed investigation including more test images is needed to evaluate the three algorithms and their performance in the four colour spaces.

6.2 Global techniques

6.2.1 Colour image segmentation using Markov random fields

Author(s): M.J. Daily ([16]).

Year: 1989.

Region definition(s): Edge points are only detected.

Basic mathematical method(s): Markov random field with edge pixels. Hopfield neural network optimisation of energy for Markov random fields.

Colour space(s): RGB, CIE(Lab) with Euclidean distance; HSI with max distance, but if $\Delta I < 25$ then the distance is ΔI.

Additional knowledge: None.

Application(s): General purpose.

Description: In this research the original image (d_{ij}) is smoothed to (f_{ij}) and in the same time edge elements, horizontal (h_{ij}), and verical (v_{ij}) are detected.

The search for output image and edge elements is performed by the minimisation of the energy functional E:

$$E = E_i + E_d + E_l + E_g \ .$$

On the basis of Hammersley–Clifford theorem ([8, 23]) the minimisation of E is equivalent to searching of maximum *a posteriori* probability (MAP) for a random Markov field modelling the given image. In such a model the image content together with edge elements content create one global state. The probability of this state is governed by the Gibbs distribution with the energy function calculated by summing the local energies over the entire image.

The interpolation term E_i of the energy is defined with weighting coefficient λ:

$$E_i = \lambda \sum_{i,j} ((f_{i,j-1} \ominus f_{i,j})^2 (1 - v_{i,j}) + (f_{i+1,j} \ominus f_{i,j})^2 (1 - h_{i,j})) \ ,$$

where $\ominus$ denotes colour distance operation.

The data term E_d of the energy with weighting coefficient α has the form:

$$E_d = \alpha \sum_{i,j} \| f_{i,j} - d_{i,j} \|^2 \ .$$

The line term E_l of the energy E has the weight β:

$$E_l = \beta \sum_{i,j} (v_{i,j} + h_{i,j}) \ .$$

The last term E_g with weight γ forces the edge strength towards hard limits, i.e. to zero or one:

$$E_g = \gamma \sum_{i,j} \left(\int_0^{v_{i,j}} g^{-1}(v) dv + \int_0^{h_{i,j}} g^{-1}(h) dh \right) \ ,$$

where g is a standard sigmoid function.

Using Hopfield's idea ([33]), we get iterative formulas for $f_{i,j}, v_{i,j}, h_{i,j}$:

$$v_{i,j} = g \left(\frac{\lambda(f_{i,j+1} \ominus f_{i,j})^2 - \beta}{\gamma} \right) \ ,$$

$$h_{i,j} = g \left(\frac{\lambda(f_{i+1,j} \ominus f_{i,j})^2 - \beta}{\gamma} \right) \ ,$$

$$f_{i,j} = \frac{\lambda(L^v_{i,j} f_{i,j+1} + L^v_{i,j-1} f_{i,j-1} + L^h_{i,j} f_{i+1,j} + L^v_{i-1,j} f_{i-1,j}) + \alpha d_{i,j}}{\lambda(L^v_{i,j} + L^v_{i,j-1} + L^h_{i,j} + L^h_{i-1,j}) + \alpha}$$

where $L^v_{i,j} = 1 - v_{i,j}$ and $L^h_{i,j} = 1 - h_{i,j}$.

This set of equations must be iterated until convergence. However, in practice, it was found that updating the line process once every 10 colour process updates is adequate.

6.2.2 Toward color image segmentation in analog VLSI: algorithm and hardware

Author(s): F. Perez and C. Koch ([56]).

Year: 1994.

Region definition(s): Edge points are only detected.

Basic mathematical method(s): Canny edge operator. Markov random field with line process.

Colour space(s): HSI and rgb.

Additional knowledge: None.

Application(s): General purpose.

Description: Algorithm starts with the intensity edge map generated from the Canny edge operator and gradually eliminates those edges not due to hue differences. The constraint is that hue disconuity would only form an edge element if pixel saturation value exceeded a threshold (20% saturation). In doing so, the authors are discounting edges due to confounding cues of highlights, shading, transparency, and shadows for moderately chromatic scenes.

Markov random field formulation

Let $(D_{i,j})$ be the hue feature of the input image and $(H_{i,j})$.

The following energy functional is minimised:

$$E = E_d + E_v + E_h + E_l \ ,$$

where

$$
\begin{aligned}
E_d &= \lambda \sum_{i,j} (H_{i,j} \ominus D_{i,j})^2 \ , \\
E_v &= \sum_{i,j} (H_{i,j+1} \ominus H_{i,j})^2 (1 - v_{i,j}) \ , \\
E_h &= \sum_{i,j} (H_{i+1,j} \ominus H_{i,j})^2 (1 - h_{i,j}) \ , \\
E_l &= \alpha \sum_{i,j} (h_{i,j} + v_{i,j}) \ .
\end{aligned}
$$

Here $v_{i,j}, h_{i,j}$ are edge elements which take only values 0 or 1. The minimisation algorithm utilizes a first–order Tikhonov stabilizing functional (Poggio [58]).

6.2.3 A Markov random field approach to data fusion and colour segmentation

Author(s): W.A. Wright ([78]).

Year: 1989.

Region definition(s): Edge pixels are only determined.

Basic mathematical method(s): Markov random field extended to cliques between colour components.

Colour space(s): RGB.

Additional knowledge: None.

Application(s): General purpose.

Description: Energy function considers not only cliques of points, but also cliques of edges inside of spectral components and cliques of coincidence relation between spectral components (image planes). Broadly speaking energy is decreased if an edge occurs in each of the spectral images, otherwise it is increased.

The minimisation problem is solved using simulated annealing schedule as described in Geman ([23]).

7 Physics based segmentation

In this section we discuss segmentation techniques employing physical models to partition an image into regions that correspond to surfaces or objects in the scene. The objective of these techniques is to segment a multispectral image at object boundaries and not at the edges of highlights and shadows in an image ([12]). This is a difficult task since the image measurements corresponding to a single surface can have considerable variation due to effects such as highlights, shading, sensor noise, nonuniform illumination, and surface texture ([30]). Most of the techniques discussed in the previous sections use uniformity criteria for region definition that are based on similar colour values corrupted by noise. Unfortunately, these algorithms often produce bad results on real images. For example, highlights may be segmented as separate objects in the scene or a curved object may be segmented into more than one surface due to shading that causes intensity variation ([30]). Physics based segmentation techniques allow the segmentation of real images based on physical models for image formation. The basic mathematical methods used by physic based techniques are often similar to those already discussed in the previous sections. For example, Healey ([28]) uses region splitting guided by preliminary edge detection to classify regions. The approaches for physics based segmentation and the approaches already mentioned in the previous sections do not differ regarding the basic mathematical techniques used for segmentation. They differ regarding the reflection models employed for segmenting colour images. Therefore, physics based segmentation is so far limited to determining changes in materials whose reflection properties are well known and can be modelled properly. Some approaches intend to be applied preliminary to the segmentation process. They try, for example, to distinguish material changes from shadow and/or highlight boundaries. Such classification is needed to obtain segmentation results representing surfaces and objects in the scene. The available papers can be broadly divided into two groups:

1. approaches designed exclusively for **inhomogeneous dielectrics**: images are segmented into regions representing the same material by employing the dichromatic reflection model;

2. **general approaches** that are not limited to inhomogeneous dielectrics: regions correspond to different materials which are, for example, metal, plastic, etc.

7.1 Inhomogeneous dielectrics

7.1.1 Image segmentation and reflection analysis through colour

Author(s): G.J. Klinker, S.A. Shafer, and T. Kanade ([40, 41, 42]).

Year: 1988, 1990.

Region definition(s): Clusters in the colour space are determined due to their characteristic form and due to physic based hypothesis. Pixels with colour vectors located inside this clusters belong to the same region.
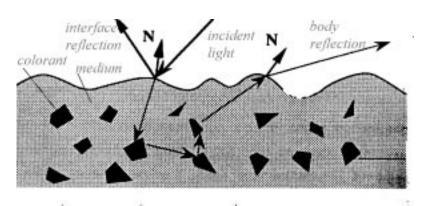
Basic mathematical method(s): Searching for colour clusters from local image areas that show the characteristic features of the body and surface reflection processes. Then region growing is applied.

Colour space(s): RGB.

Additional knowledge: Inhomogeneous dielectrics.

Application(s): Determining the number of materials in the scene.

Description: The authors investigated the influence of highlights, shading, and camera properties as, for example, colour clipping, blooming, colour balancing, and chromatic lens aberration on the results of colour image segmentation. They classify physical events with measured colour variation in the image by employing the dichromatic model for reflection from dielectrics. Shafer's Dichromatic Reflection Model ([66]) describes the light, $L(\lambda, i, e, g)$, which is reflected from a point on a dielectric, nonuniform material as a mixture of the light $L_s(\lambda, i, e, g)$ reflected at the material surface and the light $L_b(\lambda, i, e, g)$ reflected from the material body (see the following illustration drawn after ([40])).

The parameters $i, e,$ and g denote the angles of the incident and emitted light and the phase angle; λ is the wavelength. The Dichromatic Reflection Model is given by

$$L(\lambda, i, e, g) = L_s(\lambda, i, e, g) + L_b(\lambda, i, e, g).$$

Furthermore, the spectral reflection properties of L_s and L_b are separated from their geometric reflection properties by modelling them as products of spectral power distributions, $c_s(\lambda)$ and $c_b(\lambda)$, and geometric scale factors, $m_s(i, e, g)$ and $m_b(i, e, g)$, which describe the intensity of the reflected light. Thus, the

upper equation becomes

$$L(\lambda, i, e, g) = m_s(i, e, g)c_s(\lambda) + m_b(i, e, g)c_b(\lambda).$$

The authors show, for example, that a colour cluster containing matte and highlight pixels looks like a "skewed T" in colour space. Using this classification a hypothesis-based segmentation algorithm is developed. The algorithm looks, in a *bottom-up* process, for colour clusters from local image areas that show the characteristic features of the body and surface reflection processes. When it finds a "promising" cluster in an image area, a hypothesis is generated which describes the object colour and/or highlight colour in the image area and determines the shading and highlight components of every pixel in the area. The algorithm then applies the new hypothesis to the image, using a region-growing approach to determine the exact extent of the image area to which the hypothesis applies. This step verifies the applicability of the hypothesis. Thus, the analysis consists of many small interpretation cycles that combine *bottom-up* processing with response in *top- down* processing. Accurate segmentation results are presented for images of plastic objects.

7.1.2 Colour image segmentation with detection of highlights and local illumination induced by inter-reflections

Author(s): R. Bajcsy, S.W. Lee, and A. Leonardis ([4, 5, 6])).

Year: 1989, 1990.

Region definition(s): Clusters are detected based on peaks and valleys in histograms and based on observations of physical phenomena in the HSI space. Pixels with colour vectors located inside this clusters belong to the same region.

Basic mathematical method(s): Histogram thresholding in the HSI space.

Colour space(s): HSI.

Additional knowledge: Inhomogeneous dielectrics.

Application(s): Determining the number of materials in the scene.

Description: Colour image segmentation is performed in the presence of several confounding physical phenomena. It is based on material changes with detection of highlights and regarding shading, shadow, and small inter-reflections between adjacent objects. The assumption is made that the image consists of patches of object surfaces that have uniform chromaticity (not intensity), i.e., the image can be divided into regions of uniform hue and saturation regardless of the surface structure. The algorithm is derived in HSI space employing

the dichromatic reflection model for inhomogeneous dielectrics assuming that highlights have the same spectral composition as the illumination. The illumination is *whitened* using a reference plate of known spectral reflectance. The authors describe the resulting structure of colour clusters in HSI space corresponding to phenomena such as shading, shadows, highlights and inter-reflection. Based on this analysis, a hue histogram technique is used to segment individual surfaces. Further segmentation is achieved using saturation values based on the observation that

1. shading, highlights, shadow and inter-reflections change the intensity,
2. shading and shadow do not change the hue or the saturation value,
3. highlight decreases the saturation value, and
4. the inter-reflection in general cause the change in hue and saturation.

Thus, colour clusters of highlights and most of the small inter-reflections from those of body reflection can be separated by applying local thresholding on saturation values. Experiments on colour images of glossy objects demonstrate the effectiveness of the approach. The algorithm can be applied to images which are acquired under controlled illumination and which represent inhomogeneous dielectrics.

7.1.3 Colour image segmentation with highlight removal employing a direct approach

Author(s): K. Schlüns and O. Wittig ([65]).

Year: 1993.

Region definition(s): Straight line segments in a chromaticity diagram determine colours and classes of materials. Pixels with spectral components that lie on such line segment create a region.

Basic mathematical method(s): Calculation of line segments in a chromaticity diagram and applying histogram techniques.
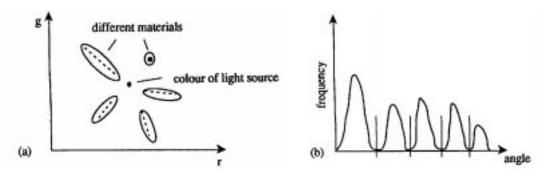
Colour space(s): RGB.

Additional knowledge: Inhomogeneous dielectrics.

Application(s): Determining the number and colours of materials in the scene.

Description: The authors propose a two-dimensional approach to classify RGB colours into groups representing different surface materials. First, the chromaticity components r and g of each pixel are determined by normalisation: r = R/(R+G+B), g = G/(R+G+B). In this representation, materials consisting of a body reflection component L_b and a specular reflection component

66

L_s (for description of symbols see 7.1.1) are characterised by line segments in the (r,g)-chromaticity diagram. The chromaticity of illumination is determined by the intersection of the straight lines that are defined by these line segments when assuming neutral interface reflection. Matte materials appear as a point in the diagram (see left figure). If the scene contains more than one inhomogeneous dielectric material, then the chromaticity of illumination can be calculated from the colour image applying Hough transformation or some other method for line recognition. However, in most applications the illumination colour can be assumed being approximately white. Such rough estimate is sufficient for the following processing steps. For each pixel, the straight line is determined that originates in the locus of the illumination's chromaticity and passes through the chromaticity of the pixel. Its slope is represented using the angle between the line and an arbitrary axis. The angles are simply related to the hue values of the materials and they can be accumulated in a histogram (see right figure). Thus, the classification process is reduced to an one-dimensional problem with many existing, robust solutions. Scattered entries with low frequency, not belonging to the body of a peak, can be further examined regarding inter-reflections by applying a local windowing technique. That end-point of a line segment in the chromaticity diagram, that lies farther away from the locus of the illumination's chromaticity, represents the body reflection chromaticity. This feature can be used, for example, for an illumination independent description of the material's colour and for highlight removal (dichromatic projection).

7.2   General approaches

7.2.1   Colour vision and image intensities: when are changes material ?

Author(s): J.M. Rubin and W.A. Richards ([62]), ([63]).

Year: 1982, 1984.

Region definition(s): Only material changes are detected.

Basic mathematical method(s): Test two conditions at discontinuities in the image.

Colour space(s): At least two spectral samples, e.g.: two channels of the RGB space.

Additional knowledge: Physical reflectance model.

Application(s): Determining material changes in the scene.

Description: Since the objective of physics based segmentation is to segment an image at object boundaries and not at the edges of shadows and highlights, one important requirement is to determine material changes in the scene. Rubin and Richards ([62]) suggest to use a rejection scheme for this purpose. The rejection scheme is based on the idea that it is easier to reject the occurrence of a change in the spectral functions due to shadow, shading, highlight, or a change in pigment density than to detect material changes directly. The assumption is formulated as the *spectral crosspoint condition* to be used for detecting material changes. This condition states that given two regions X and Y across a discontinuity and intensity samples I taken at two wavelengths λ_1 and λ_2 the following is a test for a material change:

$$(I_{X\lambda_1} - I_{Y\lambda_1})(I_{X\lambda_2} - I_{Y\lambda_2}) < 0.$$

This condition was found to be inadequate for finding material changes since spatial and spectral information are strongly intertwined in a crosspoint. Thus, a second condition was introduced, the *opposite slope sign condition* ([63]) formally described as

$$(I_{X\lambda_2} - I_{X\lambda_1})(I_{Y\lambda_2} - I_{Y\lambda_1}) < 0,$$

which again signals a material change. The condition is based on the idea that, given two spectral components measured on both sides of a shadow boundary, on each side the same component must be the maximum. This approach is easy to apply to colour images but the results are often incorrect. Several configurations of the spectral values can occur where a material change exists which can not be detected by testing the conditions mentioned above ([62]). Furthermore, Gershon, Jepson, and Tsotsos ([24]) presented some situations where a material change is detected when testing the above conditions although the discontinuity is caused by shadowing.

7.2.2 Ambient illumination and the determination of material changes

Author(s): R. Gershon, A.D. Jepson, and J.K. Tsotsos ([24]), ([25]).

Year: 1986.

Region definition(s): Only material changes are distinguished from shadow boundaries.

Basic mathematical method(s): Applying several operators to the image components and testing their relative responses.

Colour space(s): RGB.

Additional knowledge: Physical reflectance model.

Application(s): Distinguishing material changes from shadow boundaries.

Description: An approach is presented to distinguish between material changes and shadow boundaries. Employing colour reflection models, the authors analyse the cases of ideal shadows, for which ambient illumination and direct illumination have the same spectral composition, and non-ideal shadows. A pull factor P_f, that quantifies the deviation of a shadow boundary from the ideal case, is defined as

$$P_f = \frac{|gR - rG|}{R^2 + G^2}.$$

This quantity is used to estimate the likelihood of a shadow and a lit region belonging to the same material. The stronger the effect of the additional ambient illumination, the more the shadow region differs from the lit region. Thus, it becomes more unlikely that both regions belong to the same material. Biologically inspired colour operators are introduced to distinguish material changes from shadow boundaries. The operators have spatial centre-surround antagonistic organisation and their response is related to a shadow boundary's pull factor. The response of a monochromatic opponent unit, for example, R+centre/R–surround, is given by

$$RESP(\mathbf{x}; \sigma) = G(\mathbf{x}; \sigma_c) * L_R(\mathbf{x}) - G(\mathbf{x}; \sigma_s) * L_R(\mathbf{x}),$$

where $*$ denotes convolution, $L_R(\mathbf{x})$ is the logarithm of the red component of the input image, and $G(\mathbf{x}; \sigma_i)$ is a Gaussian of the form

$$G(\mathbf{x}; \sigma_i) = \frac{1}{2\pi\sigma_i^2} \exp\left[\frac{-|\mathbf{x}|^2}{2\sigma_i^2}\right].$$

The response of a double-opponent operator, for example, R+G–centre/R–G+surround, is computed as:

$$RESP(\mathbf{x}; \sigma) = G(\mathbf{x}; \sigma_c) * (L_R(\mathbf{x}) - L_G(\mathbf{x})) - G(\mathbf{x}; \sigma_s) * (L_R(\mathbf{x}) - L_G(\mathbf{x}))$$

$$= (L_R(\mathbf{x}) - L_G(\mathbf{x}))DOG(\mathbf{x}),$$

where $L_G(\mathbf{x})$ is the logarithm of the green image component and DOG denotes a DOG filter (with $\sigma_s/\sigma_c = 2.5$ in the experiments ([24, 25]). The responses of the operators are combined to distinguish material changes from shadow

69

boundaries. For instance in the red-green case, the authors compute the relative amplitude response (RAP) which is given by

$$RAP = \frac{|\text{Peak Response of R+G-/R-G+}|}{[(\text{Peak Response of R+/R-})^2 + (\text{Peak Response of G+/G-})^2]^1/2}.$$

If the RAP value is greater than the anticipated pull factor, then the discontinuity in the area is not considered a material change but rather a change caused by shadow effects. Properties of the algorithm are demonstrated by applying it to a colour image of red and green peppers.

7.2.3   Segmenting images using normalised colour

Author(s): G.E. Healey ([26, 27, 28, 29]).

Year: 1989, 1990, 1992.

Region definition(s): Connected set of pixels bounded by edge pixels are assumed to correspond to a single region (single material). Edge pixels are detected by the Canny operator applied to a single band of the image.

Basic mathematical method(s): Region splitting using maximum likelihood estimates to classify the membership of a pixel to a region (material).

Colour space(s): Any multispectral colour space normalised using the L^2 norm, especially, the normalised RGB.

Additional knowledge: No interreflections in the scene, physical reflectance model.

Application(s): Determining the number of materials in the scene.

Description: The main idea is based on a region splitting algorithm using a normalised colour space. Taking into account the physical reflectance properties of metallic surfaces and inhomogeneous dielectric materials, the sensor response depends on the sensor, the illuminant, and the reflecting material, but not on the scene geometry ([28]). In an earlier paper Healey ([27]) normalised the sensor space employing the L^1 norm. He defined a colour metric using the Mahalanobis distance. In 1992, Healey ([29]) suggested to normalise the sensor space by employing the L^2 norm because the Euclidean distance between two material lines depends in this space exclusively on the angle between the two lines. Opposed to that, the Euclidean distance depends also on the locations of the two material lines in the sensor space when the L^1 norm is employed. Due to the normalisation process, dark brown and light brown pixels are, for example, segmented into the same material class. These pixels have spectral reflectance functions that are nearly a scalar multiple of one another. Therefore, it is assumed that they could correspond to a single material viewed under

70

different geometries. All classes of materials are assumed to be equally likely and the distribution of the measured sensor vectors is modelled for each material class c_i using a multivariate normal density function $\mathcal{N}(\mu_i; \Sigma_i)$ described by

$$p(S/c_i) = \frac{1}{(2\pi)^{N/2}\|\Sigma_i\|^{1/2}} \exp^{\frac{-1}{2}(s-\mu_i)^T \Sigma_i^{-1}(S-\mu_i)},$$

where μ_i is the N-element mean vector for material class c_i and Σ_i is the N x N covariance matrix $\Sigma_i = E[(S - \mu_i)(S - \mu_i)^T]$, and $0 \leq i \leq M - 1$ supposing that M materials occur in the scene. The objective is to classify a new sensor vector v as belonging to one of the M material classes. Each distribution $p(S/c_i)$ is normalised to a multivariate normal distribution $\mathcal{N}(\hat{\mu}_i; \hat{\Sigma}_i)$ to suppress the effects of geometry in the scene onto the segmentation process. The normalisation is described by

$$\hat{\mu}_i = \frac{\mu_i}{\|\mu_i\|}, \ \hat{\Sigma}_i = \frac{\Sigma_i}{\|\Sigma_i\|^2}, \text{ and } \hat{v} = \frac{v}{\|v\|}$$

where $\|.\|$ denotes the L^2 norm. The problem is to assign $\hat{v}$ to one of the M material classes described by the distributions $\mathcal{N}(\hat{\mu}_i; \hat{\Sigma}_i)(0 \leq i \leq M - 1)$. From Bayes decision theory, the set of functions

$$g_i(\hat{v}) = \log p(\hat{v}/c_i) + \log p(c_i)$$

serve as minimum error rate discriminant functions for classification. If $p(\hat{v}/c_i)$ is modelled using $\mathcal{N}(\hat{\mu}_i; \hat{\Sigma}_i)$, then a set of discriminant function is obtained given by

$$g_i(\hat{v}) = -\frac{1}{2}(\hat{v} - \hat{\mu}_i)^T \hat{\Sigma}_i^{-1}(\hat{v} - \hat{\mu}_i) - \frac{1}{2}\log|\hat{\Sigma}_i| .$$

The colour image is segmented into several classes of material using the discriminant functions. A mean normalised colour vector is computed for each area containing no edge pixels and the discriminant functions are computed based on the definition above. If, for example, N materials are already found and the values of all corresponding discriminant functions are less than a given threshold then a new material class is predicted. Otherwise, the region is labelled to belong to the existing class for which the value of the discriminant function is the largest. If the measured sensor values inside a region are at least as large in all colour bands as the measured sensor values for a body reflection region at the same scene geometry, the region is merged with the adjacent region to avoid classifying highlight regions as material classes. This approach has been developed for N spectral samples, i.e., it is not limited to RGB images. Moreover, it is easy to imagine that the results in determining different classes of materials improve when using more than three spectral samples. These spectral samples can be easily generated by using a black & white camera and N different spectral filters with low bandpass. Furthermore,

the algorithm is inherently parallel and can be mapped onto high performance parallel hardware ([26]).

7.2.4   Colour image segmentation with simple edge classification

Author(s): A. Koschan

Year: ∗ ∗ ∗∗

Region definition(s): Connected set of pixels bounded by simply classified edge pixels are assumed to correspond to a single region (single material). Edge pixels are detected by the Cumani operator.

Basic mathematical method(s): Detection of colour edges by the Cumany operator. Complementing the edge detection result and merging of regions due to simple observations.

Colour space(s): RGB, HSI.

Additional knowledge: No interreflections in the scene.

Application(s): Determining the number of materials in the scene.

Description: Colour edges are located applying the Cumani operator ([15]) in the RGB space. The operator is based on an extension of edge detectors using the second- order differential derivative to the case of multiple band images. The edge detection result is complemented to obtain bounded regions. Then edges are tried to be classified avoiding the misclassification of highlight, shading or shadow. Pixels adjacent to both sides of an detected edge are compared in the HSI space for this classification. Two cases are distinguished regarding observations presented in ([5]) and ([56]) and based on own experiments. The following conditions are assumed for pixels in two adjacent regions.

1. There is a small change in hue values, a big change in intensity values, and the darker area is slightly less saturated than the lighter area. Then it is very likely that the discontinuity is caused by shadow or shading.

2. There is a big change in intensity. The lighter area has very high intensity and very low saturation. The darker area has average intensity and average saturation. Then it is likely that the discontinuity is caused by highlight.

If the pixels in two regions that are separated by an edge satisfy one of the conditions mentioned above then the regions are merged.

8   Conclusions

From an extensive analysis of available literature on colour image segmentation we can draw some conclusions:

1. Colour images allow for more reliable image segmentation than for gray scale images.

2. Applying of hue feature is very successfull in many applications, however it should be used with care when intensities are small.

3. Most clearly specified, algorithmically efficient, and robust are methods designed for the particular small applications assuming well specified knowledge about the scene.

4. General purpose algorithms are not robust and usually not algorithmically efficient.

5. All techniques are dependent on parameters, constants and thresholds which are usually fixed on the basis of few experiments. Tuning, adapting of parameters is rarely performed.

6. As a rule, authors ignore comparing their novel ideas with existing ones.

7. As a rule, authors do not estimate the algorithmic complexity of their methods.

8. It seems that seperating processes for region segmentation and for object recognition is the reason of failure of general purpose segmentation algorithms.

9. The main problem with the combination of region segmentation and colour object recognition is the need of colour constancy. Although some approaches already exist to achieve colour constancy (e.g.: ([48]), ([21]), ([20]), ([72]), etc.), the methods are computationally expensive and/or they need assumptions that are rather restrictive. The discussion of colour constancy is out of scope of this review but a robust and efficient algorithm is needed for that purpose.

10. Several different colour spaces are employed for image segmentation. Nevertheless, no general advantage of one of the colour spaces with regard to the other colour spaces has been found yet.

11. Physics based approaches try to overcome the problem of misclassifying highlighted and shaded areas in the image. Unfortunately, nearly all techniques can only be employed to a restricted set of materials and, furthermore, they are computationally expensive. Thus, they are far away from either being a general solution to image segmentation nor being suitable for real-time segmentation. Nevertheless, promising results might be expected in the future.

References

[1] Abou I.E. and Pratt W.K. (1979) "Quantitative design and evaluation of enhancement/thresholding edge detectors", Proc. of the IEEE, vol. 67, pp. 756–763.

[2] Allen J.T. and Huntsberger T. (1989) "Comparing color edge detection and segmentation methods", Proc. IEEE 1989 Southeastcon, pp. 722–728.

[3] Amadasun M. and King R.A. (1988) "Low–level segmentation of multispectral images via agglomerative clustering of uniform neighbourhoods", *Pattern Recognition*, vol. 21, no. 3, pp. 261–268.

[4] Bajcsy R., Lee S.W., and Leonardis A. (1989) "Image segmentation with detection of highlights and interreflections using color", Technical Report GRASP LAB 182 MS–CIS–89–39, Dept. of Computer and Info. Science, University of Pennsylvania, USA.

[5] Bajcsy R., Lee S.W., and Leonardis A. (1990) "Color image segmentation with detection of highlights and local illumination induced by inter- reflections", Proc. 10th Int. Conf. on Pattern Recognition 1990, Atlantic City, New Jersey, USA, 16 – 21 June 1990, pp. 785–790.

[6] Bajcsy R., Lee S.W., and Leonardis A. (1990) "Color image segmentation and color constancy", Proc. SPIE vol. 1250 Perceiving, Measuring and Using Color, 1990, pp. 245–255.

[7] Baker D.C., Hwang S.S., and Aggarwal J.K. (1989) "Detection and segmentation of man-made objects in outdoor scences: concrete bridges", *Journal of the Optical Society of America A*, vol. 6, no. 6, pp. 938– 951.

[8] Besag J. (1974) "Spatial interaction and the statistical analysis of lattice systems", *Journal of Royal Statistical Society Ser. B*, vol. 36, pp. 192–236.

[9] Bezdek J.C. (1981) *Pattern recognition with fuzzy objective function algorithms*, Plenum Press, New York.

[10] Bonsiepen L. and Coy W. (1991) "Stable segmentation using color information" in *Computer Analysis of Images and Patterns*, ed. R. Klette, Proc. of CAIP'91, Dresden, September 17–19, 1991, pp.77–84.

[11] Brand M. (1993) "A short note on local region growing by pseudophysical simulation", Proc. CVPR'93, New York City, USA, 15 – 18 June 1993, pp. 782–783.

[12] Brill M.H. (1991) "Photometric models in multispectral machine vision", Proc. SPIE vol. 1453 Human Vision, Visual Processing, and Digital Display II, pp. 369–380.

[13] Celenk M. (1988) "A recursive clustering technique for colour picture segmentation", Proc. Int. Conf. on Computer Vision and Pattern Recognition, Ann Arbor, MI, USA, 5 –9 June 1988, pp. 437–444.

[14] Chapron M. (1992) "A new chromatic edge detector used for color image segmentation", Proc. 11th Int. Conf. on Pattern Recognition, Den Hague, Netherlands, 30 Aug. – 3 Sept. 1992, vol. III, Conf. C: Image, Speech, and Signal Analysis, pp. 311–314.

[15] Cumani A. (1991) "Edge detection in multispectral images", *Computer Vision, Graphics, and Image Processing: Graphical Models and Image Processing*, vol. 53, no. 1, pp. 40–51.

[16] Daily M.J. (1989) "Colour image segmentation using Markov random fields", Proc. CVPR'89, San Diego, Ca., USA, 4–8 June 1989, pp. 304–312.

[17] DiZenzo S. (1986) "A note on the gradient of a multi-image", *Computer Vision, Graphics, and Image Processing*, vol. 33, pp. 116–125.

[18] Ferri F. and Vidal E. (1992) "Colour image segmentation and labeling through multiedit–condensing", *Pattern Recognition Letters*, vol. 13, no. 8, pp. 561–568.

[19] Fesharaki M.N. and Hellestrand G.R. (1993) "Real-time colour image segmentation", Technical Report SCS&E 9316, School of Computer Science and Engineering, University of New South Wales, Australia, Dec. 1993.

[20] Finlayson G.D., Drew V, and Funt B.V. (1993) "Diagonal transform suffice for color constancy", Proc. 4th ICCV, Berlin, Germany, 11 –14 May 1993, pp. 164–171.

[21] Forsyth D.A. (1990) "A novel algorithm for color constancy", *Int. Journal of Computer Vision*, vol. 5, no. 1, pp. 5–35.

[22] Gauch J. and Hsia C.-W. (1992) "A comparison of three color image segmentation algorithms in four color spaces", Proc. SPIE vol. 1818, Visual Communications and Image Processing '92, pp. 1168–1181.

[23] Geman S. and Geman D. (1984) "Stochastic relaxation, Gibss distributions, and the Bayesion restoration of images", *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 6, pp. 721–741.

[24] Gershon R., Jepson A.D., and Tsotsos J.K. (1986) "The effects of ambient illumination on the structure of shadows in chromatic images", Technical Report RBCV–TR–86–9, Dept. of Computer Science, University of Toronto, Ontario, Canada, 1986.

[25] Gershon R., Jepson A.D., and Tsotsos J.K. (1986) "Ambient illumination and the determination of material changes", *Journal of the Optical Society of America A*, vol. 3, no. 10, pp. 1700–1707.

[26] Healey G.E. (1989) "A parallel color algorithm for segmenting images of 3-d scenes", Proc. DARPA Image Understanding Workshop 89, Palo Alto, Ca., USA, 23 – 26 May 1989, pp. 1038–1041.

[27] Healey G.E. (1989) "Color discrimination by computer", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 19, no. 6, pp. 1613–1617.

[28] Healey G.E. (1990) "Using physical color models in 3-d machine vision", Proc. SPIE vol. 1250 Perceiving, Measuring and Using Color, 1990, pp. 264– 275.

[29] Healey G.E. (1992) "Segmenting images using normalized color", eds. G.E. Healey, S.A. Shafer, and L.B. Wolff (1992) "Physics-based vision principles and practice color", Jones and Bartlett Publ., Boston, USA, pp. 166–198.

[30] Healey G.E. (1992) "Color image segmentation", eds. G.E. Healey, S.A. Shafer, and L.B. Wolff (1992) "Physics-based vision principles and practice color", Jones and Bartlett Publ., Boston, USA, pp. 99–100.

[31] Hild M., Shirai Y., and Asada M. (1992) "Initial segmentation for knowledge indexing", Proc. 11th Int. Conf. on Pattern Recognition 1992, Den Hague, Netherlands, 30 Aug.– 3 Sept. 1992, vol. I, Conf. A: Computer Vision and Applications, pp. 587–590.

[32] Holla K. (1982) "Opponent colors as a 2-dimensional feature within a model of the first stages of the human visual system", Proc. 6th Int. Conf. on Pattern Recognition, Munich, Germany, 19 –22 Oct. 1982, pp. 561–563.

[33] Hopefield J.J. and Tank D.W. (1985) "Neural computation of decisions in optimiztion problems", *Biological Cybernetics*, vol. 52, no.14, pp. 141–154.

[34] Huang C.L., Cheng T.Y., and Chen C.C. (1992) "Colour images' segmentation using scale space filter and Markov random field", *Pattern Recognition*, vol. 25, no. 10, pp. 1217–1229.

[35] Hueckel M. (1973) "A local visual operator which recognizes edges and lines", *Journal of the ACM*, vol. 20, no. 4, pp. 634–647.

[36] Huntsberger T.L. and Descalzi M.F. (1985) "Colour edge detection", *Pattern Recognition Letters*, vol. 3, pp. 205–209.

[37] Huntsberger T.L., Jacobs C.L., and Cannon R.L. (1985) "Iterative fuzzy image segmentation", *Pattern Recognition*, vol. 18, no. 2, pp. 131–138.

[38] Ismaili I.A. and Gillies D.F. (1994) "Colour image segmentation using regression analysis in RGB space", *Machine Graphics & Vision*, vol. 3, nos.1/2, pp. 373-384.

[39] Jordan III J.R. and Bovik A.C. (1988) "Computational stereo vision using color", *IEEE Control Systems Magazine*, June, pp. 31–36.

[40] Klinker G.J. (1988) "A physical approach to color image understanding", Ph.D. Thesis, Technical Report CMU–CS–88–161, Computer Science Department, Carnegie Mellon University, USA, May 1988.

[41] Klinker G.J., Shafer S.A., and Kanade T. (1988) "Image segmentation and reflection analysis through color", Proc. IUW'88, vol. II, Cambridge, Massachusetts, USA, 6 – 8 April 1988, pp. 838–853.

[42] Klinker G.J., Shafer S.A., and Kanade T. (1990) "A physical approach to color image understanding", *Int. Journal of Computer Vision*, vol. 4, no. 1, pp. 7–38.

[43] Lanser S. and Eckstein W. (1992) "A modification of Deriche's approach to edge detection", Proc. 11th Int. Conf. on Pattern Recognition 1992, Den Hague, Netherlands, 30 Aug.– 3 Sept. 1992, vol. III, Conf. C: Image, Speech, and Signal Analysis, pp. 633–637.

[44] Lanser S. (1993) "Kantenorientierte Farbsegmentation im CIE-Lab Raum", Proc. 15. DAGM-Symposium Mustererkennung 1993, eds. S.J. Pöppl and H. Handels, Lübeck, Germany, 27 – 29 September 1993, pp. 639–646, (in German).

[45] Lauterbach B. and Anheier W. (1993) "Segmentierung farbiger kartographischer Vorlagen in empfindungsgemäßen Farbräumen", Proc. 15. DAGM-Symposium Mustererkennung 1993, eds. S.J. Pöppl and H. Handels, Lübeck, Germany, 27 – 29 Sept. 1993, pp. 733–740, (in German).

[46] Lim Y.W. and Lee S.U. (1990) "On the color image segmenatation algorithm based on the thresholding and the fuzzy c-means techniques", *Pattern Recognition*, vol. 23, no. 9, pp. 935–952.

[47] Lin X. and Chen S. (1991) "Color image segmentation using modified HSI system for road following", Proc. IEEE Conf. on Robotics and Automation, Sacramento, California, April 1991, pp. 1998–2003.

[48] Maloney L.T. and Wandell B.A. (1986) "Color constancy: a method for recovering surface spectral reflectance", *Journal of the Optical Society of America A*, vol. 3, no. 1, pp. 29–33.

[49] Meyer F. (1992) "Colour image segmentation", Proc. IEE Int. Conf. on Image Processing and its Applications, Maastricht, Netherlands, 7 – 9 April 1992, pp. 303–306.

[50] Nevatia R. (1977) "A color edge detector and its use in scene segmentation", *IEEE Trans. on Systems, Man and Cybernetics*, vol. 7, no. 11, pp. 820–826.

[51] Nevatia R. (1982) *Machine perception*, Prentice-Hall, Englewood Cliffs.

[52] Ohlander R., Price K., and Reddy D.R. (1978) "Picture segmentation using a recursive region splitting method", *Computer Graphics and Image Processing*, vol. 8, pp. 313–333.

[53] Ohta Y.I., Kanade T., and Sakai T. (1980) "Color information for region segmentation", *Computer Graphics and Image Processing*, vol. 13, pp. 222–241.

[54] Pal N.P. and Pal S.K. (1993) "A review om image segmentation techniques", *Pattern Recognition*, vol. 26, no. 9, pp. 1277–1294.

[55] Panjwani D.K. and Healey G. (1993) "Unsupervised segmentation of textured colour images using Markov random field models", Proc. CVPR'93, New York, USA, 15–17 June 1993, pp. 776–777.

[56] Perez F. and Koch C. (1994) "Toward color image segmentation in analog VLSI: algorithm and hardware", *International Journal of Computer Vision*, vol. 12, no. 1, pp. 17–42.

[57] Phong B.T. (1975) "Illumination for computer generated pictures", *Communications of the ACM*, vol. 18, no. 6, pp. 311–317.

[58] Poggio T., Torre V. and Koch C. (1985) "Computational vision and regularization theory", *Nature*, 317(6035), pp. 314–319.

[59] Priese L. and Rehrmann V. (1992) "A fast hybrid color segmentation method", Fachberichte Informatik 12/92, Universität Koblenz-Landau, Germany.

[60] Priese L. and Rehrmann V. (1993) "On hierarchical color segmentation and applications", Proc. CVPR'93, New York, USA, 15 – 17 June 1993, pp. 633–634.

[61] Priese L. and Rehrmann V. (1993) "A fast hybrid color segmentation method", Proc. 15. DAGM-Symposium Mustererkennung 1993, eds. S.J. Pöppl and H. Handels, Lübeck, Germany, 27 – 29 September 1993, pp. 297– 304.

[62] Rubin J.M. and Richards W.A. (1982) "Color vision and image intensities: when are changes material ?", *Biological Cybernetics*, vol. 45, pp. 215–226.

[63] Rubin J.M. and Richards W.A. (1984) "Color vision and image intensities: representing material changes", AI Memo 764, MIT Artificial Intelligence Laboratory, Cambridge, Massachusetts, USA.

[64] Schettini R. (1993) "A segmentation algorithm for colour images" *Pattern Recognition Letters*, vol. 14, pp. 499–506.

[65] Schlüns K. and Wittig O. (1993) "Photometric stereo for non-Lambertian surfaces using color information", Proc. 7th Int. Conf. on Image Analysis and Processing 7 ICIAP, ed. S. Impedovo, Capitolo, Monopoly, Italy, 20 – 22 Sept. 1993, pp. 505–512.

[66] Shafer S.A. (1985) "Using color to separate reflection components", *COLOR research and application*, vol. 10, no. 4, pp. 210–218.

[67] Skarbek W. (1993) *Methods for digital image representation*, Akademicka Oficyna Wydawnicza PLJ, Warszawa (in Polish).

[68] von Stein H.-D. and Reimers W. (1983) "Segmentation of color pictures with the aid of color information and spatial neighborhoods", ed. H.W. Schüssler (1983) "Signal processing II: theories and applications", Elsevier Science Publ., North-Holland, pp. 271–273.

[69] Taylor R.I. and Lewis P.H. (1992) "Colour image segmentation using boundary relaxation", Proc. 11th IAPR Int. Conf. on Pattern Recognition, Den Hague, Netherlands, 30 Aug.–3 Sept. 1992, vol.III, conf. C: Image, Speech, and Signal Analysis, pp. 721–724.

[70] Tominaga S. (1986) "Color image segmentation using three perceptual attributes", Proc. CVPR'86, Miami Beach, Florida, USA, 22 – 26 June 1986, pp. 628–630.

[71] Tominaga S. (1990) "A colour classification method for color images using a uniform color space", Proc. 10th. Int. Conf. on Pattern Recognition, vol. I, Atlantic City, New Jersey, 16 – 21 June 1990, pp. 803–807.

[72] Tominaga S. and Wandell B.A. (1990) "Component estimation of surface spectral reflectance", *Journal of the Optical Society of America A*, vol. 7, no. 2, pp. 312–317.

[73] Trivedi M. and Bezdek J.C. (1986) "Low-level segmentation of aerial images with fuzzy clustering", *IEEE Trans. on Systems, Man, and Cybernetics*, vol. 16, no. 4, pp. 589–598.

[74] Tseng D.-C. and Chang C.-H. (1992) "Color segmentation using perceptual attributes", Proc. 11th Int. Conf. on Pattern Recognition 1992, Den Hague,

Netherlands, 30 Aug. – 3 Sept. 1992, vol. III, Conf C: Image, Speech, and Signal Analysis, pp. 228–231.

[75] Umbaugh S.E., Moss R.H., Stoecker W.V., and Hance G.A. (1993) "Automatic colour segmentation algorithms with application to skin tumor feature identification", *IEEE Engineering in Medicine and Biology*, vol. 12, no. 3, pp. 75–82.

[76] Vlachos T. and Constantinides A.G. (1992) "A graph-theoretic approach to colour image segmentation and contour classification", Proc. IEE Int. Conf. on Image Processing and its Applications, Maastricht, Netherlands, 7 – 9 April 1992, pp. 298–302.

[77] Westman T., Harwood D., Laitinen T., and Pietikäinen M. (1990) "Color segmentation by hierarchical connected components analysis with image enhancement by symmetric neighborhood filters", Proc. 10th Int. Conf. on Pattern Recognition 1990, Atlantic City, New Jersey, USA, 16 – 21 June 1990, pp. 796–802.

[78] Wright W.A. (1989) "A Markov random field approach to data fusion and colour segmentation", *Image and Vision Computing*, vol. 7. no. 2, pp. 144–150.

[79] Wyszecki G. and Stiles W. S. (1982) *Color science*, Wiley, New York.

[80] Yachida M. and Tsuji S. (1971) "Application of colour information to visual perception", *Pattern Recognition*, vol. 3, pp. 307–323.

[81] Young R. (1986) "Simulation of human retinal function with the Gaussian derivative model", Proc. Int. Conf. on Computer Vision and Pattern Recognition, Miami Beach, Florida, USA, 22 – 26 June 1986, pp. 564–569.