

## TECHNICAL NOTE

## A METHOD FOR SEGMENTING FOOD COLOUR IMAGES

Giorgio Peri, Roberto Romaniello

## 1. Introduction

Segmentation of food image, i.e. partitioning the digital image into disjoint and homogeneous regions, is an essential step in computer vision and automatic pattern recognition processes based on image analysis of foods. Generally, in food images segmentation, there are only two regions: the foreground (food image itself) and the background [4]. Despite this simplification, the segmentation process of a colour image is still an inherently complex problem because it involves the determination of an optimal decision surface in the 3D colour space so that an image pixel is classified as a foreground point if its colour point is in one side of the decision surface, otherwise it is considered as a background point [1].

In this note is proposed a method for segmentation of food colour images in regions that are the object of interest and the background. The basic idea is to reduce the calculation complexity inherent to the determination of an optimal decision surface in the 3D colour space with the assumption that this surface is a plane orthogonal to one of the reference axis of the colour space. The proposed method has three steps: i) determination of the optimal decision plane for the segmentation of an image; ii) coarse segmentation of the image; iii) morphological operations in order to correct possible errors in the segmented image.

## 2. Scheme of the method

In the standard colour space (R,G,B) each pixel of an image is represented by a colour point  $\mathbf{c} = (r, g, b)$ , where  $r$ ,  $g$  and  $b$  define the intensity of the red, green, and blue components of the pixel colour. It follows that an image with  $N$  pixels is represented with a set  $S$  of  $N$  colour points in the RGB colour space.

In general, these colour points are not all distinguished among them, so to a colour image corresponds a discrete density function  $P(\mathbf{c})$  indicating the normalized frequency (or weight) of each colour point in the RGB colour space.

Since the red, green, and blue intensities of the pixels are quantized usually using 256 different levels, in the RGB colour space there are 768 planes each one orthogonal to the reference axis.

## 2.1 Determination of the optimal decision plane

If a plane  $\Gamma$  divides the  $S$  set of colour points in two subsets of colour points  $S_1$  and  $S_2$  the sum of the weighted variances of the colour points  $\mathbf{c} \in S_1$  and  $\mathbf{c} \in S_2$ , called partition error  $E(\Gamma)$  [6], is given by:

$$E(\Gamma) = \sum_{\mathbf{c} \in S_1} P(\mathbf{c})(\mathbf{c} - \mathbf{m}_1(\Gamma))^2 + \sum_{\mathbf{c} \in S_2} P(\mathbf{c})(\mathbf{c} - \mathbf{m}_2(\Gamma))^2 \quad (1)$$

with

$$\mathbf{m}_1(\Gamma) = \frac{\sum_{\mathbf{c} \in S_1} \mathbf{c}P(\mathbf{c})}{\sum_{\mathbf{c} \in S_1} P(\mathbf{c})} \quad \text{e} \quad \mathbf{m}_2(\Gamma) = \frac{\sum_{\mathbf{c} \in S_2} \mathbf{c}P(\mathbf{c})}{\sum_{\mathbf{c} \in S_2} P(\mathbf{c})} \quad (2)$$

where  $\mathbf{m}_1(\Gamma)$  and  $\mathbf{m}_2(\Gamma)$  are the centroid (or gravity centres) of the colour points respectively  $\mathbf{c} \in S_1$  e  $\mathbf{c} \in S_2$ .

The optimal decision plane for the segmentation of an image is individuated by the partition error minimization criteria [6]. In fact, with the computation of the partition error associated to each of the 768 planes orthogonal to reference axis of the colour space, the optimal decision plane  $\Gamma_{\text{optimum}}$  is given by:

$$\Gamma_{\text{optimum}} = \arg \min_{\Gamma} E(\Gamma) \quad (3)$$

A very efficient method for the error minimization process consists in rearrange the equation (1) in the follow equivalent form [7]:

$$E(\Gamma) = \sum_{\mathbf{c} \in S_1 \cup S_2} \mathbf{c}^2 P(\mathbf{c}) \frac{\left( \sum_{\mathbf{c} \in S_1} \mathbf{c} P(\mathbf{c}) \right)^2}{\sum_{\mathbf{c} \in S_1} P(\mathbf{c})} + \frac{\left( \sum_{\mathbf{c} \in S_2} \mathbf{c} P(\mathbf{c}) \right)^2}{\sum_{\mathbf{c} \in S_2} P(\mathbf{c})} \quad (4)$$

Memory presented the 11.06.2006; accepted the 16.05.2008

GIORGIO PERI, full professor; ROBERTO ROMANIELLO, Phd student. Dipartimento di Scienze delle Produzioni, dell'Ingegneria, della Meccanica e dell'Economia applicate ai Sistemi Agro-zootecnici, Università degli Studi di Foggia.

Since the first term to the second member of the equation (4) is a constant, the minimization of the partition error implies the maximization of the summation of the second and the third term at the second member of the equation (4). Thus, given the discrete density function  $P(c)$  and calculated for each decision plane orthogonal to reference axis the quantity:

$$C_1(\Gamma) = \sum_{c \in S_1} P(c) \quad (5a)$$

$$C_2(\Gamma) = \sum_{c \in S_2} P(c) \quad (5b)$$

$$C_3(\Gamma) = \left( \sum_{c \in S_1} cP(c) \right)^2 \quad (5c)$$

$$C_4(\Gamma) = \left( \sum_{c \in S_2} cP(c) \right)^2 \quad (5d)$$

the optimal decision plane  $\Gamma_{optimum}$  is also given by:

$$\Gamma_{optimum} = \arg \max_{\Gamma} \left( \frac{C_3(\Gamma)}{C_1(\Gamma)} + \frac{C_4(\Gamma)}{C_2(\Gamma)} \right) \quad (6)$$

that for each possible decision plane requires the execution of only two division operations and a summation operation of pre-calculated quantity using 5a, 5b, 5c, and 5d.

## 2.2 Coarse segmentation

The coarse segmentation of the image is obtained through the partition of the its colour points set  $S$  in two subsets  $S_1$  and  $S_2$  using the optimal decision plane identified by (6) and by assignment of a binary symbol (for example, 0) to the pixels represented from the colour points  $c \in S_1$  and of the complementary binary symbol (for example, 1) to the pixels represented from the colour points  $c \in S_2$ . This pixels classification in the image plane, produces a binary image in which there are present two classes of pixels, each of which is constituted of pixels with the same binary symbol.

## 2.3 Morphological operations

In the binary image may be present isolated pixels and/or in small groups wrongly classified; it occurs if the optimal decision surface for the image segmentation is not a plane, or if it is a plane not orthogonal to a reference axis of the RGB colour space. Also pixel misclassifications occur if the two colour points agglomerations representing the image regions in the RGB colour space are partly overlapped, and therefore not fully separable. Given that the regions in the segmented image must be two and each of them must be made up of connected pixels [3], the correction of the binary image produced by the coarse segmentation process is done with two morphological operations known as area-opening and flood-filling [5]. The first of these morphological operations assigns the bi-

nary symbol 0 to the pixels with binary symbol 1 whose area, in terms of number of pixels, is less than a predetermined threshold and assures, therefore, the connection of the pixels in the class of those with assigned the binary symbol 0. The second of these morphological operations assigns the symbol binary 1 to pixels with binary symbol 0 that are present in the class of pixels with symbol binary 1 and assures, therefore, the connection of the pixels in the class of those with the assigned binary symbol 1. The threshold value for the area-opening operation was set equal to  $N/100$ , where  $N$  is the pixels number of the image.

## 3. MATLAB program

The method was coded in MATLAB 7.1 (The Mathworks Inc., U.S.A.) and was implemented in a personal computer with Pentium M processor (1,86 GHz), 2 GB RAM and Windows XP Professional as operating system.

## 4. Materials and methods

The method performances were evaluated with the segmentation of 40 food colour images belonging to types with very different morphological and chromatic characteristics; the types of foodstuff considered were: meat, bakery products, fruits and tubers. These images (10 for each type of product) were acquired in a custom acquisition room with lighting system constituted of 4 fluorescent lamps (Osram TLD 65 – 15 W) and colour temperature of 6500 K through a camera (PowerShot A520, Canon Inc., Japan) with spatial resolution of 2272 x 1704 pixels and colour depth of 8 bits per each of the three components of pixels colour. After the capturing step the images were saved in the TIFF format (Tagged Image File Format) without compression, and thereafter segmented unmodified, and after the reduction of spatial resolution to 1600 x 1200 pixels, 1280 x 1024 pixels and 1024 x 768 pixels.

Three performance indexes of the method were considered: object pixels fraction correctly classified in the segmented image, background pixels fraction correctly classified in the segmented image, and image processing time. With reference to the first two performance indices in a segmented image the fraction of the object pixels properly classified  $P_o$  and the background pixels fraction correctly classified  $P_s$  are given by:

$$P_o = \frac{n_1}{n_2} \quad e \quad P_s = \frac{n_3}{n_4} \quad (7)$$

where

- $n_1$  is the pixels number of the objects correctly classified;
- $n_2$  is the pixels actual number of the object;
- $n_3$  is the pixels number of the background correctly classified;

- $n_4$  is the pixels actual number of the background.

The pixels actual number of the object and the pixels actual number of the background in the images were derived from the corresponding binary images representing the optimal visual result of the segmentation. For each one of the 40 images considered the mentioned binary image was obtained using a graphic elaboration software (Paint Shop Pro 7, Jasc Software, USA).

With reference to the third performance index, the image processing time was calculated by the sum of the coarse segmentation time and the binary image correction time.

## 5. Results

Table 1 lists the average values and standard deviations of the pixels fraction of the object correctly classified in the segmented images of the 4 types of foodstuff taken into account. The maximum and the minimum of these average values are respectively  $0.9901 \pm 0.0042$  and  $0.9882 \pm 0.0061$ . Therefore, the corresponding average classification errors of object pixels are respectively about 0.99% and 1.18%. Table 2 lists the average values and standard deviations of the pixels fraction of the background correctly classified in the segmented images of the 4 types of foodstuff taken into account. The maximum and the minimum of these average values are respectively  $0.9998 \pm 0.0001$  and  $0.9931 \pm 0.0002$ . Therefore, the corresponding average classification errors of the background pixels are respectively about 0.02% and 0.69%.

Type of foodstuff	Average value and standard deviation of the pixel fraction of the object correctly classified
Meat	$0.9882 \pm 0.0061$
Bakery products	$0.9891 \pm 0.0044$
Fruit	$0.9901 \pm 0.0042$
Tubers	$0.9898 \pm 0.0077$

TABLE 1 - Mean values and standard deviations of the foreground pixels fraction correctly classified in the segmented images of the 4 foodstuff types.

Type of foodstuff	Average value and standard deviation of the pixel fraction of the background correctly classified
Meat	$0.9995 \pm 0.0003$
Bakery products	$0.9997 \pm 0.0002$
Fruits	$0.9998 \pm 0.0001$
Tubers	$0.9931 \pm 0.0002$

TABLE 2 - Mean values and standard deviations of the background pixels fraction correctly classified in the segmented images of the 4 foodstuff types.

Type of foodstuff	Average value and standard deviation of the image processing time (s)
Meat	$1.72 \pm 0.03$
Bakery products	$1.75 \pm 0.03$
Fruits	$1.74 \pm 0.02$
Tubers	$1.73 \pm 0.02$

TABLE 3 - Mean values and standard deviations of the image processing time of the 4 foodstuff types.

Spatial resolution (pixel x pixel)	Average value and standard deviation of the image processing time (s)
2272 x 1702	$1.74 \pm 0.07$
1600 x 1200	$0.87 \pm 0.03$
1280 x 1024	$0.61 \pm 0.03$
1024 x 768	$0.36 \pm 0.01$

TABLE 4 - Mean values and standard deviations of the image processing time for segmenting all images with spatial resolution of 2272x1704 pixels and with lower spatial resolutions.

Table 3 lists the average values and standard deviations of the image processing time of the 4 types of foodstuff taken into account. The maximum and the minimum of these averages are respectively  $1.75 \pm 0.03$  s and  $1.72 \pm 0.03$  s with a maximum deviation of 0.09 s. These results indicate that the image processing time does not depend, in fact, from the morphological and the colour characteristics of the object of interest. Table 4 lists the average values and standard deviations of processing time of all images acquired with spatial resolution 2272 x 1704 and after the reduction of spatial resolution to lower values (1600 x 1200, 1280 x 1024 and 1024 x 768). An examination of Table 4 highlights the dependence of the average image processing time from spatial resolution, or from number of colour points that represent the images in the RGB colour space. This number of colour points is 3866944, 1920000, 1310720 and 786432 for spatial resolutions 2272 x 1704, 1600 x 1200, 1024 and 1280 x 1024 x 768 respectively.

As shown in Figure 5 the mentioned dependence of the average image processing time by the number of colour points that represent them in RGB colour space is linear with determination coefficient equal to 0.9951.

## 6. Conclusions

In this work is proposed a method for segmentation of food colour images in regions that are the object of interest and the background. This method is easy to implement using MATLAB. The results obtained with

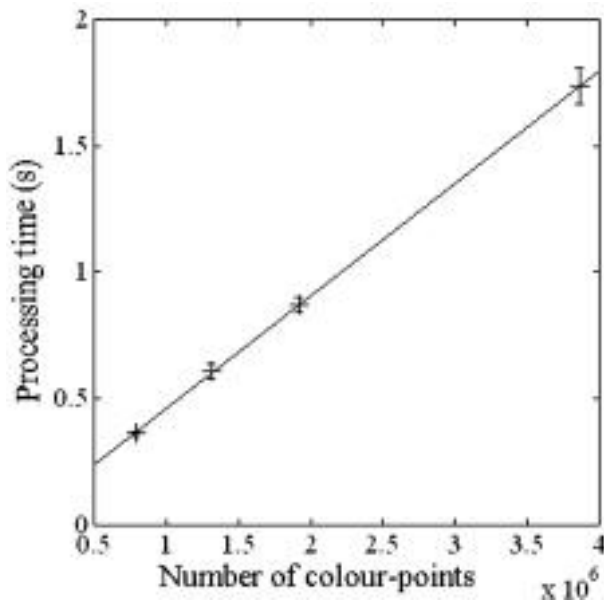


Fig. 5 - Relationship between the image processing time and the number of the colour points ( $N$ ) representing the image in the RGB colour space.

the segmentation of 40 food images with very different chromatic and morphological characteristics indicates that the proposed method gives accurate results, with classification errors of the object pixels and of the background pixels equal to about 1% and about 0.4% respectively. In addition, the proposed method is efficient because the images processing times are on average about 0.4 s 1.7 s respectively in the case of images with medium spatial resolution (1024 x 768 pixels) and in the case of images with relatively high spatial resolution (2272 x 1704 pixels).

## 7. References

[1] CELENK M.U., HAAG M.U., *Optimal thresholding for*

*color images*, SPIE Conference on Nonlinear Image Processing IX (1998), 250-259.

[2] GONZALEZ R., WINTZ O., *Digital image processing*, 2<sup>a</sup> ed. (2002), Prentice-Hall International, Upper Saddle River (USA).

[3] JAIN A.K., *Fundamentals of digital image processing*, 1<sup>a</sup> ed. (1989), Prentice-Hall International, London (UK).

[4] MERY D., PEDRESCHI F., *Segmentation of colour food images using a robust algorithm*, Journal of Food Engineering (2005) 66, 353-360.

[5] SOILLE P., *Morphological image analysis*, 2<sup>a</sup> ed. (2003), Springer-Verlag, Heidelberg (DE).

[6] WAN S.J., WONG S.K.M., PRUSIKIEWICZ P., *An algorithm for multidimensional data clustering*, ACM Transactions on Mathematical Software (1988) 14, 153-162.

[7] WU X., *Efficient statistical computations for optimal color quantization*, in Graphics Gem II, 1<sup>a</sup> ed. (1991), Academic Press, San Diego (USA).

## SUMMARY

In this note, a method for segmenting foods from their backgrounds in colour images is presented. The proposed method has three steps: i) determination of the optimal decision plane for the segmentation of an image; ii) coarse segmentation of the image; iii) morphological operations in order to correct possible errors in the segmented image.

The method was implemented in MATLAB and tested on 40 colour images of foodstuff with very different morphological and chromatic characteristics, including meat, baked products, fruit and tubers.

The experimental results are presented and the performance of method in the segmentation process is assessed. The method has shown to be both effective and efficient also for colour images with high spatial resolution.

**Keywords:** colour images, food, image processing, segmentation.