Fitting Contours of Marbling in Iberian Ham Images

A. Plaza[1], E. Cernadas[1,3], M.L. Durán[1] and M.J. Petrón[2]

[1] *Departamento de Informática, Escuela Politécnica de Cáceres*
[2] *Departamento de Zootecnia, Facultad de Veterinaria*
Universidad de Extremadura, Avda. de la Universidad s/n, 10.071 Cáceres, SPAIN
[3]*Departamento de Lenguajes y Sistemas Informáticos, Escuela Superior de Ingeniería Informática*
Campus Universitario As Lagoas, Ourense, SPAIN
email: aplaza@unex.es

Abstract: *Cured Iberian ham is one of the most valuable meat products of Spain and Portugal, with a first-rate consumer acceptance. Its sensorial characteristics are basically due to high intramuscular fat content and marbling, i.e. number, position, shape and size of intramuscular fat streaks. Our current attempts are to investigate the relationship between spatio-geometrical characteristics of fat streaks and Iberian ham sensorial quality. To achieve this goal, not only the detection of streaks is important, but also the identification of their contours with high accuracy. In previous work we have developed a method for this purpose, based on two steps: obtaining an initial estimate of the position of the streaks and refining the estimate by region growing techniques. In this paper we concentrate on the latter issue, evaluating the performance of several region growing techniques in the task of contour-fitting of marbling in Iberian ham images. A statistical study of pixel and region classification is provided and discussed.*

1. INTRODUCTION

Iberian ham is one of the most valuable dry-cured meat products of Spain and Portugal, with a first-rate consumer acceptance. It is produced from uncooked hams of Iberian pigs, following a prolonged traditional method that requires between 1 and 2 years ripening. Its sensorial characteristics are basically due to high intramuscular fat content and marbling. Currently, chemical processing is the only proved way to determine fat level of cured Iberian hams, but this technique is tedious, destroying, and does not offer any information about fat distribution [1].

Some researchers have studied a set of dry-cured Iberian hams from pigs with different feeding system [2]. Their chemical analysis revealed similar fat levels, but hams from pigs fed on acorns and pasture obtained better qualifications in a sensorial analysis of the samples carried out by trained human testers. So, ham marbling, i.e. number, position, shape and size of intramuscular fat streaks, must influence its high sensorial characteristics. Then, the design of a non-destructive methodology to classify Iberian ham from the viewpoint of sensorial quality, independently of the subjective and variable criteria of testers, would be of great interest to Iberian ham industries.

Intramuscular fat streaks can be seen as curvilinear patterns (see figure 1, left). Our current attempts are to investigate the relationship between objective spatio-geometrical characteristics of the streaks and Iberian ham sensorial quality. To achieve this goal, we should recognize the marbling in Iberian ham images. In this task, not only the detection of fat streaks is important, but also the identification of their contours with high precision, in order to calculate quality-indicative parameters as the relative size of the largest streaks in the slice, streak shape descriptors, etc.

In the marbling recognition process, Iberian ham images are segmented into regions of interest and background. Despite the large body of literature, segmentation is a difficult task that is severely affected by image acquisition details (such as noise and illumination conditions) and high-level information (such as expectations of the contents of a scene) [3]. In previous work, we have developed and evaluated a segmentation method to detect fat streaks [4], [5]. This method consists of two steps: i) obtaining an initial estimate of the position of the streaks and ii) refining the estimate by region growing techniques. In this paper, we are only interested in testing the accuracy of the refinement stage, so, we compare the performance of several contour-fitting algorithms in our particular application. This is done by using ground truth images to generate a set of seed points which represent a true estimate of the position of the streaks.

The paper is organized as follows: in the following section we describe different approaches to region refinement. Section 3 describes our experiment and discusses the results of applying the proposed region growing techniques to our database of Iberian ham images. Section 4 provides the conclusions at which we have arrived.

2. METHODS

Compared to the substantial volume of research devoted to image segmentation, region refinement has received much less attention. The existing techniques can be classified from different points of view:

a) Global region-growing techniques, if the growth criteria are the same for all the regions in the image, or local region-growing techniques, if these criteria are region-dependent.

b) Parameterized or parameter-free techniques, depending on the use or not of certain external parameters, determined from heuristic knowledge, to control the growing process.

c) Regarding the intrinsic characteristics of the region-growing techniques, they can be region-oriented, boundary-oriented or hybrid.

In this section, we describe the proposed methods taking in mind the previously addressed classification.

2.1. Region-Oriented Methods

Region-oriented methods rely on the postulate that neighbouring pixels within one region have similar intensities. Let I(P) be the intensity at pixel P, Ev(P) the expected intensity at pixel P, estimated from the seed region characteristics, and T a fixed threshold. Standard region-oriented approaches [6], [7] iteratively merge seed regions with adjacent pixels if either the intensity of these pixels or their dissimilarity with respect to the expected value is below a fixed, user-defined threshold, i.e. $I(P) < T$ or $|E_v(P) - I(P)| < T$.

The major limitation of the previous approach is that it incorporates a parameter. As a result, these methods are highly sensitive to changes in this value, which must be carefully tuned to satisfy the requirements of the particular application. We propose a parameter-free region-based method where Ev(P) is the mean of the seed region and T is automatically obtained from statistical characteristics of the original image. In a previous work [5], some methods [8] to determine T were evaluated (Tsai method, based on the moment-preserving principle, Kapur method, based on entropy considerations of the gray-level histogram, and Otsu method). The best results for the detection problem were found using Otsu method, which works as follows: let σ_W^2, σ_B^2 and σ_T^2 be the within-class, between-class and total variance, respectively. An optimal threshold T is determined by maximizing one of the following criterion functions with respect to T [9]:

$$\lambda = \frac{\sigma_B^2}{\sigma_W^2} \qquad \eta = \frac{\sigma_B^2}{\sigma_T^2} \qquad \kappa = \frac{\sigma_T^2}{\sigma_W^2}$$

Rosin [10] also proposes a region-oriented method which does not require any parameters. In this method, the labelling of an adjacent pixel is made by calculating the region/pixel and the background/pixel dissimilarity. Let $E_{V\_Region}(P)$ and $E_{V\_Background}(P)$ be the expected values at a pixel belonging to a region of interest and the background, respectively. If $|E_{V\_Region}(P) - I(P)| < |E_{V\_Background}(P) - I(P)|$, the pixel is incorporated to the region. The expected values depend on the region model used, which can be data-driven or model-driven. In the data-driven approach, the expected values are estimated from the seed regions (using the mean operation, for instance) and change adaptatively as the seeds grow. On the other hand, the model-driven approach assumes the availability of labeled training data (ground truth) from which the expected values are estimated. This can be done using a leaving-one-image-out approach, i.e. calculating the mean intensities of the regions of interest and the background in all the available images except the one under consideration and using these intensities as the expected values in the current image.

2.2. Boundary-Oriented Methods

Boundary-based approaches to region growing use the postulate that the pixel values change rapidly at the edges of regions [3]. The objective here is to obtain closed curves representing the boundaries between regions. Gradient operators like the Sobel or Roberts filter fail to produce closed contours, but the Laplacian of Gaussian operator is able to obtain all the zero-crossings of the original image. In our proposed method, the zero crossings are used as watersheds [11] for the expansion of the seeds.

2.3. Hybrid Methods

While both the region and boundary methods have their advantages and disadvantages, their problems are not necessarily identical. Gradient-based boundary finding is more affected by noise than region-based methods because the gradient is very noise-sensitive. Further, since conventional boundary finding relies on changes in the gray-level rather than its actual value, this method is less sensitive to changes in the gray-scale distribution. Many authors [12], [13], [14] have proposed hybrid techniques which combine boundary and region criteria. We propose a rule-based method where the decision to merge a seed region and an adjacent pixel depends on two parameters: the region/pixel dissimilarity and the gradient at the position of the pixel, which is calculated using the two Sobel operators Gx(P) and Gy(P):

$$\text{Merge adjacent pixel P if } \begin{cases} |E_v(P) - I(P)| < A \\ \text{and} \\ \sqrt{G_x(P)^2 + G_y(P)^2} < B \end{cases}$$

$$\text{In our particular application, we consider } \begin{cases} A = 20 \\ B = 250 \end{cases}$$

3. Results and Discussion

3.1. Description of the Experiment

Fifteen Iberian ham slices from the *biceps* muscle have been digitized with a general purpose scanner at spatial resolution of 100 pixels per inch and gray-level resolution of 8 bits depth. The area in pixels of the slices ranges from 131.109 to 203.779, i.e. 57 cm^2 to 72

cm^2. All visible intramuscular streaks have been annotated by an expert in food technology. The total number of streaks annotated per image ranges from 21 to 47 (up to 484 in the 15 images) and their areas range from 20 to 5.314 pixels, i.e. 1.3 mm^2 to 343 mm^2. There are 32.46 average annotated streaks per image; 14.13 with area smaller than 100 pixel units, 15.6 between 100 and 1000 pixels, and 2.73 larger than 1000 pixels. Figure 1 shows an original Iberian ham image and its correspondent set of annotated intramuscular fat streaks, which are represented as white regions against a black background.

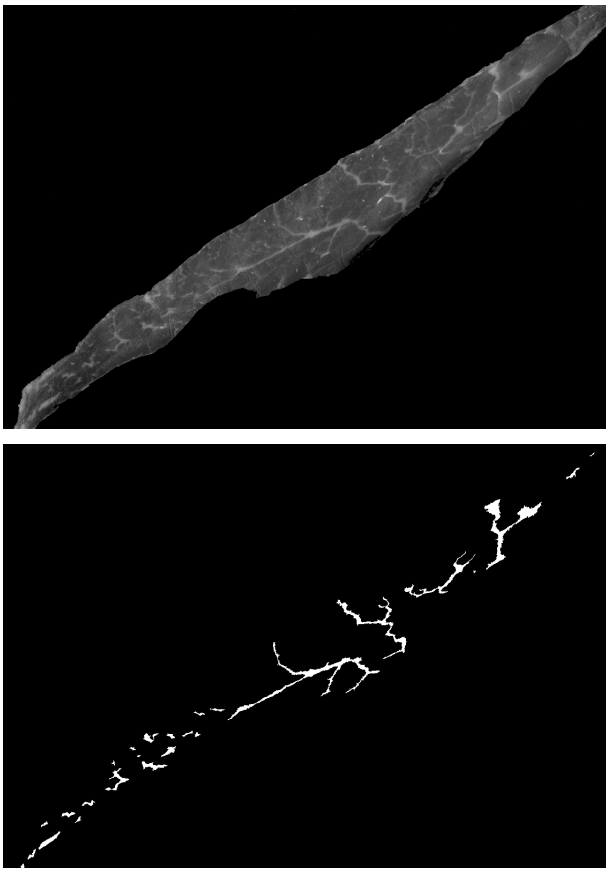

Figure 1. Original Iberian ham image (upper) and correspondent annotated image (lower).

To evaluate the performance of the addressed region growing techniques, we automatically choose a random point inside each annotated streak and start the growing in the original image from this seed point. The true initial estimate allows us to isolate the boundary-fitting capacity of the algorithms from the fact that the performance of these methods usually falls if the seeds are not accurately located. We have also considered the selection of three random points inside each annotated streak in our experiment in order to prove if the contour accuracy depends on the number of seeds included into each region.

3.2. Results and Comparison

The proposed region growing methods have been applied to our database of images. We assess their performance in terms of region classification accuracy by calculating the measures of Hoover et al. [15].

A correct detection instance is counted when a detected region is at least T% overlapped with an annotated region and vice-versa. Figures 2 and 3 show the scores of correct detections for the proposed region growing techniques, starting from one and three seed points, respectively. Values of the tolerance T ranging from 10 to 90 are considered to simulate the effect of demanding more accuracy in true boundaries detection. For comparative purposes, the results obtained are separated regarding the size of the fat streaks to be detected (three different intervals are considered: smaller, medium-sized and larger streaks).

After a detailed analysis of the results presented in figures 2 and 3, some remarks can be made:

1. The use of three seed points does not improve significantly the contour precision obtained. So, contour accuracy does not depend on the number of seeds into each region.
2. At very soft tolerance (T=10%) some regions are not detected. This fact indicates that a small percentage of seeds never grow.
3. The Laplacian method is clearly superior to the other ones when detecting small or medium-sized fat streaks with moderate accuracy; otherwise, the performance of this method falls dramatically. On the other hand, Otsu and hybrid methods produce results which are comparable, being the latter approach slightly better only when very high accuracy is demanded.

In general, neither the data-driven nor the model-driven methods improve the results obtained by the other tested methods (only with large streaks and very high contour precision, the model-driven approach performs better).

We began this experiment in the belief that robust region-growing techniques would lead us to better results, but this is not necessarily true in our specific application. Thus, taking in mind other aspects like prior knowledge needed, it is important to emphasize that, although Otsu, hybrid and model-driven methods have similar performance, Otsu approach does not incorporate a priori information about the application and/or heuristic parameters.

As mentioned before, region classification depends on overlapping criteria. Regarding that the percentage of missed area is given by (100-T)%, a region is classified as correct if it satisfies a vague criterion, which is stricter as T increases. For the purpose of determining the relative area of the largest fat streaks in each Iberian ham slice, the performance of the techniques does not only depend on the number of correctly detected regions, but also on the percentage of correctly labelled pixels.

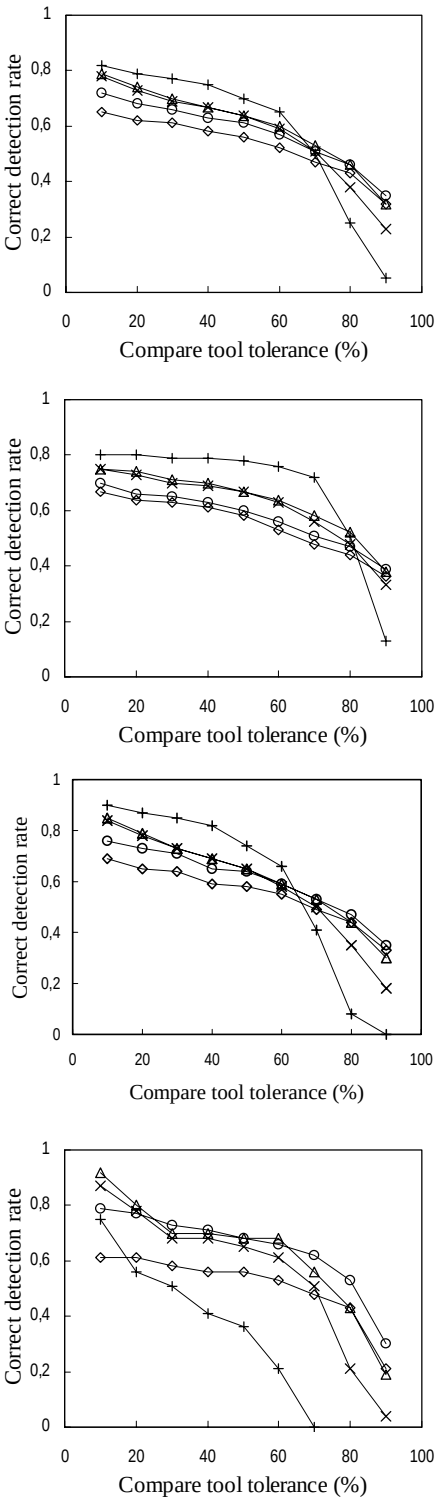

Figure 2. Correct detection rate for Otsu (×), data-driven (♦), model-driven (●), Laplacian (+) and hybrid (▲) region growing methods at different tolerances, considering all regions (upper), regions whose size is between 1 and 100 pixels (middle-upper), regions between 100 and 1000 pixels (middle-lower) and regions larger than 1000 pixels (lower). The methods start from one seed point.

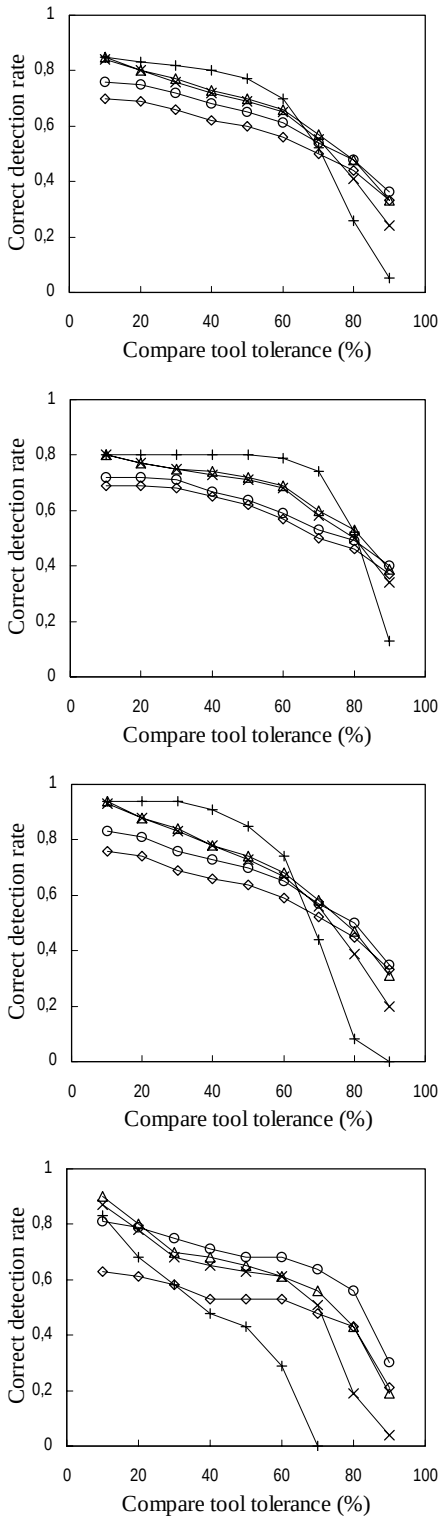

Figure 3. Same as figure 2 but the methods start from three seed points.

Then, apart from region classification, we assess the performance of the proposed methods in terms of pixel classification accuracy. Sensitivity (S) represents the rate of pixels annotated by experts in food technology which are correctly detected, and positive predictivity (PP) represents the rate of correctly detected pixels

among all pixels segmented into fat streaks (see appendix). Table 1 shows the results obtained for the proposed algorithms.

Region-growing method	*Sensitivity* (S)	*Predictivity* (PP)
Laplacian	*0.62*	*0.56*
Hybrid	*0.69*	*0.59*
Otsu	*0.73*	*0.59*
Data-driven	*0.91*	*0.32*
Model-driven	*0.74*	*0.57*

Table 1. Sensitivity and positive predictivity after applying the addressed region growing techniques to our dataset. The methods start from one seed point.

Since we have experimentally proved that the total annotated and detected areas are always comparable in the resulting images, S and 1-PP measure the percentage of correctly and incorrectly classified area at a pixel level, respectively. The best approach would be the one that satisfies two considerations:

1. The sensitivity must be as high as possible.
2. Both the percentage of misclassified area (1-S) and erroneously classified area (1-PP) must be as small as possible and comparable.

Based on this reasoning and taking into account the results addressed in table 1, data-driven method achieves the highest sensitivity at the expense of a high rate of erroneously classified area *(1-0.32=0.68)*, while the rest of the approaches provide, in general, better compromises, .i.e. quite high sensitivity and $(1-S) \cong (1-PP)$.

4. CONCLUSIONS

We have compared and discussed several region-growing techniques for the purpose of contour-fitting of marbling in Iberian ham images. We began this experiment in the belief that robust region-growing techniques would lead us to better results. The main conclusion of this work is that the previous remark may be true in general-purpose applications, but it is not necessarily true in our specific application. Boundary-based region growing reveals to be superior when small and medium-sized fat streaks are to be fitted and moderate contour accuracy is demanded; otherwise the performance of this method falls drastically in relation to other tested methods. This fact leads us to think that better solutions to the whole detection process could be achieved by combining methods, i.e. i) initial growing by boundary-based techniques and ii) refinement of the resulting regions by some of the other approaches. Nevertheless, further research is still needed in order to prove this remark.

ACKNOWLEDGEMENT

Funding from *Junta de Extremadura* (local government) under the IPR98A03P project and FIC98B015 PhD fellowship is gratefully acknowledged.

APPENDIX: STATISTICAL MEASURES

Let $i = 1..M$ be the number of images in the database, Np_{A_i} the number of pixels corresponding to regions in A_i (the annotated image), Np_{R_i} the number of pixels corresponding to regions in R_i (the resulting image after applying a region growing technique) and Np_{O_i} the number of pixels of $A_i \cap R_i$. The sensitivity is defined:

$$S = \frac{1}{M}\sum_{i=1}^{M}\frac{Np_{O_i}}{Np_{A_i}}$$

and the positive predictivity is defined:

$$PP = \frac{1}{M}\sum_{i=1}^{M}\frac{Np_{O_i}}{Np_{R_i}}$$

REFERENCES

[1] Antequera, T., López-Bote, C.J., Córdoba, J.J., García, C., Asensio, M.A., Ventanas, J., Díaz, Y., “Lipid oxidative changes in the processing of Iberian pig hams”, Food Chemistry, vol. 45, no. 105, pp. 105-110, 1992.

[2] Ruiz, J., Ventanas, J., Cava, R., Timón, M.L., García, C., “Sensory characteristics of Iberian Ham: Influence of Processing Time and Slice Location”. Food Research International, vol. 31, no. 1, pp. 53-58, 1998.

[3] Haralick, R.M. and Shapiro, L.G., Computer and Robot Vision, Addison-Wesley, 1993.

[4] Plaza, A., Cernadas, E., Durán, M.L., Sánchez, Antequera, T., “Directional Morphological Operators to Detect Fat Streaks in Iberian Ham”, VIII National Symposium on Pattern Recognition and Image Analysis, pp. 47-48, Bilbao, Spain, 1999.

[5] Plaza, A., Cernadas, E., Durán, M.L., Sánchez, J.M., Petrón, M.J., “Recognizing Marbling in Iberian Ham”, Third International Conference on Computer Vision, Pattern Recognition and Image Processing, Atlantic City, USA, Feb. 2000 (accepted).

[6] Adams, R. and Bischof, L., “Seeded Region Growing”. IEEE Trans. PAMI, vol. 16, no. 6, pp. 641-647, 1994.

[7] Chang, Y. and Li, X., “Adaptative Image Region-Growing”. IEEE Trans. PAMI, vol. 3, no. 6, pp. 868-872, 1994.

[8] Pal, N.R. and Pal, S.K., “A Review on Image Segmentation Techniques”, Pattern Recognition, vol. 26, no. 9, pp. 1277-1294, 1993.

[9] Otsu, N., “A Threshold Selection Method from Gray-Level Histograms”, IEEE Trans. Systems, Man, and Cybernetics, vol. 9, no. 1, pp. ,1979.

[10] Rosin, P.L., “Refining Region Estimates”, International Journal of Pattern Recognition and Artificial Intelligence, vol. 12, no. 6, pp. 841-866, 1998.

[11] Vincent, L. and Soille, P., “Watersheds in Digital Spaces: An Efficient Algorithm Based on Inmersion Simulations”, IEEE Trans. PAMI, vol. 13, no. 6, pp. 583-598, 1991.

[12] Pavlidis, T. and Liow, Y., “Integrating Region Growing and Edge Detection”. IEEE Trans. PAMI, vol. 12, no. 3, pp. 225-233, 1990.

[13] Gambotto, J., “A New Approach to Combining Region Growing and Edge Detection”. Pattern Recognition Letters, 14, pp. 869-875, 1993.

[14] Chakraborty, A., Staib, L.H., Duncan, J.S., “Deformable Boundary Finding in Medical Images by Integrating Gradient and Region Information”, IEEE Trans. Medical Imaging, vol. 15, no. 6, pp. 859-870, 1996.

[15] Hoover, A., Jean-Baptiste, G., Jiang, X., Flynn, P.J., Bunke, H., Goldgof, D.B., Bowyer, K., Eggert, D.W., Fitzgibbon, A., Fisher, R.B., “An Experimental Comparison of Range Image Segmentation Algorithms”, IEEE Trans. PAMI, vol. 18, no. 7, pp. 673-689,1996.