Recognizing Marbling in Iberian Ham

A. Plaza, E. Cernadas, M.L. Durán, J.M. Sánchez, M.J. Petrón*
Grupo de Imagen Digital, Departamento de Informática, Escuela Politécnica
**Departamento de Zootecnia, Facultad de Veterinaria*
Universidad de Extremadura, Avda. de la Universidad s/n, 10071, Cáceres (Spain)
aplaza@unex.es

Abstract

The Iberian pig is an autochthonous animal bred from the south-western area of Spain. Dry-cured Iberian ham presents a special flavor, which makes it a very worthy product among consumers. Its special characteristics are mainly due to marbling and fat content.

Currently, chemical procedure is the only proved way to evaluate fat content in Iberian ham. Nevertheless, food technology experts believe that marbling, i.e., number and size of intramuscular fat streaks, also influences its sensorial quality. In this study we focus on the detection and contour identification of fat streaks in digital Iberian ham images. A statistical study of the proposed methods is provided.

1. Introduction

The Iberian pig is an autochthonous animal bred from the south-western area of Spain. Its meat is mainly used to produce dry-cured products, in particular cured hams. These hams are very appreciated by consumers and their special characteristics are basically due to high intramuscular fat content and marbling.

Currently, chemical procedure is the only proved way to determine fat level of cured Iberian hams. This method is tedious, destroying and does not offer any information about fat distribution [1]. Nevertheless, some researchers have studied a set of dry-cured Iberian hams from pigs with different feeding system. 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 testers [2]. So, ham marbling, i.e. number, position, shape and size of intramuscular fat streaks in the ham must influence its high sensorial characteristics. Then, the design of a new 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.

Our attempts are to investigate objective physical parameters, which can be extracted from the analysis of digital Iberian ham images. To achieve this goal, not only the detection of fat streaks is important, but also the identification of their true contours, in order to calculate parameters as the relative size of one streak in relation to others, streak shape descriptors, etc.

Fat streaks can be considered as orientated line patterns. Several approaches to line detection have been described in the literature [3], [4], [5]. In this study, two line extractors at different scales and orientations combined with region growing techniques are applied to digital Iberian ham images in order to obtain scale-orientation information that will be used for automatic detection and identification of fat streaks.

2. Data

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 (256 gray levels). The area in pixels of the slices ranges from 131.109 to 203.779, i.e. 57 cm^2 to 72 cm^2. All intramuscular fat streaks have been annotated in the images by an expert in food technology. The total number of streaks annotated per image ranges from 21 to 47 (up to 454 in the fifteen images) and their areas range from 5.314 to 20 pixels, i.e. 343 mm^2 to 1.3 mm^2.

3. Methods

Identification of fat streaks in digital Iberian ham images is performed in two steps. Firstly, line extractors are applied to the original images at different scales and orientations in order to obtain scale-orientation information that will be used to generate a set of seed regions. Secondly, region growing techniques are locally applied to the above seeds.

3.1 Application of line extractors

Two line extractors (called line operator and directional morphological operator) have been developed. Both work as follows: an *L*x*L* kernel, whose active region consists of an orientated local line of *W* pixels width passing through the target pixel, is moved across the original image, pixel by pixel. The output value assigned

to the target pixel in the resulting image will depend on the operating principle of the line extractor addressed.

The line operator improves line signal to background noise ratio by taking the average gray level of the pixels lying on the active region and subtracting the average intensity of all the pixels in the locally orientated neighborhood.

The directional morphological operator is based on a similar principle to the line operator approach [6]. This operator improves line signal to background noise ratio by taking the minimum gray level of the pixels lying on the active region. The detection proceeds by a series of morphological openings using this directional structuring element.

In both cases, line orientation is obtained by applying the operator at *n* directions (8 in our experiment). The orientation producing the highest output (i.e. the maximum line strength) is taken as the line direction. Lines of different widths are detected by applying the operators at multiple scales.

3.2 Boundary fitting techniques

Four seeded region growing techniques, based on two different principles (region-based approaches and boundary-based approaches) are discussed.

Region-based approaches rely on the postulate that neighboring pixels within one region have similar gray-level. Our region-based algorithm follows a pixel-aggregation scheme, starting from a seed region and examining all its neighboring pixels, incorporating any of them that meets a growth acceptance criterion. When new pixels are accepted, they are adjoined to the seed region and the process is repeated with the resulting new region. The growth terminates when no acceptable neighbors exist.

The selection of an appropriate threshold is crucial to the success of this algorithm. In the published methods, threshold determination is usually based on histogram analysis. We have studied three different approaches to automatic selection of a threshold from the gray-level histogram of the original images: i) Otsu method, based on discriminant analysis, ii) Tsai method, based on the moment-preserving principle, and iii) Kapur Method, based on entropy considerations [7].

Boundary-based approaches to region growing rely on the postulate that pixel values change rapidly at the edges of regions, which can be detected by the application of a gradient operator such as the Sobel or Roberts filter. High values of the filters provide candidates for region boundaries. Nevertheless, these filters fail to produce closed contours due to sharp changes in line direction and noise. In order to obtain closed contours, a Laplacian of Gaussian zero-crossing edge detector is applied to the original image. Zero threshold is used to detect all zero-crossings in the image [8]. The region growing technique is performed as follows: every closed contour surrounding a seed region is marked as fat streak, and every contour with no seed region inside is misestimated.

4. Results

For the purpose of automatic detection of fat streaks, the proposed methods have been applied to our database of 15 digital images. The results obtained are compared by FROC curves [9], where the true positive rate (number of correctly detected fat streaks divided by the number of annotated streaks) is graphed against the average number of false positives (detection of surfaces which are not in the image) per image. Different values of the decision variable result in different compromises between true and false positive values. In our study, ten equidistant thresholds falling into the dynamic range of the images obtained after processing original images with line extractors have been considered.

We assume that a detected region is counted as true positive when both areas (detected and annotated) are at least *T*% pixels overlapped one each other. Several criteria have also been introduced with respect to false positives counting: i) detected regions smaller than the minimum annotated streak (approximately 20 pixels in our dataset) are ignored, and ii) regions very close one each other (less than 5 pixels) are merged.

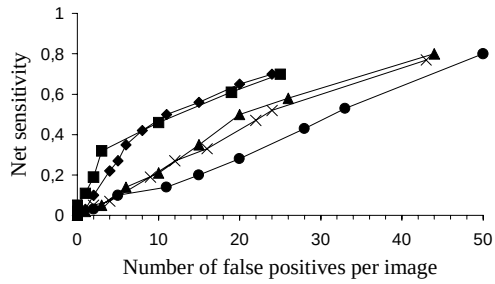

Figure 1. Comparison of fat streak detection performance for: simple thresholding (●) and original images processed by median filters (×), non-directional morphological operators (▲), directional morphological operators (■) and line operators (♦).

Figure 1 shows the results of applying the proposed method with line operators and directional morphological operators to our dataset of digital Iberian ham images, considering *T*=50%. As the original images present quite contrast, other intuitive techniques such as a simple thresholding of the original image and processing the original image with median filters or non-directional morphological operators are also applied in order to compare with the proposed approaches. The curves show that the line and directional morphological operator methods produce results which are much better than the results obtained by the other methods.

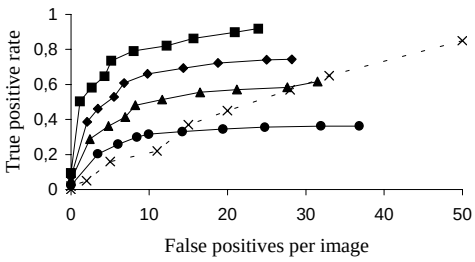

Figure 2. Weighted up comparison of fat streak detection performance for: simple thresholding at 50% tolerance (×) and line operators at 50% (■), 70% (♦), 80% (▲) and 90% (●).

An improved performance is obtained if detection is weighted up in relation to the size of the streaks within each image. In the gross FROC curve shown in figure 2, detected streaks by the line operator are weighted up by R_a/R_t, where R_a is the fat streak area in pixels and R_t is the summarized area of annotated streaks in the image. Four different values of the tolerance variable T, ranging from 50% to 90%, have been considered (simple thresholding at 50% is also addressed for the purpose of comparison). The results obtained are significantly better than those shown in figure 1. Similar results are obtained when directional morphological operators are applied to our dataset of images.

With respect to region growing, classification of segmented regions is done in five classes: correct detection, over-segmented streak, under-segmented streak, missed streak and noise streak [10]. Results for six tolerance thresholds, ranging from 50% to 95% (which indicates the effect of demanding more accuracy in true boundaries detection), are shown in figures 3 and 4.

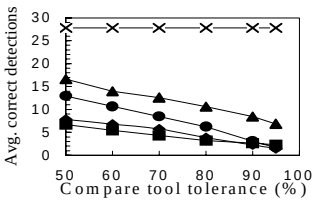

Figure 3. Average correct detections per image after applying the region-growing methods: Otsu (▲), Tsai (♦), Kapur (■) and Laplacian (●) from the seeds obtained by the line detector, on 15 digital Iberian ham images. The ideal case (×) is also addressed.

The region growing algorithm with the best global performance would obviously be the one with the highest scores for the average number of correct detections and the lowest value for all the error metrics. After analyzing the results shown in figures 3 and 4, some conclusions can be made: Otsu method achieves slightly superior overall performance than Laplacian method, and the other methods present poor performance in our application.

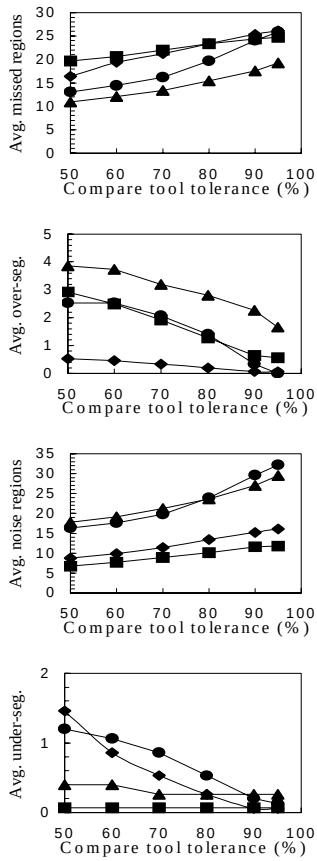

Figure 4. Average missed regions, over-segmentations, under-segmentations and noise regions per image after applying the region-growing methods: Otsu (▲), Tsai (♦), Kapur (■) and Laplacian (●) from the seeds obtained by the line detector, on 15 digital Iberian ham images.

5. Discussion

Detection performance of two line extractors has been evaluated. The sensitivities of both methods are comparable and show much better results than other intuitive techniques. The gross FROC shown in figure 2 improves sensitivities without changing the number of false positives. From this curve, it can be noted that the overall performance of our methods improves significantly when detections are weighted up by streak sizes. This fact indicates that the largest streaks are always detected.

On the other hand, the number of false positives is a priori significant for a given sensitivity. Nevertheless, we have experimentally proved that the accumulated area of false positives is much smaller than the accumulated area of true positives. Also, there is an interval where the accumulated areas of false positives and true negatives are comparable. This fact can be used to calculate the total fat percentage contained in streaks (in pixel units) as the sum

of true and false positive areas, which will be used to evaluate the relative fat content of the largest fat streaks. We have also proved that this area is nearly equal to the total area of annotated streaks. Both considerations lead to conclude that the relative fat content of the largest streaks can be estimated.

Figure 5 shows a quantitative analysis of the sizes of all detected, missed, noise, over-segmented and under-segmented regions after applying the region growing technique which presents the best performance (Otsu method) at 70% tolerance. Each bin in these graphs contains regions whose size is 10% of the maximum region, ordered by pixel size.

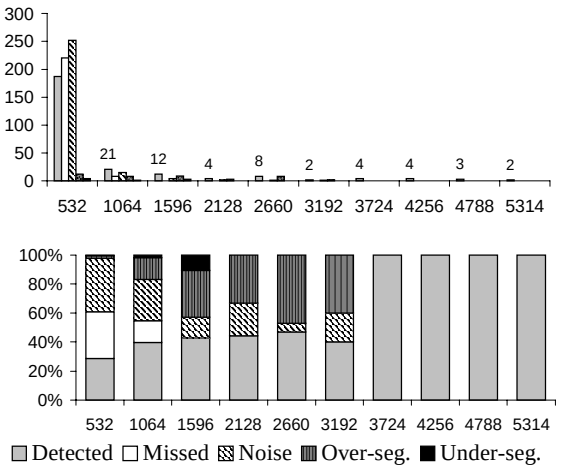

Figure 5. Number (upper) and rate (lower) of regions within each class (correct detections, over-segmentations, under-segmentations, missed streaks, noise streaks) for Otsu method at 70% compare tolerance.

The graphs show that missed and noise regions are predominantly smaller in size than detected regions. The proposed region growing algorithm most often misses small regions (on the order of 1000 pixels or less) and most of the false positives are also concentrated in the lowest bins. It is important to emphasize that the largest streaks in the image are correctly classified at this intermediate tolerance. This fact can be extrapolated to stricter tolerance criteria, which leads us to think that, although acceptable shape description is achieved for the largest streaks by our methods, the effect of over-segmentation and noise is still significant for the smallest streaks. We also believe that region connectivity techniques should be applied in the future to improve performance of the evaluated region growing techniques.

6. Conclusions and future work

Two approaches for the detection of fat streaks in digital Iberian ham images have been described. Their performance is comparable and much better than other intuitive techniques. Our methods allow us to estimate the relative size of the largest fat streaks, which could be a factor to predict Iberian ham sensorial quality.

Comparison of some published region growing techniques for the purpose of fat streak boundary fitting in digital Iberian ham images is also presented. The overall performance of Otsu method is superior to the other methods. Although this method can give acceptable shape description of the largest streaks in the image, over-segmentation and noise effect is significant for the smallest streaks. In order to achieve a complete shape description of all streaks, additional work is still needed. This work will be focused on connecting over-segmented regions and decreasing noise detections.

Acknowledgement

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

References

[1] Antequera, T., López-Bote, C.J., Córdoba, J.J., García, C., Asensio, M.A., Ventanas, J. and Díaz, Y., "Lipid oxidative changes in the processing of Iberian pig hams", *Food Chem.*, Vol. 45, No. 105, 1992.

[2] Cava, R., *Influencia de la Alimentación sobre los Fenómenos Oxidativos desarrollados durante la Maduración del Jamón de Cerdo Ibérico*, Doctoral Thesis, Universidad de Extremadura, Spain, 1997.

[3] Zwiggelaar, R., Parr, T.C., Boggis, C.R.M., Astley, S., Taylor, C.J., "Statistical modelling of lines and structures in mammograms". *Proc. SPIE, Medical Imaging* 30-34 1997.

[4] Merlet, N., Zerubia, J., "New Prospects in Line Detection by Dynamic Programming". *IEEE Trans. PAMI*, vol. 18, no. 4, April 1996.

[5] Yamada, H., Yamamoto, K., Hosokawa, K., "Directional Mathematical Morphology and Reformalized Hough Transformation for the Analysis of Topographic Maps", *IEEE Trans. PAMI*, vol. 15, no. 4, April 1993.

[6] Plaza, A., Cernadas, E., Durán, M.L., Sánchez, J.M., Petrón, M.J.., "Directional Morphological Operators to Detect Fat Streaks in Iberian Ham", *VIII NSPRIA*, Bilbao, Spain, 1999.

[7] Sahoo, P.K., Soltani, S., Wong, A.K.C., "A Survey of Thresholding Techniques", *Computer Vision, Graphics, and Image Processing* 41,233-260, 1988.

[8] Haralick, R.M. and Shapiro, L.G., *Computer and Robot Vision*, Volume I, Addison-Wesley, 1992.

[9] MacMillan, N.A., Creelman, C.D., *Detection Theory: A User´s Guide*, Cambridge University Press, 1991.

[10] 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, July 1996.