Multispectral Image Segmentation by Energy Minimization for Fruit Quality Estimation*

Adolfo Martínez-Usó, Filiberto Pla, and Pedro García-Sevilla

Dept. Lenguajes y Sistemas Informáticos, Jaume I Univeristy
{auso,pla,pgarcia}@uji.es
http://www.vision.uji.es

Abstract. This article presents the results of an unsupervised segmentation algorithm in multispectral images. The algorithm uses a minimization function which takes into account each band intensity information together with global edge criterion. Due to the unsupervised nature of the procedure, it can adapt itself to the huge variability of intensities and shapes of the image regions. Results shows the effectiveness of the method in multispectral fruit inspection applications and in remote sensing tasks.

1 Introduction

The main motivation of the developed work has been to obtain a method able to segment images of fruits for their quality classification in visual inspection processes using multispectral information. Particularly, this application problem implies the following requirements:

1. The method has to be able to segment *multispectral images* obtained from several wavelength ranges.
2. An *unsupervised method* would be needed due to manifold variables which can arise in fruit images. Thus, any prior knowledge should be avoided for the segmentation procedure.
3. The segmentation method has to be mainly based on *multispectral intensity* and *edge criteria*, in order to define the segmented region boundaries as accurately as possible.

The traditional effectiveness of a multiresolution Quadtree (QT) structure is combined with the multispectral intensity and edge information in a hierarchical representation. This leads us to an efficient strategy to solve the problem.

The method we are presenting starts from [4] as a natural development to multispectral images. Although this algorithm has achieved good results, it is quite obvious that multispectral images may offer us a lot of new and useful information. More concretely, in fruit inspection tasks we realize that there exist defects that can only be detected in certain bands of the spectrum and most of

* This work has been partly supported by grants DPI2001-2956-C02-02 from Spanish CICYT and IST-2001-37306 from the European Union

J.S. Marques et al. (Eds.): IbPRIA 2005, LNCS 3523, pp. 689–696, 2005.

the defects have a specific range of bands where they can be better discriminated. Therefore, fruit inspection by means of multispectral images acquires a high importance when we want to increase the classification quality and to obtain more accurate results in our application.

Multispectral images involve larger amount of information, noise, etc., so a preprocessing step has been included in order to improve the final results and accelerate the whole process:

1. On the one hand, we use an invariant representations of the input values to obtain multispectral images independent from some lighting and geometrical conditions [5]. These representations are very useful in those tasks that require invariance to shadows, orientation, reflections, etc.
2. On the other hand, we accelerate the whole process by means of the band selection proposed in [7]. Together with this acceleration, we have been able to keep the quality of the segmentation results just like if no band selection would have been applied.

The main contribution of the presented work is the proposed multispectral energy function that efficiently combines intra-region features with border information. The proposed framework yields satisfactory results, particularly in fruit inspection tasks.

2 Variational Image Segmentation

Variational methods for image segmentation develop algorithms and their mathematical analysis to minimize the segmentation energy E represented by a real value. The segmentation energy measures how smooth the regions are, the similarity between the segmented image and the original one and the similarity between the obtained edges and the discontinuities of the original image.

The Mumford-Shah model [6] has been regarded as a general model within variational segmentation methods. According to Mumford-Shah's conjecture, the minimal segmentation exists but it is not unique; for each image a set of minimal segmentations exists. This model looks for a piecewise smoothed image u with a set of discontinuities, edges of the original image g by means of minimizing the segmentation energy in the following equation, where K is the set of discontinuities in the image domain Ω representing the *edges* of g:

$$E(u,K) = \int_{\frac{\Omega}{K}} (|\nabla u(x)|^2 + (u-g)^2)dx + length(K) \qquad (1)$$

Since Mumford-Shah's work, several approaches appeared that suggested modifications to the original scheme. Recent works change equation (1) in order to improve the results. In this sense, the boundary function, which is binary in the Mumford and Shah's formulation, was changed by a continuous one which obtains a clearly defined boundary in [2]. Furthermore, in [1] the authors analyze some possible generalizations of the Mumford-Shah functional for color images. They suggest that these changes accentuate different features in edge detection and restoration.

In general, formulating variational methods have several advantages:

1. A variational approach returns explicitly a measure of the quality of the segmentation. Therefore, on the one hand we are able to know how good the segmentation is, on the other we may use it as a quantitative criterion to measure the segmentation quality.
2. Many segmentation techniques can be formulated as a variational method.
3. Finally, a variational approach provides a way to implement non-supervised processes by looking for a minimum in the segmentation energy.

3 Energy Minimization of the Quadtree Image Representation

In this work, a function to minimize the segmentation energy is proposed. With this assumption, it is important to point out that we cannot guarantee to find a global minimum but, the experimental results obtained show that the solutions are very satisfactory. In addition, we use a statistically robust functional that takes into account any resolution or scale change producing the same segmentation results in each case.

Let u be a smoothed image and a set of discontinuities of the original image g, let R_i be a set, with $0 < i \leq \|\Omega\|$ and $R_i \subseteq \Omega$. R_i is a family of r regions in Ω such that $\bigcup_{i=1}^{r} R_i = \Omega$ and $R_i \bigcap R_j = \emptyset$ for $i \neq j$. Let B_i represent the border of region R_i, that is, $R_i' = R_i - B_i$ is the inner part of region R_i. Finally, let γ be certain very small value that avoids dividing by zero.

Thus, $\forall x \in R_i$ let us consider the function $E_i(R_i, B_i)$:

$$E_i(R_i, B_i) = \int_{R_i} (|u(x) - m_i|)\, dx + \frac{\int_{R_i'} |\nabla g(x)|}{\int_{B_i} |\nabla g(x)| + \gamma}\, dx \tag{2}$$

In the image u with N bands, $u(x) = (u_1(x), ..., u_N(x))$ represents the intensity values in each band $u_j(x)$ of an R_i element x, and m_i is a central measure for the intensity value of R_i. The final segmentation energy is expressed as $E(\Omega) = \sum_i E_i(R_i, B_i) + \|\Omega\|$.

The QT structure allows us to divide an image within a complete multiresolution tree representation including neighboring information. This spatial information can be further used by a clustering strategy which groups the QT leaves using the intensity values in each band and the edge information.

Let us see the following discrete version of (2) with the same nomenclature,

$$E_i(g) = k \cdot H_i + (1-k) \cdot G_i + \lambda \cdot length(\Omega) \tag{3}$$

that returns the energy at each region. H_i and G_i are terms as follows:

$$H_i = \frac{\sum_{R_i} D(u(x), m_i)}{\sigma_{image}} \qquad\qquad G_i = \frac{\sum_{R_i - B_i} |\nabla g(x)|}{\sum_{B_i} |\nabla g(x)| + \gamma} \tag{4}$$

Specifically, in the QT representation, R_i is the set of leaves of the QT belonging to region i and B_i represents the boundary leaves in R_i, $0 < i \leq r$, with r being the number of regions at each iteration. The function D calculates a distance between two intensity sets of pixels (Euclidean, Manhattan, etc.). The value $|\nabla g(x)|$ returns the gradient magnitude at the pixel x. Note that the parameter $0 < k < 1$ allows to define the weight between the color and boundary information and λ allows the method to give more or less importance to the number of regions. Finally, the value σ_{image}, which is the sum of the standard deviations of each plane in the image, contributes to normalize the first term and makes the function statistically robust. Thus, in the energy functional (3) we can distinguish three components:

1. The first one takes into account the homogeneity of each region by means of a distance from each pixel to a central measure of the region it belongs (usually the median). So, the smaller this result, the more homogeneous the cluster is. This component represents our specific constraint to each region with the intensity information in each spectral band.
2. The second component promotes that the gradient magnitude will be low in the inner leaves and high in the boundary ones, that is, promotes regions with no edges inside. In this sense, gradient information is used by means of a boundary map which is made from the maximum gradient value found in each band.
3. The third term helps the function to punish a large number of regions and it is frequently used in variational image segmentation as in [2][1].

It is important to point out that the energy functional (3) has been designed in order to achieve a statistically robust behavior, being invariant to scale or/and intensity changes.

4 The Algorithm

4.1 The Preprocessing Step

1. The collection of multispectral images were obtained by an imaging spectrograph (RetigaEx, Opto-knowledged Systems Inc., Canada). The spectral range extended from 400 to 720 nm in the visible, with a spectral resolution of 10 nm, obtaining a set of 33 spectral bands for each image.
2. Multispectral applications with a limited time to finish usually have to deal with the band selection problem in order to characterize the problem without loss of accuracy. In this sense, considering that multispectral images contain information represented by means of a set of two dimensional signals, in [7], it is proposed a band selection from the point of view of information theory measures. So, dealing with the relationships among the set of bands representing the multispectral image, the optimal subset of spectral image bands can be obtained.
   Using a database of oranges, in [7] is shown that a 97.4% of classification performance is achieved with 6 bands, whereas using th 33 bands they manage a 98.5% in their unsupervised classification approach. We use the same database and select the same 6 resulting bands as in [7].

3. In [5], authors presented a set of illumination invariants to work with multispectral images. The results obtained showed the good performance of those invariants in order to avoid changes in the illumination intensity, highlights and object geometry. Thus, after the band selection, we apply the transformation:
$$C_n^{new} = \frac{C_n - \min(C_1, \ldots, C_N)}{\sum_j (C_j - min(C_1, \ldots, C_N))} \tag{5}$$
to obtain an invariant representation where C_i is the pixel value in band i and $1 \leq i \leq N$.

4.2 The Segmentation Process

The *QT* structure allows us to divide an image within a complete multiresolution tree representation including neighboring information. This spatial information can be further used by a clustering strategy which joins the *QT* leaves using intensity and edge information. The multiresolution process allows us to analyze the image with a coarse-to-fine sequence described as follows:

1. We construct a hierarchical structure level by level. It is important to clarify that, talking about a *QT*, the meaning of *level* is a set of leaves with the same size. Thus, the first level will be the first four children leaves that descend from the whole image and so on. Therefore, while the previous named *QT* is created, each level is revised by the functional (3) in order to revise the clustering at that resolution. Each cluster created in any level will be taken into account in the next levels. Finally, when we finish the *QT* construction, the *salient regions* have been detected in a coarse way.
2. Focusing the attention on the *salient regions* (the coarse ones that have been labeled), they will be taken as the most significant groupings of the image. So, we continue the process expanding each cluster by means of a region growing method where each cluster applies the functional (3) to its neighboring regions. This second step will take care of shaping the edges of each region by intensity and edge criteria.

Note that we use the functional (3) described in Sect. 3 in both of the previous steps but, whereas in the first one this application is guided by a hierarchical structure in order to develop each resolution level, in the second one the application of the functional follows a region growing strategy to achieve the final regions in a more accurate way.

Before summarizing the segmentation process, it is important to point out what are the main ideas the proposed method is based on:

1. We look for different features provided by the invariants in a first step. In addition, we accelerate the process by selecting an optimal set of spectral bands without loss of quality.
2. To guide the segmentation process, the following questions have to be solved:
   (a) *the way to continue the segmentation process.*
   (b) *how long the process have to continue.*

The first question is solved by means of a multiresolution analysis of the image with a QT structure. Multiresolution is able to decompose the image in several resolution levels developing a coarse-to-fine process from the *salient regions* to the final shapes of each region. On the other hand, the question (b) will be determined by the functional (3) described in Sect. 3. It will be minimized in a progressive way until the functional energy stops decreasing.

The whole segmentation process is summarized in the following iterative algorithm where each cluster is compared to all its neighboring clusters and it will be merged when the segmentation criterion is satisfied (see Sect.3).

1. Apply the band selection (6 from 33) and change the representation to the one described in [5].
2. Make the boundary map as an edge information reference.
3. Construct an oversegmented representation of the image, that is, expand the QT until every square region have all pixels with the same intensity. After this, create an ordered list according to region sizes.
4. The algorithm computes the functional (3) for each region and its neighbor regions in an iterative sequence that may be seen as a coarse-to-fine segmentation process.
5. If the whole image energy has decreased, reorder the list of regions by size and repeat the previous step. Arranging regions according to their size gives more importance to bigger regions and represents a spatial constraint that facilitates merging small regions with big ones.
6. Ignore very small regions (not identifiable with any defect) and merge them to their most similar neighbors.

This clustering process stops when no other merging can be performed without increasing the energy of the segmentation.

5 Results

Fruit image segmentation is used as input in a further process to characterize and classify the fruit surface. These visual inspection applications identify and detect different types of defects and parts of the fruit.

Fig.1 and Fig.2 show examples of oranges with two different types of defects on their surface, overripe defect and scratch defect respectively. In both images, the first row shows some image bands. These bands are not the ones selected by the band selection method, but they were chosen for illustration purposes.

Second row shows the edge results, where the darker the edge line is, the greater difference between the spectral values of neighboring regions. The first image on the left is the result after the segmentation algorithm is applied over all 33 bands. As we can see, this segmentation tends to finish with too many regions due to illumination and geometrical factors. When the segmentation algorithm is applied on the invariant representation (second images on the second rows) we obtain satisfactory results. In the last two images where the band selection and the invariant representation were used, the quality of the results increased, since

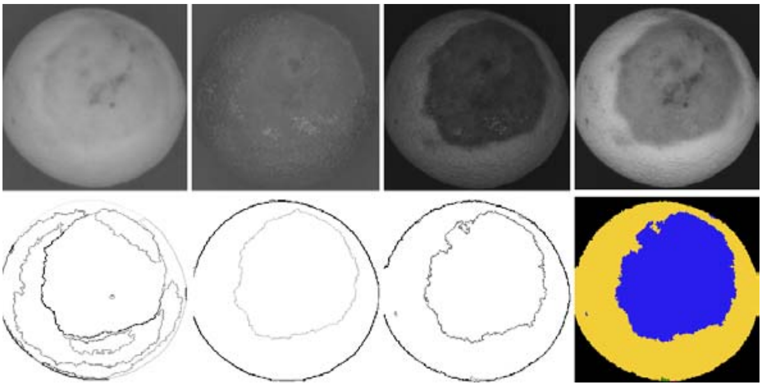

Fig. 1. First row corresponds to the bands 0, 8, 16 and 32 of an orange with an overripe defect in its surface (that is, 400, 480, 560 and 720 nm in the visible spectrum). Second row shows, from left to right, results using 33 bands, illumination invariants, 6 bands and illumination invariants and the final segmentation represented with random colors.

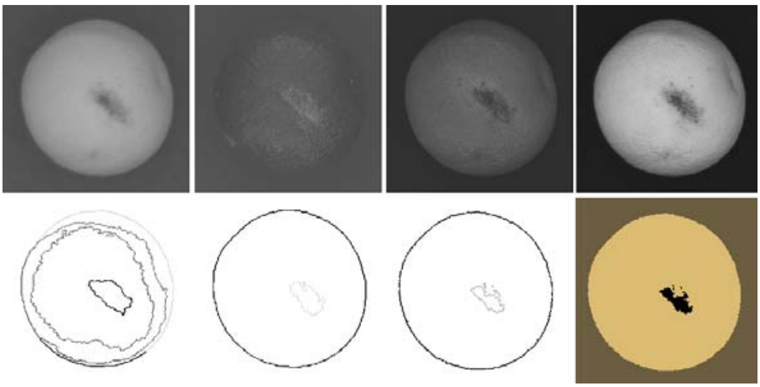

Fig. 2. As Fig.1, first row corresponds to the bands 0, 8, 16 and 32, but this orange has a scratch defect in its surface. Second row shows the edge results and the final segmentation in the same order as Fig.1.

the boundaries are more accurately defined and some new regions are discovered. Moreover, note how the segmentation process adapts to the regions of each image due to its unsupervised nature.

Thus, the segmentation obtained has found the different variations of the stains on the orange surface and this will allow the extraction of region descriptors for their classification in fruit quality assessment applications.

In order to show the performance of the presented method to other application fields, Fig.3 shows a multispectral image from satellite and the results obtained. This image has 220 bands and some of them are quite noisy (the special features of this image are explained in [3]). As we can see, the final segmentation has separated the brighter regions and the darker ones drawing the boundaries quite accurately.

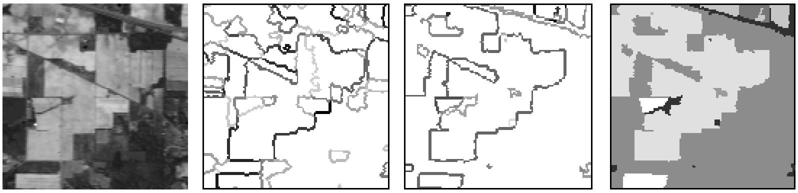

Fig. 3. Satellite image results. From left to right, original image (band 120), results using the band selection, results using the same selected bands and illumination invariants and the final segmentation represented with random colors.

6 Conclusions

In this paper, we have presented a multispectral segmentation algorithm based on a minimization function that uses the intensity and edge information. This algorithm has been combined with a preprocessing step which improves and accelerates the whole process. The results obtained show how the algorithm can adapt to the different situations and variability of intensity regions, being able to segment areas and locating the borders due to the use of gradient information during the segmentation process. Thus, this unsupervised segmentation strategy can locate different regions and find their contours satisfactorily.

Enclose to the minimization process we have developed a preprocessing step, based on a band selection and illumination invariants for multispectral images, which improves notably the accuracy of the results and a QT representation that guides the minimization process and increases the efficiency by means of a segmentation at different resolution levels.

References

1. A. Brook, R. Kimmel, and N.A. Sochen. Variational restoration and edge detection for color images. *Journal of Mathematical Imaging and Vision*, 18(3):247–268, 2003.
2. G.A. Hewer, C. Kenney, and B.S. Manjunath. Variational image segmentation using boundary functions. *IEEE Transactions on Image Processing*, 7(9):1269–1282, 1998.
3. Luis O. Jimenez and A. Landgrebe. Hyperspectral data analysis and supervised feature reduction via projection pursuit. *IEEE TGRS*, 37(6):2653–2667, 1999.
4. A. Martinez-Uso, Filiberto Pla, and Pedro Garcia-Sevilla. Color image segmentation using energy minimization on a quadtree representation. *LNCS*, (3211):25–32, 2004.
5. R. Montoliu, F. Pla, and Arnoud K. Klaren. Multispectral invariants. Technical Report, DLSI, Universitat Jaume I, Castellon, Spain, 2004.
6. D. Mumford and J. Shah. Optimal approximations by piecewise smooth functions and associated variational problems. *CPAM*, 42(4), 1989.
7. Jose M. Sotoca, Filiberto Pla, and Arnoud K. Klaren. Unsupervised band selection for multispectral images using information theory. *ICPR*, 2004.