
Image Segmentation Through an Iterative Algorithm of the Mean Shift

Roberto Rodríguez Morales, Didier Domínguez,
Esley Torres and Juan H. Sossa

Additional information is available at the end of the chapter

<http://dx.doi.org/10.5772/50474>

1. Introduction

Image analysis is a scientific discipline providing theoretical foundations and methods for solving problems appearing in a range of areas as diverse as biology, medicine, physics, astronomy, geography, chemistry, meteorology, robotics and industrial manufacturing, among others.

Inside any image analysis system, an aspect of vital importance for pattern recognition and image interpretation that has to be taken into account is segmentation and contour extraction. Both problems can be really difficult to face due to the variability in the form of the objects and the variation in the image quality. An example can be found in the case of biomedical images which are frequently affected by noise and sampling, that can cause considerable difficulties when rigid segmentation methods are applied [Chin-Hsing et al., 1998; Kenong & Levine, 1995; Koss et al., 1999; Rodríguez et al., 2002].

Many segmentation techniques are available in the literature and some of them have been widely used in different application problems. Most of these segmentation techniques were motivated by specific application purposes. Many different approaches for image segmentation there are; which mainly differ in the criterion used to measure the similarity of two regions and in the strategy applied to guide the segmentation process. The definition of suitable similarity and homogeneity measures is a fundamental task in many important applications, ranging from remote sensing to similarity-based retrieval in large image databases.

Segmentation is an important part of any computer vision and image analysis system, wherein regions of interest are identified and extracted for future processing. Of the quality

of segmentation depends, on great measure, the good performance of higher level analysis steps such as recognition and interpretation.

However, in spite of the most complex algorithms developed until now, segmentation continues to be very application dependent. A single method that can solve the multitude of present problems there is not. It still remains a complex problem with no exact solution that by means of traditional low-level techniques, such as: thresholding, region growing and other classical operations requires a considerable amount of interactive guidance in order to attain satisfactory results. Automating these model-free approaches is difficult because of complexity, shadows, and variability within and across individual objects.

For years, the most suitable algorithms have been the iterative methods. These cover a variety of techniques, ranging from the mathematical morphology based methods, the deformable models up to thresholding based methods. However, one of the problems of these iterative techniques is the stopping criterion, for which many strategies have been proposed [Vincent & Soille, 1991; Cheriet et. al., 1998; Chenyang et. al., 2000].

Mean shift (MSH) is a robust technique which has been applied in many computer vision tasks, as by example: image segmentation, visual tracking, etc. [Shen & Brooks, 2007]. MSH technique was proposed by Fukunaga and Hostetler [Fukunaga et. al., 1975] and largely forgotten until Cheng's paper [Cheng, 1995] rekindled interest in it. MSH is a versatile non-parametric density analysis tool and it can provide reliable solutions in many applications [Comaniciu, 2002]. In essence, MSH is an iterative mode detection algorithm in the density distribution space. The MSH procedure moves to a kernel-weighted average of the observations within a smoothing window. This computation is repeated until convergence is obtained at a *local* density mode. This way the density modes can be located without explicitly estimating the density. An elegant relation between the MSH and other techniques can be found in [Shen & Brooks, 2007].

The iterative algorithm that is used in this chapter is based on the mean shift and in several works was previously introduced and applied [Rodríguez & Suarez, 2006; Rodríguez, 2008; Domínguez & Rodríguez, 2009; Domínguez & Rodríguez, 2011; Rodríguez et. al., 2011a; Rodríguez et. al., 2011b; Rodríguez et. al., 2012]. In those papers, entropy was used as a stopping criterion. Entropy is not a new concept in the information theory field and it has been used in image restoration, edge detection and as an objective evaluation method for image segmentation [Zhang, 2003].

In this chapter is presented a research, using standard images and real images, based on a segmentation algorithm which used an iterative computation of the mean shift filtering. A comparison of the obtained results was carried out, according to the number of iterations and the degree of homogenization of the segmented images. Also, a comparison of the obtained results with our algorithm with other segmentation methods already established was carried out.

The aim of this chapter is to present the advances that the authors have obtained in the field of the image segmentation. Also, some strategies that constitute suitable tools are presented, which it can be used in many system of image analysis where methods of segmentation are required. The main contribution of this chapter is to analyze how the quality of the segment-

ed images varies for different values of the window sizes (hr and hs) and the stopping criterion. Many of the obtained results were compared with other methods.

This chapter continues as follows: In Section 2 the most significant theoretical aspects on mean shift are detailed. In Section 3, we shortly introduce the entropy concept and we also give some comments on this. The iterative algorithm of the mean shift is described in Section 4. In Section 5 the used standard images are presented. Moreover, some of the characteristics of the real images are described. In Section 6 the experimental results are exposed, and also an analysis and discussion of these are carried out. Finally, in Section 7 the most important conclusions of this chapter are given.

2. Theoretical aspects

The iterative procedure to compute the mean shift is introduced as normalized density estimate of the gradient. By employing a differentiable kernel, an estimate of the density gradient can be defined as the gradient of the kernel density estimate; that is,

$$\hat{\nabla} f(x) = \nabla \hat{f}(x) = \frac{1}{nh^d} \sum_{i=1}^n \nabla K\left(\frac{x-x_i}{h}\right) \tag{1}$$

Conditions on the kernel $K(x)$ and the window radio h are derived in [Fukunaga & Hostetler, 1975] to guarantee asymptotic unbiasedness, mean-square consistency, and uniform consistency of the estimate in the expression (1). For a radial symmetry kernel,

$$K(x) = C_k(\|x\|^2)$$

where the profile is $r = \|x\|^2$, then; for example, for *Epanechnikov* kernel (other choices are possible as will be seen below),

$$K_E(x) = \begin{cases} 1/2C_d^{-1}(d+2)(1-\|x\|^2) & \text{if } \|x\|^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

The density gradient estimate becomes,

$$\hat{\nabla} f_E(x) = \frac{1}{n(h^d C_d)} \cdot \frac{d+2}{h^2} \sum_{x_i \in S_h(x)} (x_i - x) = \frac{n_x}{n(h^d C_d)} \cdot \frac{d+2}{h^2} \left(\frac{1}{n_x} \sum_{x_i \in S_h(x)} (x_i - x) \right) \tag{2}$$

where the region $S_h(x)$ is a hypersphere of radius h having volume $h^d C_d$, centered at x , and containing n_x data points; that is, the uniform kernel. The last term in expression (2) is called the *sample mean shift*,

$$M_{h,U}(x) = \frac{1}{n_x} \sum_{x_i \in S_h(x)} (x_i - x) = \frac{1}{n_x} \sum_{x_i \in S_h(x)} x_i - x \tag{3}$$

The quantity $\frac{n_x}{n(h^d c_d)}$ is the kernel density estimate $\hat{f}_U(x)$ (the uniform *kernel*) computed with the hypersphere $S_h(x)$, and thus we can write the expression (2) as:

$$\hat{\nabla} f_E(x) = \hat{f}_U(x) \cdot \frac{d+2}{h^2} M_{h,U}(x) \tag{4}$$

which yields,

$$M_{h,U}(x) = \frac{h^2}{d+2} \frac{\hat{\nabla} f_E(x)}{\hat{f}_U(x)} \tag{5}$$

Expression (5) shows that an estimate of the normalized gradient can be obtained by computing the sample mean shift in a uniform kernel centered on x . In addition, the mean shift has the direction of the gradient of the density estimate at x when this estimate is obtained with the *Epanechnikov* kernel. Since the mean shift vector always points towards the direction of the maximum increase in the density, it can define a path leading to a local density maximum; that is, to a mode of the density (see Fig. 1).

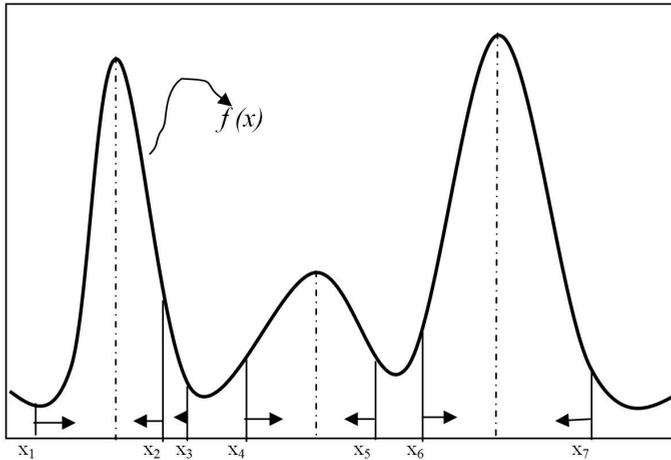


Figure 1. Gradient mode clustering.

In addition, expression (5) shows that the mean is shifted towards the region in which the majority of the points reside. Since the mean shift is proportional to the local gradient estimate, it can define a path leading to a stationary point of the estimated density, where these

stationary points are the modes. Moreover, as it was pointed out the normalized gradient in expression (5) introduces a desirable adaptive behavior, since the mean shift step is large for low density regions corresponding to valleys, and decreases as x approaches a mode. This is possible to see in a clear way in Figure 2.

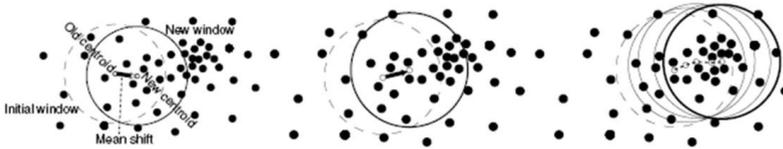


Figure 2. Local maxima of the probability density given by samples.

Mathematically speaking, this is justified since $\frac{\hat{\nabla}f_E(x)}{\hat{f}_U(x)} > \hat{\nabla}f_E(x)$. Thus the corresponding step size for the same gradient will be greater than that nearer mode. This will allow observations far from the mode or near a local minimum to move towards the mode faster than using $\hat{\nabla}f_E(x)$ alone.

In [Comaniciu & Meer, 2002] it was proven that the *mean shift procedure*, obtained by successive:

- computing the mean shift vector $M_h(x)$
- translating the window $S_h(x)$ by $M_h(x)$,

guarantees convergence.

Therefore, if the individual mean shift procedure is guaranteed to converge, it is hoped that a recursively procedure of the mean shift also converges. In other words, if we consider an iterative procedure like the individual sum of many procedures of the mean shift and each individual procedure converges; then, the iterative procedure also converges. The question that continues open is when to stop the recursive procedure. The answer is in the use of the entropy, as it will be shown in next Section.

2.1. Generalization

Employing the profile notation the density estimate can be written as [Comaniciu, 2000],

$$\hat{f}_k(x) = \frac{1}{nh^d} \sum_{i=1}^n k\left(\left\|\frac{x-x_i}{h}\right\|^2\right) \quad (6)$$

By denoting with $g = -k'$, that is, the profile defined by the derivative of profile k with the sign changed (we assume that the derivative of k exists $\forall x \in [0, \infty)$), excepting a finite set of points), then the density gradient estimate (see expression (1)) becomes,

$$\begin{aligned} \hat{\nabla} f_K(x) &= \nabla \hat{f}_K(x) = \frac{2}{nh^{d+2}} \sum_{i=1}^n (x - x_i) k' \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \\ &= \frac{2}{nh^{d+2}} \left[\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \right] \left[\frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x \right] \end{aligned} \quad (7)$$

where $\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)$ is assumed to be nonzero.

One can observe that the derivate of the *Epanechnikov* profile is the uniform profile, while the derivate of the normal profile remains as exponential.

The last bracket in expression (7) contains the mean shift vector computed with a kernel $G(x)$ defined by $G(x) = cg \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)$, where c is a normalization constant, that is,

$$M_{h,G}(x) = \frac{\sum_{i=1}^n x_i g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)}{\sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right)} - x = \frac{\sum_{i=1}^n x_i G \left(\frac{x - x_i}{h} \right)}{\sum_{i=1}^n G \left(\frac{x - x_i}{h} \right)} - x \quad (8)$$

Then, the density estimate at x becomes,

$$\hat{f}_G(x) = \frac{1}{nh^d} \sum_{i=1}^n G \left(\frac{x - x_i}{h} \right) = \frac{c}{nh^d} \sum_{i=1}^n g \left(\left\| \frac{x - x_i}{h} \right\|^2 \right) \quad (9)$$

By using the expressions (8) y (9), the expression (7) becomes,

$$\hat{\nabla} f_K(x) = \hat{f}_G(x) \cdot \frac{2}{h^2 c} M_{h,G}(x) \quad (10)$$

from where it follows that,

$$M_{h,G}(x) = \frac{h^2 c}{2} \frac{\hat{\nabla} f_K(x)}{\hat{f}_G(x)} \quad (11)$$

Expression (11) is a generalization of the mean shift vector. This allows to use other kernels; for example, Gauss kernel, which gives wonderful results.

On the other hand, a digital image can be represented as a two-dimensional array of p -dimensional vectors (pixels), where $p=1$ in the gray level case, three for color images, and $p > 3$ in the multispectral case. As was pointed in [Comaniciu & Meer, 2002] when the location and range vectors are concatenated in the joint spatial-range domain of dimension $d = p + 2$, their different nature has to be compensated by proper normalization of parameters h_s and h_r . Thus, the multi-variable kernel is defined as the product of two radially symmetric kernels and the Euclidean metric allows a single bandwidth for each domain, that is:

$$K_{h_s, h_r}(x) = \frac{C}{h_s^2 h_r^p} k\left(\left\|\frac{x^s}{h_s}\right\|^2\right) k\left(\left\|\frac{x^r}{h_r}\right\|^2\right) \tag{12}$$

where x^s is the spatial part, x^r is the range part of a feature vector, $k(x)$ the common profile used in both domains, h_s and h_r , the employed kernel bandwidths, and C the corresponding normalization constant.

One can observe in Figure 2 that the l_2 norm is implicitly used in order to define the neighborhoods of pixels. From a mathematical point of view the concept of norm is associated with the size of the elements of a given space. Given a linear space L over a field K and an element $x \in L$ is defined as norm of x , denoted $\|x\|$, a finite functional which satisfies some conditions [Domínguez & Rodríguez, 2009]. As we have pointed out, when $S_h(x)$ is defined

as expression (2) implicitly makes use of the l_2 norm defined as, $\|x\|_2 = \sqrt{\sum_{j=1}^d x_j^2}$, $x \in \mathfrak{R}^d$, since $S_h(x) = \{x' : \|x - x'\|_2 \leq h\}$.

Note that in order to verify the condition $x_i \in S_h(x)$ in (2), for each x_i it is necessary to calculate $\|x - x_i\|_2$ which entails conducting d elevations to the second power, $d-1$ sums and calculating one square root. Verifying the same condition using the l_∞ norm, defined as $\|x\|_\infty = \max_j |x_j|$ only involves calculating the maximum value in the module of components of the difference vector $x - x_i$.

In [Domínguez & Rodríguez, 2009; Domínguez & Rodríguez, 2011], we carried out a theoretical and practical study related with this issue. We proved the convergence of the mean shift by using the l_∞ norm. The convergence of mean shift for discrete data was proved in [Comaniciu, 2000] using the l_2 norm for defining the hypersphere $S_h(x)$. The following theorem guarantees the convergence when it replaces the l_2 norm by the l_∞ norm. The proof is similar to the theorem proved in [Comaniciu, 2000] and it can be found in [Domínguez & Rodríguez, 2011].

Theorem 1

Let $\hat{f}_E = \left\{ \hat{f}_k(y_k, K_E) \right\}$ the sequence of density estimates obtained using Epanechnikov kernel and computed in the points y_k defined by the successive locations of the mean shift procedure with uniform kernel and $N(x)$, denoting $\|x\|_N$, a norm that satisfies $N(x) \leq \|x\|_2$, $\forall x \in \mathbb{R}^d$. If the hypersphere $S_h(y_k)$ is defined using $N(x) \forall k \in \mathcal{I}$, then the sequence is convergent.

As a direct consequence of this theorem, the mean shift algorithm converge using the l_∞ norm when defining the hypersphere $S_h(x)$ because $\|x\|_\infty \leq \|x\|_2$, $\forall x \in \mathbb{R}^d$.

3. Entropy

From the point of view of digital image processing, entropy of an image I is defined as:

$$E(I) = - \sum_{x=0}^{2^B-1} p(x) \log_2 p(x) \quad (13)$$

where B is the total quantity of bits of the digitized image and by agreement $\log_2(0)=0$; $p(x)$ is the probability of occurrence of a gray-level value. Within a totally uniform region, entropy reaches the minimum value. Theoretically speaking, the probability of occurrence of the gray-level value, within a uniform region is always one. In practice, when one works with real images the entropy value does not reach, in general, the zero value. This is due to the existent noise in the image. Therefore, if we consider entropy as a measure of the disorder within a system, it can be used as a good stopping criterion, by the use of the mean shift filtering, for an iterative process. Entropy within each region diminishes in measure in that the regions become more homogeneous, and at the same time in the whole image, until reaching a stable value. When convergence is reached, a totally segmented image is obtained, because the mean shift filtering is not idempotent. In addition, as in [Comaniciu & Meer, 2002] was pointed out, the mean shift based image segmentation procedure is a straightforward extension of the discontinuity preserving smoothing algorithm and the segmentation step does not add a significant overhead to the filtering process.

The choice of entropy as a measure of goodness deserves several observations. Entropy reduction diminishes the randomness in corrupted probability density function and tries to counteract noise. Then, by following this analysis, as the segmented image is a simplified version of the original image, entropy of the segmented image should be smaller. Recently, it was empirically found that the entropy of the noise diminishes faster than that of the signal [Suyash et. al., 2006]. Therefore, an effective criterion to stop would be when the relative rate of change of the entropy from one iteration to the next, falls below a given threshold. All these observations were the main motivation in seeking a segmentation procedure from the iterations of the mean shift filtering. This new algorithm is much simpler [Rodríguez & Suarez, 2006].

4. Algorithms

In general, an image captured with a real physical device is contaminated with noise and in most cases a statistical model of white noise is assumed, mean zero and variance σ . For smoothing or elimination of this form of noise many types of filters have been published, the most classic being the low pass filter. This filter indiscriminately replaces the central pixel in a window by the average or the weighted average of pixels contained therein. The end result with this filtering is a blurred image; since this reduces the noise but also important information is taken away from the edges. However, there are low pass filtering techniques that preserve the discontinuities and reduce abrupt changes near local structures. A diverse number of approaches have been published taking into consideration the use of adaptive filtering. These range from an adaptive Wiener filter, local isotropic smoothing, to an anisotropic filtering. The mean shift works in the spatial-range domain, but differs from it in the use of local information. The algorithm that was proposed in [Comaniciu & Meer, 2002] for filtering through mean shift is as follows:

Let $\{x_i\}_i$ and $\{z_i\}_i, i=1, \dots, n$ be the input and filtered images in the joint spatial-range domain. For each pixel $p \in x_i, p=(x, y, z) \in \mathfrak{R}^3$, where $(x, y) \in \mathfrak{R}^2$ and $z \in [0, 2^\beta - 1], \beta$ being the quantity of bits/pixel in the image. The filtering algorithm comprises the following steps:

For each $i=1, \dots, n$

1. Initialize $j=1$ and $y_{i,j} = p_i$.
2. Compute the mean shift in order to obtain the mode where the pixel converges; that is, the calculation of the mean shift is carried out until convergence, $y = y_{i,c}$.
3. Store at Z_i the component of the gray level of calculated value: $Z_i = (x_i^s, y_{i,c}^r)$, where x_i^s is the spatial component and $y_{i,c}^r$ is the range component.

4.1. Segmentation algorithm by recursively applying the mean shift filtering

4.1.1. Algorithm No. 1

Let $ent1$ be the initial value of the entropy of the first iteration. Let $ent2$ be the second value of the entropy after the first iteration. Let $errabs$ be the absolute value of the difference of entropy between the first and the second iteration. Let $edsEnt$ be the threshold to stop the iterations; that is, to stop when the relative rate of change of the entropy from one iteration to the next, falls below this threshold. Then, the segmentation algorithm comprises the following steps:

1. Initialize $ent2 = 1, errabs = 1, edsEnt = 0.001$.
2. While $errabs > edsEnt$, then
3. Filter the image according to the steps of the previous algorithm; store in $Z^{[k]}$ the filtered image.
4. Calculate the entropy from the filtered image according to expression (8); store in $ent1$.

5. Calculate the absolute difference with the entropy value obtained in the previous step; $errabs = |ent1 - ent2|$
6. Update the value of the parameter; $ent2 = ent1$; $Z^{[k+1]} = Z^{[k]}$

It can be observed that, in this case, the proposed segmentation algorithm is a direct extension of the filtering algorithm, which ends when the entropy reaches stability. The effectiveness of this algorithm will be proven along this chapter. In this work the thresholding value ($edsEnt$) was empirically obtained. Recent investigations have proven that smaller values of the threshold do not affect, qualitatively nor quantitatively in dependence on original image, the final result of the segmentation. One will be able to see these results in this chapter. More details and discussion on this issue will be given in the next section.

In [Christoudias et. al., 2002], it was stated that the recursive application of the mean shift property yields a simple mode detection procedure. The modes are the local maxima of the density. Therefore, with the new segmentation algorithm, by recursively applying mean shift, convergence is guaranteed. Indeed, the proposed algorithm is a straightforward extension of the filtering process. In [Comanociu, 2000], it was proven that the mean shift procedure converges. In other words, one can consider the new segmentation algorithm as a concatenated application of individual mean shift filtering operations. Therefore, if we consider the whole event as linear, the recursive algorithm converges.

4.1.2. Algorithm No. 2: Binarization algorithm by recursively applying the mean shift filtering

This algorithm is very similar to the algorithm No. 1, only that in this occasion two steps are added. This continue of this way,

1. Initialize $ent2 = 1$, $errabs = 1$, $edsEnt = 0.001$.
2. While $errabs > edsEnt$, then
 - 2.1. Filter the image according to the steps of the previous algorithm; store in $Z[k]$ the filtered image.
 - 2.2. Calculate the entropy from the filtered image according to expression (6); store in $ent1$.
 - 2.3. Calculate the absolute difference with the entropy value obtained in the previous step; $errabs = |ent1 - ent2|$.
 - 2.4. Update the value of the parameter; $ent2 = ent1$; $Z[k+1] = Z[k]$.
3. To carry out a parametric logarithm ($parlog = 70$, this is the parameter).
4. Binarization: to assign to the background the white color and to the objects the black color.

In the experimentation was proven that the final result is not very sensitive to this parameter, because a variation in the range from 60 to 90 led to the same result [Rodriguez, 2008].

5. Used standard images and utilized real images. Some characteristics

In Figure 3 a representation of the used standard images for this research appear. Some characteristics on these standard images can be commented.



Figure 3. Standard images. (a) Cosmonaut, (b) Baboon, (c) Barbara, (d) Bird, (e) Cameraman, (f) Peppers, (g) Lake, (h) Mountain, (i) Lena.

For example, one can observe that some of these images are rich in high frequencies (Baboon and Barbara), other are rich in low frequencies (Bird and Peppers, these have more homogeneous zones) while other images have both, low and high frequencies (Cosmonaut and Cameraman). These characteristics will influence on the behavior of iterative algorithm, in particular, on the number of iterations. This issue will be deeply analyzed in Section of experimental results.

Other real images used in this work can be seen in Figure 4. These images are biopsies, which represent an angiogenesis process in malignant tumors. These were included in paraffin by using the immunohistoquimic technique with the complex method of avidina biotina. Finally, monoclonal CD34 was contrasted with green methyl to accentuate formation of new blood vessels (angiogenesis process). These biopsies were obtained from soft parts of

human bodies and the images were captured via MADIP system with a resolution of 512x512x8 bit/pixels [Rodríguez et. al., 2001].

Several notable characteristics of these images there are; which are common to typical images that we encounter in the tissues of biopsies. For example, the intensity is slightly darker within the blood vessel than in the local surrounding background. It is emphasized that this observation holds only within the local surroundings. In addition, due to acquisition protocol, the images are corrupted with a lot of noise. For more details on these images refer to [Rodríguez et. al., 2005].

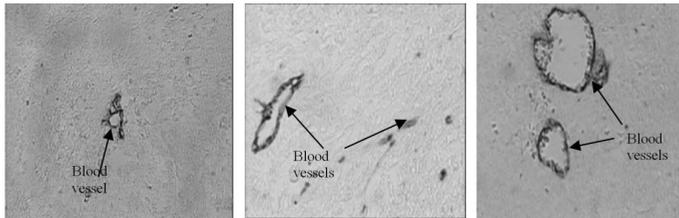


Figure 4. These images represent the angiogenesis process. The blood vessels are marked with arrows.

6. Experimental results and discussion

Image segmentation, that is, classification of the image intensity-level values into homogeneous areas is recognized to be one of the most important steps in any image analysis system. Homogeneity, in general, is defined as similarity among the pixel values, where a piecewise constant model is enforced over the image [Comaniciu and Meer, 2002].

All the segmentation experiments in this work were performed by using a uniform kernel. In order to be effective the comparison of the obtained results with our algorithm and with the *EDISON* system [Christoudias et. al., 2002], the same parameters (hr and hs), in both procedures, were used.

The value of hs is related to the spatial resolution of the analysis while the value hr defines the range resolution. It is necessary to note that the spatial resolution hs has a different effect on the output image when compared to the gray level resolution (hr , *spatial range*). Only features with large spatial support are represented in the segmented image with our algorithm when hs is increased. On the other hand, only features with high contrast survive when hr is large. Therefore, the quality of segmentation is controlled by the spatial value hs and the range (gray level) hr , resolution parameters defining the radii of the (3D)/2D windows in the respective domains. As our algorithm is a direct extension of the filtering algorithm similar behavior was also reported in [Comaniciu and Meer, 2002]. In addition, as our algorithm does not need parameter M , for the effects of the comparison the same one was set to $M = 1$ in the *EDISON* system.

The first preliminary results when applying our algorithm were published in the year 2006 [Rodríguez & Suarez, 2006]. In those researchers a quantitative comparison was not carried out, the comparison was only visual. The aim of that moment was alone to give to know the existence of our algorithm and to carry out a comparison with another established already [Christoudias et. al., 2002]. A deeper explanation on the characteristics of our algorithm was published in the year 2011[Rodríguez et. al., 2011a]. Nevertheless, two examples of the results reached in the year 2006 appear in the Figure 5 and 6.

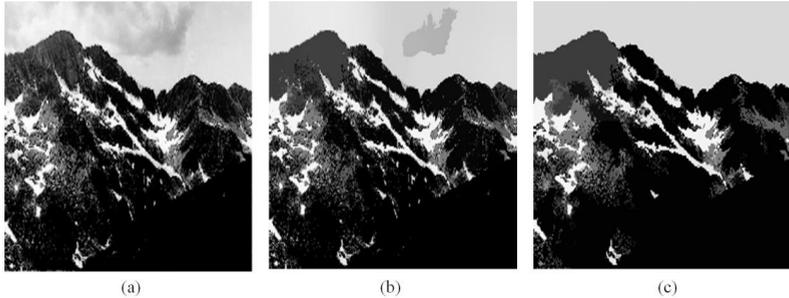


Figure 5. a) Original image, (b) Segmented image according to our algorithm $(hs, hr) = (12, 15)$, (c) Segmented image by using *EDISON* system $(hs, hr, M) = (12, 15, 1)$.

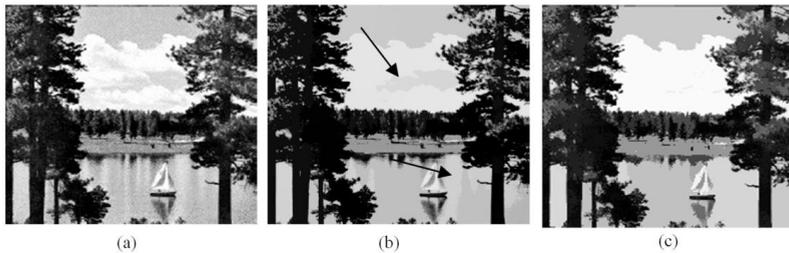


Figure 6. a) Original image, (b) Segmented image by our strategy $(hs, hr) = (12, 15)$, (c) Segmented image according to *EDISON* system, $(hs, hr, M) = (12, 15, 1)$. The arrows in the Fig. 2(b) indicate better segmented regions.

From the point of view of the final result, the image segmented with our algorithm has a more natural aspect. In many occasions, given the application, segmentation imposes certain conditions (elimination of regions, pruning or integration of certain maxima, etc). This can originate a biased image with regard to the original image. With our algorithm the resolution is only imposed on the segmentation process; that is, the parameters hr and hs . For this reason, our algorithm does not make mistakes; that is, a segmented image, very different to the original image, is not obtained. This is one of the most important experimental results obtained with our algorithm.

It is important to point out that with both algorithms (the proposed one and the *EDISON* system) very similar results were obtained (only differences in very few regions; see the ar-

rows). The substantial difference between both algorithms is the one shown in Fig. 6(c), it was necessary to carry out a filtering step and later on a segmentation step. In this last step, one can have certain complexity when adjacency graphs and hierarchical techniques are used [Comaniciu, 2000]. With our algorithm the segmented image is directly obtained from the filtering process. However, it is necessary to have in mind that segmentation is very dependent on the application. For this reason, in order to compare our proposal with *EDISON* system, the most remarkable differences were looked for.

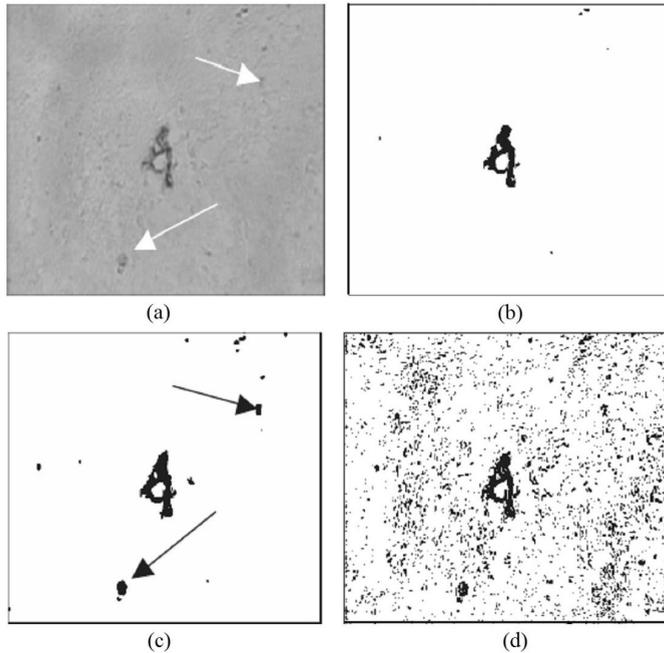


Figure 7. a) Original image, (b) Binarized image by using our new algorithm, (c) Binarized image by using graph [Rodríguez, 2008], (d) Binarized image via Otsu's method [Otsu, 1978].

Note in Figure 5 that the clouds and the sky were better isolated with our algorithm. This result is explained by the fact that our algorithm is a direct extension of the filtering process and, therefore, it does not produce many mistakes. In [Grenier et. al., 2006] the mean shift filtering was also iteratively applied in order to increase the smoothing effect. However, the difference with our algorithm is that in that work a stopping criterion was not given. The authors iterated the mean shift 10 times before starting the segmentation process.

A direct application of our algorithm for the binarization of blood vessels in an angiogenesis process was published in [Rodríguez, 2008]. Two examples appear in the Figures 7 and 8. In the year 2005 were obtained a similar result with a more complicated algorithm [Rodríguez et. al., 2005].

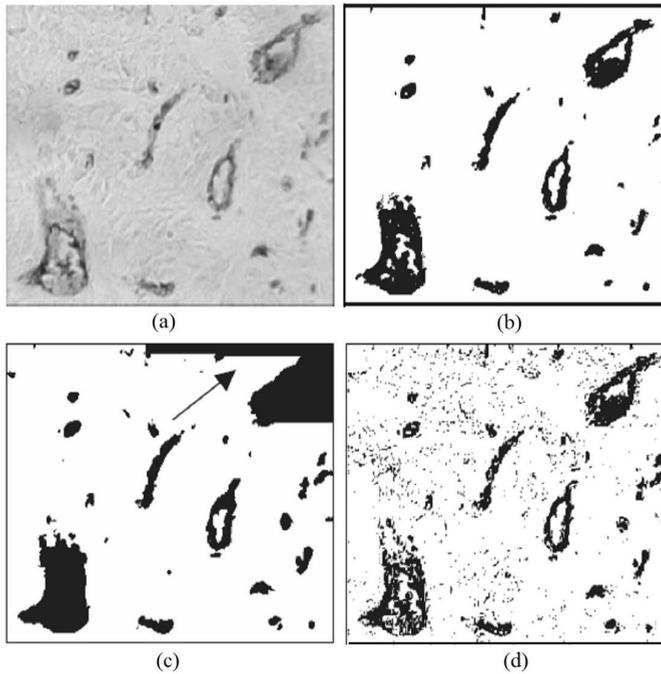


Figure 8. a) Original image, (b) Binarized image by using our new algorithm, (c) Binarized image by using graph, (d) Binarized image via Otsu's method.

It is evident to observe that the binarized image by using the new algorithm has a better appearance than the obtained image by using graph. Note, in this case, that the binarization algorithm by using graph made a mistake (see arrow in Figure 8 (c)). In practice, it has been proven that this behavior did not always happen with all images and in the corners of the images this was manifested fundamentally. We note in Figures 7 and 8 that the binarized image using the algorithm No. 2 was cleaner and it did not accentuate the spurious information that appears in the original image (see arrows in Figure 7(a)). According to criterion of pathologists these objects (spurious, with little contrast) were originated by a problem in the preparation of the samples. The obtained result by using Otsu's method is evident, a lot of noisy arose. This best result with the new binarization algorithm is because the same one is a direct extension of the filtering process. The parameter used to carry out the parametric logarithm was similar to 70 and this value was the same for all the binarized images. For more details on these results see [Rodriguez, 2008].

As it was expressed previously the l_2 norm is implicitly used in order to define the neighborhoods of pixels when working with the mean shift. An interesting issue is to analyze that it happens when substituting the l_2 norm by the l_∞ norm. In such a sense, we will show a series of experiments conducted with the aim of comparing, in terms of execution time and degree

of homogenization, the obtained results by two segmentation algorithms. The graphic of the time spent by both algorithms on a group of standard images is presented and analyzed. In order to carry out the comparison of the obtained results the same parameters ($h_r = 15$ and $h_s = 4$) in both procedures, by using the l_∞ norm and the l_2 norm were used.

In Fig. 9 (c), the segmentation of the image *Astro* is presented by using the algorithm *No.1* described in Section 2.3.1. In Fig. 9 (b) the result using the algorithm that makes use of the l_∞ norm is shown. In Fig. 9 (b) it can be seen that the segmented image using the l_∞ norm presents a greater degree of homogenization. Comparing these images visually, it is evident that the use of the l_∞ norm leads to a greater similarity in the value of intensity of certain groupings of pixels (see Earth zone). This greater degree of homogenization can be seen as an advantage if the algorithm is used in an application where one wants to extract the figure of the astronaut. However, in an application where the objects of interest are the clouds, their elimination would become a drawback. This corroborates that the segmentation is heavily dependent on the application.

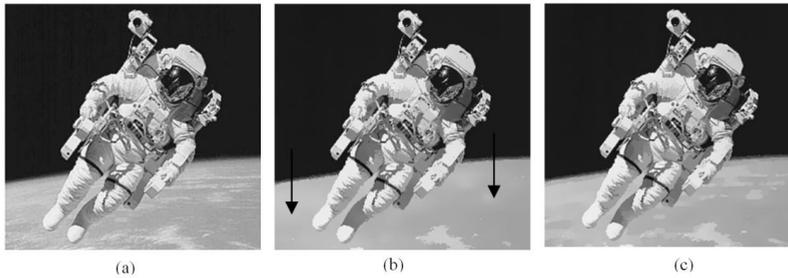


Figure 9. a) Original image, (b) Segmented image using the l_∞ norm, (c) Segmented image using the l_2 norm.

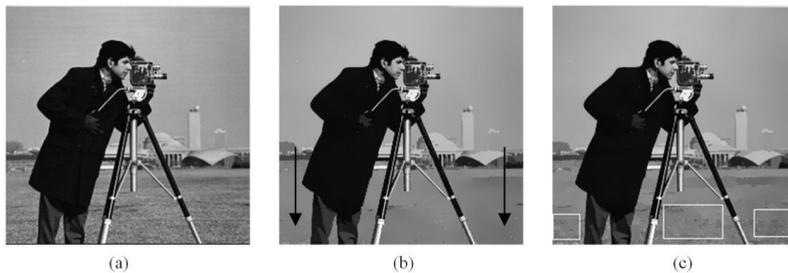


Figure 10. a) Original image, (b) Segmented image using the l_∞ norm, (c) Segmented image using the l_2 norm.

Other example of segmentation is presented in Figure 10 by using standard images. As in the previous example, there is again a greater homogenization when the neighborhood by using the l_∞ norm is defined. In this case, from a standpoint of a visual comparison, in Fig-

ure 10(b) the arrows indicate parts of major homogenization. For example, in the image of Figure 10(c), where the l_2 norm is used, the boxes indicate parts which have a lesser degree of homogeneity between the pixels that represent the grass of the field.

Figure 11 shows a graphic of the execution times of the algorithms that make use of the l_∞ and l_2 norms. The values of the runtime for each image using the l_2 norm in the definition of $S_h(x)$ are represented by circles, while the squares represent the runtime associated with the l_∞ norm. As is shown in Figure 11, in general, the runtime of the algorithm that makes use of the l_∞ norm is higher than using the l_2 norm.

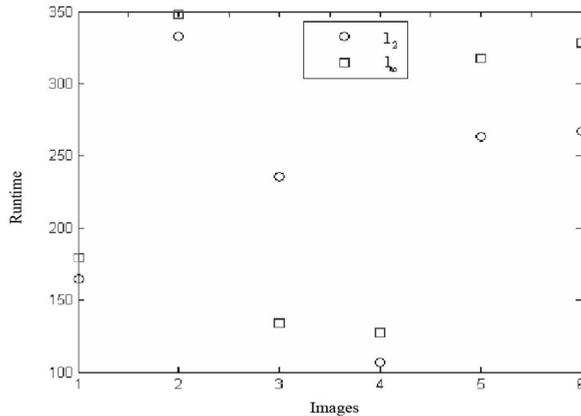


Figure 11. Runtime of the algorithms for standard images.

The greater homogenization observed using the l_∞ norm to define $S_h(x)$ suggests the search for values h_s and h_r in order to obtain more efficient results and smaller runtime. The readers interested in deepening in these results to see [Domínguez & Rodríguez, 2009].

Another issue that attracted the attention of the authors was the theoretical demonstration of the mean shift when the l_∞ norm is used. The convergence of the algorithm by using the l_∞ norm was empirically shown through an extensive experimentation [Domínguez & Rodríguez, 2009]. In [Domínguez & Rodríguez, 2011] was proven a theorem which guarantees the convergence of the l_∞ norm instead of the l_2 norm in order to define the region $S_h(x)$. The convergence of mean shift for discrete data was proved in [Comaniciu, 2000] using the l_2 norm for defining the hypersphere $S_h(x)$.

Table 1 shows the obtained results using $h_r=8$ and $h_s=2$. As can be seen, for these values, execution times were lower using the l_2 norm. The values were comparable with those obtained using the l_∞ norm in order to define the neighborhoods of the pixels and the maximum difference between the runtimes was 96,876 seconds, which was obtained with the image Baboon.

Image	Norm	hr	hs	Time
Cosmonaut256	l_2	8	2	186.1410
Cosmonaut256	l_∞	8	2	207.6880
Baboon256	l_2	8	2	76.4370
Baboon256	l_∞	8	2	173.3130
Barbara256	l_2	8	2	96.6570
Barbara256	l_∞	8	2	156
Bird256	l_2	8	2	82.7810
Bird256	l_∞	8	2	78.6250
Cameraman256	l_2	8	2	124.0470
Cameraman256	l_∞	8	2	209.2030
Peppers256	l_2	8	2	94.0780

Table 1.

Image	Norm	hr	hs	Time
Cosmonaut 256	l_2	15	4	164.8590
Cosmonaut 256	l_∞	15	4	179.2660
Baboon256	l_2	15	4	332.5160
Baboon256	l_∞	15	4	348.3130
Barbara256	l_2	15	4	235.2970
Barbara256	l_∞	15	4	134.0780
Bird256	l_2	15	4	107.2810
Bird256	l_∞	15	4	127.0630
Cameraman256	l_2	15	4	263.4380
Cameraman256	l_∞	15	4	317.1870
Peppers256	l_2	15	4	267.2650
Peppers256	l_∞	15	4	328.3590

Table 2.

In Table 2, it can be seen that for window sizes $hr=15$ and $hs=4$ the runtime was in favour of the l_2 norm (difference of 61,094 seconds). This result was obtained with the image Peppers.

However, one can observe that in most of the images the difference of the runtimes were decreased when the values $h_r=8$ and $h_s=2$ were used (see Table 1). Moreover, in case of image Barbara the runtime using the l_∞ norm was smaller than the runtime using l_2 norm. The difference was 101.219 seconds.

This suggests the use of the l_∞ norm in segmentation of high-resolution images, which may be necessary in many practical cases; it can be an interesting tool in order to obtain more efficient results. It was evidenced, through an extensive experimentation using standard images, that the use of the l_∞ norm, instead of l_2 norm, decreases the runtime of the mean shift when the values of bandwidths h_s and h_r increase. For more details on this issue see [Domínguez & Rodríguez, 2011].

Another application of our algorithm was in the medical image segmentation. It is of noticing that the mean shift can be considered as a segmentation unsupervised method. Unsupervised methods, which do not assume any prior scene knowledge which can be learned in order to help segmentation process, it are obviously more challenging than the supervised ones.

In order to have more clarity of the medical images that will be segmented, some details of the original images are given. Studied images were of arteries, which have atherosclerotic lesions and these were obtained from different parts of the human body. These arteries were contrasted with a special tint in order to accentuate the different lesions in arteries. The arteries were digitalized directly from the working desk via MADIP system with a resolution of 512x512x8 bit/pixels [Rodríguez et. al., 2001]. For more details on the characteristics of these images one can see [Rodríguez & Pacheco, 2007]. Another lesion type that were isolated is caused by disease of the visual system, glaucoma. "Glaucoma" is a term used for a group of diseases that can lead to damage to the optic nerve and result in blindness.



Figure 12. a) Original image, b) Unsupervised segmentation by using our algorithm. The arrows mark the isolated lesions.

In Figure 12, an example of segmentation on artery by using our algorithm is shown. Although another segmentation method was already applied to other atherosclerotic lesions

[Rodríguez & Pacheco, 2007]; here one can observe the obtained result when applying our unsupervised strategy.

In Figure 12, one can note that the lesion IV that appears in the original image was isolated (see arrows in Fig. 12 (a)). According to the criterion of physicians this is a good result, because the algorithm is able to isolate the lesion without any previous condition. Moreover, one also can see that the segmented image with the mean shift algorithm is totally free of noise. This is another important aspect when the mean shift filtering is used.

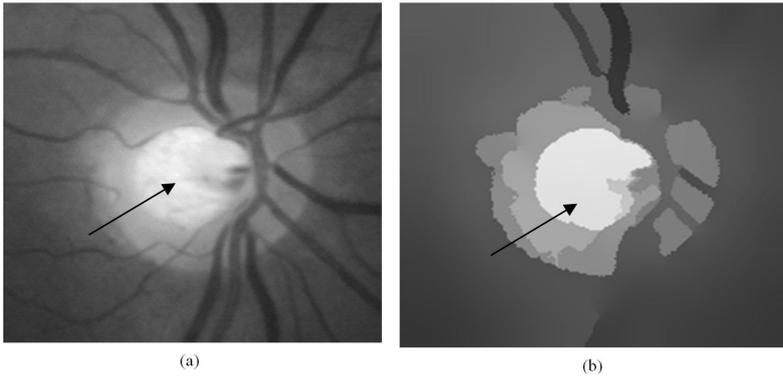


Figure 13. a) Original image. b) Segmentation by using our unsupervised strategy. The arrow indicates the isolated lesion.

Another example is shown in Figure 13. In this case, the main objective is to isolate the oval from the vascular net of the eye (see arrow). This is of great importance for the study of the glaucoma disease. According to the criterions of physicians, the discrimination of this area is of great importance in order to know the advancement of the disease. In this case, this zone is isolated appropriately. A quantitative comparison of all the shown experimental results can be found in the presented references of the own authors [Rodríguez & Tovar, 2010].

Other issue of interest of the authors was to study the behaviour of the algorithm, taking into consideration the number of iterations and the degree of homogenization of the segmented images, for different values of the stopping threshold and window sizes. First, we go to analyze what happens when the values of the stopping threshold goes decreasing, and later on we will carry out an analysis when varying the window sizes.

The segmentation of the Astro's image for different values of the stopping threshold is shown in Figure 14. One can appreciate that the number of iterations increased in an abrupt way when the parameter $edsEnt$ diminished from 0.001 to 0.0001 . However, from the point of view of a visual analysis a substantial change is not noticed in these segmented images (see Figures (f) and (g)). Homogenization is very similar. For such a reason in [Rodríguez et al., 2012], a comparison through the XOR was carried out in order to better appreciate the difference among these images.

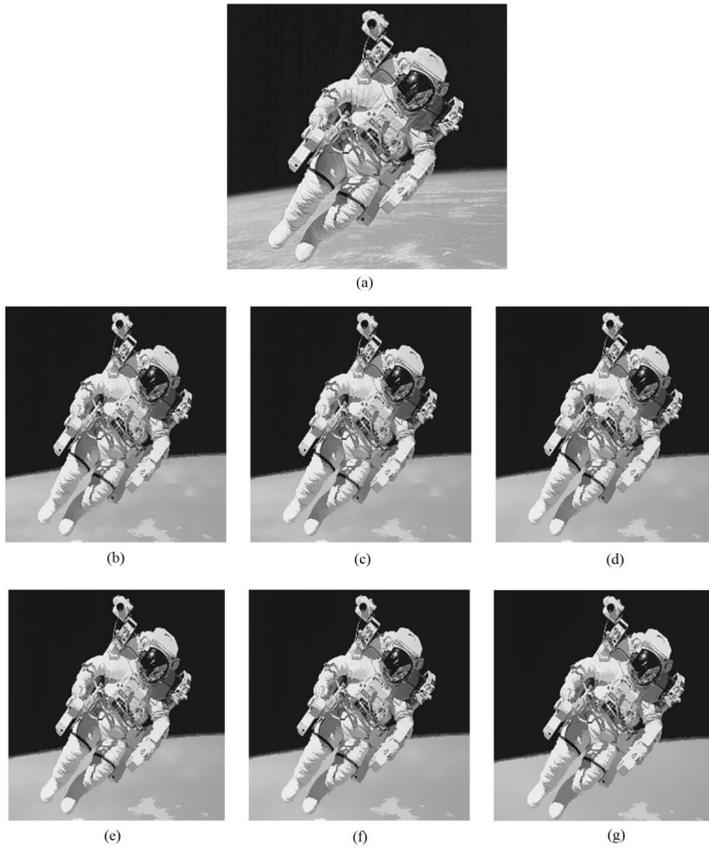


Figure 14. a) Original image (Astro), (b) Segmentation for $edsEnt = 0.1$, 2 iterations, (c) Segmentation for $edsEnt = 0.05$, 2 iterations, (d) Segmentation for $edsEnt = 0.01$, 4 iterations, (e) Segmentation for $edsEnt = 0.005$, 5 iterations, (f) Segmentation for $edsEnt = 0.001$, 7 iterations, (g) Segmentation for $edsEnt = 0.0001$, 60 iterations.

In this research, all the segmentation procedures were carried out by using a uniform kernel. We used the same window size in all the experiments ($hr = 15$, $hs = 4$), with the aim that the comparison of the obtained results was valid for different values of the stopping threshold (parameter $edsEnt$).

One can observe that, in dependence on the image features the number of iterations varied and the same one has not a lineal behaviour. Figure 15, it presents the obtained segmentation results with the baboon's image. To observe, for example, that for $edsEnt = 0.005$ the image of Figure 14 (e) was obtained in 5 iterations, while for that same value, the image of Figure 15 (e) was attained in 14 iterations. This is due to the quantity of low or high frequencies of the original image. Note that, the image of Figure 14 is rich in low frequencies (this image has more homogeneous areas). Opposite happens with the image of Figure 15 (this image is rich in high frequencies).

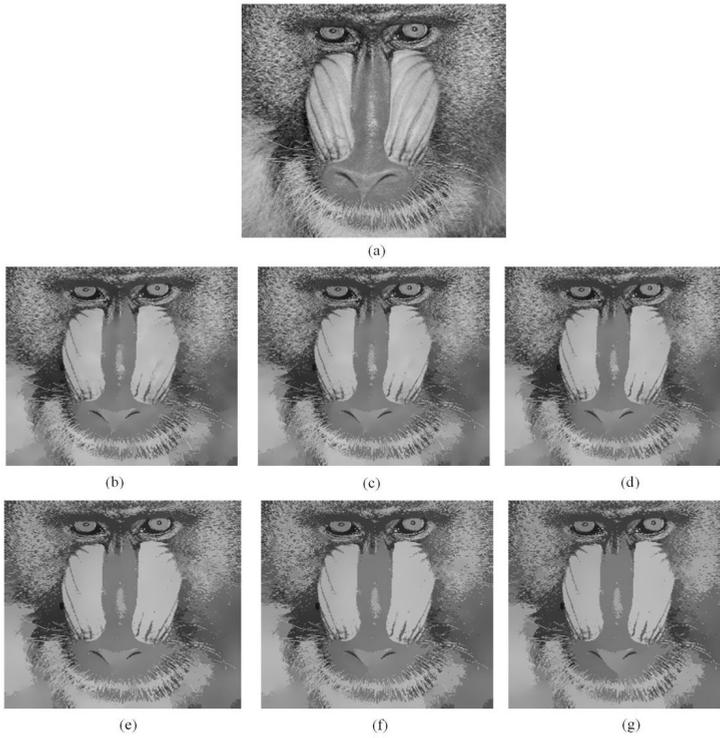


Figure 15. a) Original image (Baboon), (b) Segmentation for $edsEnt = 0.1$, 2 iterations, (c) Segmentation for $edsEnt = 0.05$, 3 iterations, (d) Segmentation for $edsEnt = 0.01$, 5 iterations, (e) Segmentation for $edsEnt = 0.005$, 14 iterations, (f) Segmentation for $edsEnt = 0.001$, 27 iterations, (g) Segmentation for $edsEnt = 0.0001$, 57 iterations.

Image	<i>EdsEnt</i>	No. Iterac	<i>EdsEnt</i>	No. Iterac	<i>EdsEnt</i>	No. Iterac
Astro	0.1	2	0.05	2	0.01	4
Baboon	0.1	2	0.05	3	0.01	5
Bird	0.1	2	0.05	2	0.01	3
Barbara	0.1	2	0.05	2	0.01	2
Image	<i>EdsEnt</i>	No. Iterac	<i>EdsEnt</i>	No. Iterac	<i>EdsEnt</i>	No. Iterac
Astro	0.005	5	0.001	7	0.0001	60
Baboon	0.005	14	0.001	27	0.0001	57
Bird	0.005	4	0.001	18	0.0001	19
Barbara	0.005	8	0.001	8	0.0001	42

Table 3. Values of the stopping threshold (*EdsEnt*) and number of iterations.

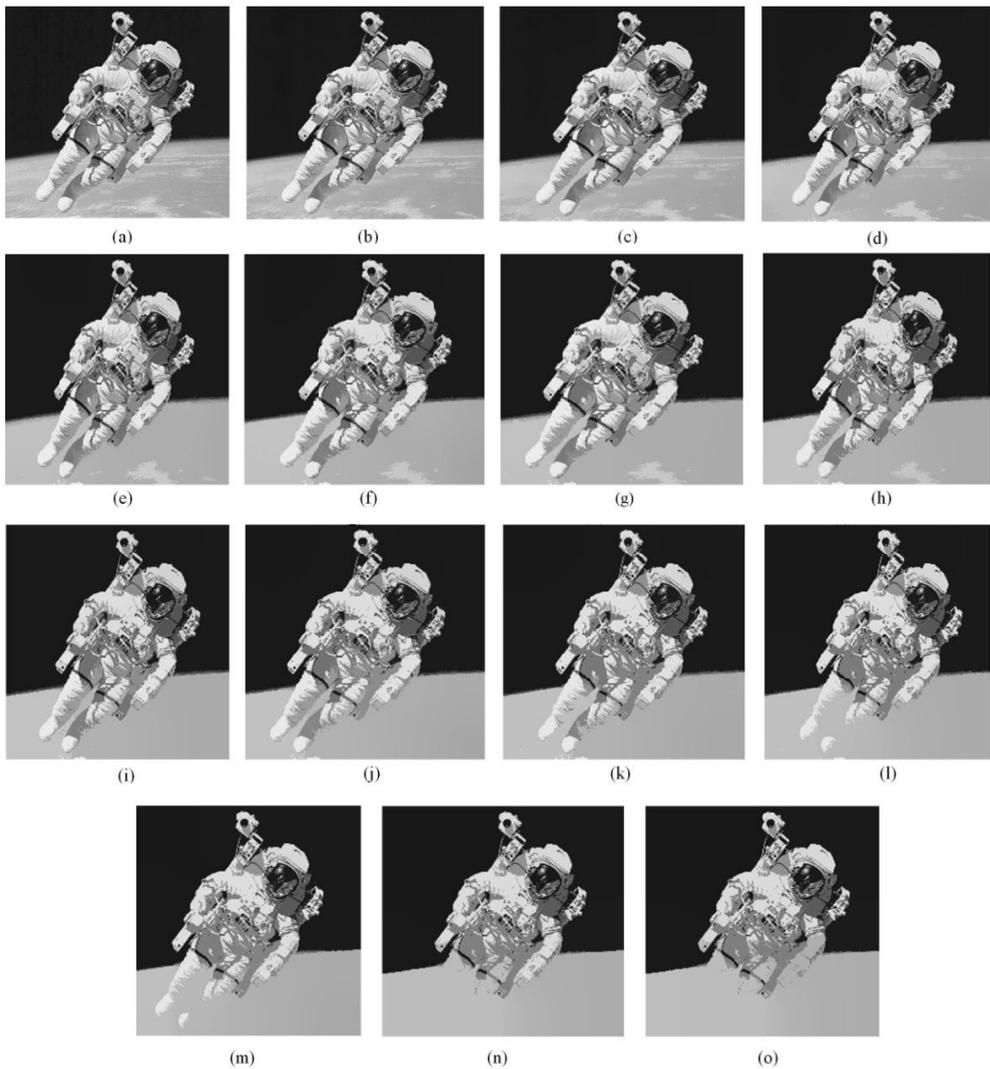


Figure 16. a) Original image (Astro), (b) Segmentation for $hr = 5$ and $hs = 2$, 6 iterations, (c) Segmentation for $hr = 7$ and $hs = 4$, 7 iterations, (d) Segmentation for $hr = 9$ and $hs = 6$, 12 iterations, (e) Segmentation for $hr = 11$ and $hs = 8$, 65 iterations, (f) Segmentation for $hr = 13$ and $hs = 10$, 70 iterations, (g) Segmentation for $hr = 15$ and $hs = 12$, 11 iterations, (h) Segmentation for $hr = 17$ and $hs = 14$, 39 iterations, (i) Segmentation for $hr = 19$ and $hs = 16$, 55 iterations, (j) Segmentation for $hr = 21$ and $hs = 18$, 7 iterations, (k) Segmentation for $hr = 23$ and $hs = 20$, 7 iterations, (l) Segmentation for $hr = 25$ and $hs = 22$, 37 iterations, (m) Segmentation for $hr = 27$ and $hs = 24$, 4 iterations, (n) Segmentation for $hr = 29$ and $hs = 26$, 23 iterations, (o) Segmentation for $hr = 31$ and $hs = 28$, 17 iterations. The arrow indicates the cloud permanency for that window size (image (h)).

It is necessary to point out that, for large values of the stopping threshold ($edsEnt$), the number of iterations had a very similar behaviour for images rich in low frequencies as well as for images rich in high frequencies (see Table 3 and Figures 14 and 15). On the other hand, one can appreciate that, in this image (see Figure 15), the number of iterations also increased abruptly when the stopping threshold was from $edsEnt = 0.001$ to 0.0001 . However, between the two images (see Figures 15 (f) and (g)), the difference is not visually appreciated. This appreciation was also analysed with more detail in [Rodríguez et. al., 2012].

This same study was carried out, fixing the value of stopping threshold, for different values of window sizes (parameters hs and hr). In this case the selected stopping threshold was $edsEn = 0.001$ (see algorithm No. 1). The segmentation of the Astro's image for different parameters hr and hs , in Figure 16 is represented.

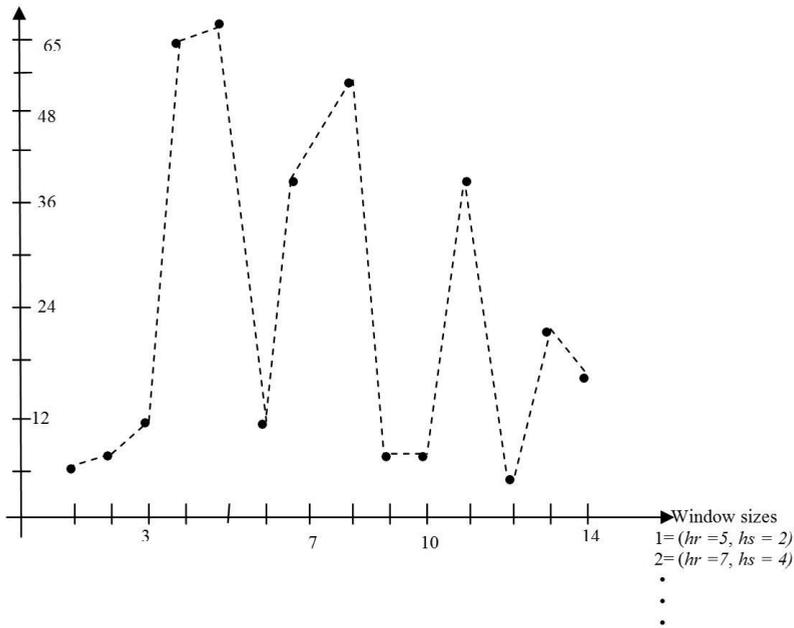


Figure 17. Graph that represents the number of iterations vs. the window sizes. Note the undulant behavior of the number of iterations with regard to the window sizes.

Some interesting comments arise of the obtained results that appear in Figure 16. Observe that, proportionally to the increase of the width of the radii hr and hs , the homogenization degree increased in the segmented image. However, the number of iterations had a behavior that fluctuated. In other words, the number of iterations did not increase linearly, in this image, when increased the radius hr and hs . Note that, in small radius the segmentation was very rude. One can observe that visually, homogeneous areas were not denoted with regard to the original image (see Figures 16 (b), (c) and (d)).



Figure 18. a) Original image (Lena), (b) Segmentation for $hr = 5$ and $hs = 2$, 14 iterations, (c) Segmentation for $hr = 7$ and $hs = 4$, 16 iterations, (d) Segmentation for $hr = 9$ and $hs = 6$, 25 iterations, (e) Segmentation for $hr = 11$ and $hs = 8$, 48 iterations, (f) Segmentation for $hr = 13$ and $hs = 10$, 25 iterations, (g) Segmentation for $hr = 15$ and $hs = 12$, 20 iterations, (h) Segmentation for $hr = 17$ and $hs = 14$, 26 iterations, (i) Segmentation for $hr = 19$ and $hs = 16$, 24 iterations, (j) Segmentation for $hr = 21$ and $hs = 18$, 21 iterations, (k) Segmentation for $hr = 23$ and $hs = 20$, 26 iterations, (l) Segmentation for $hr = 25$ and $hs = 22$, 13 iterations, (m) Segmentation for $hr = 27$ and $hs = 24$, 5 iterations, (n) Segmentation for $hr = 29$ and $hs = 26$, 8 iterations, (o) Segmentation for $hr = 31$ and $hs = 28$, 10 iterations. The arrows indicate the permanency of stains for those window sizes.

However, for window sizes very big the earth is totally homogeneous (see arrow in Fig. 16 (i)). In other words, starting from $hr = 19$ and $hs = 16$, the earth is totally uniform (see Figures 16 (j), (k) and (l)). Moreover, one visually does not observe difference among these images. Alone, in the images of Figures (n), (m) and (o), the feet of the cosmonaut were combined with the gray level of the earth. In order to verify this visual impression, a comparison of these images through the *Xor* was carried out in [Rodríguez et. al., 2011b]. On the other hand, it is possible to see that, in all these segmented images (for big window sizes), the behavior of the number of iterations was also oscillating; that is, these did not grow proportionally when increased the window sizes (see Fig. 17).

Figure 18 shows the obtained results with the image of Lena. Observe in Figure 18 that, the same as in the previous example, proportionally to the increase of the width of the radius hr and hs , the degree of uniformity increased in the segmented images. However, in this case for small radii, contrary to the previous example, the numbers of iterations were bigger. The explanation of this behavior, it comes given by the characteristics of the original image (high frequencies in the image). On the other hand, the number of iterations in this example also had an oscillating behaviour, that is; these did not increase proportionally with the window sizes.

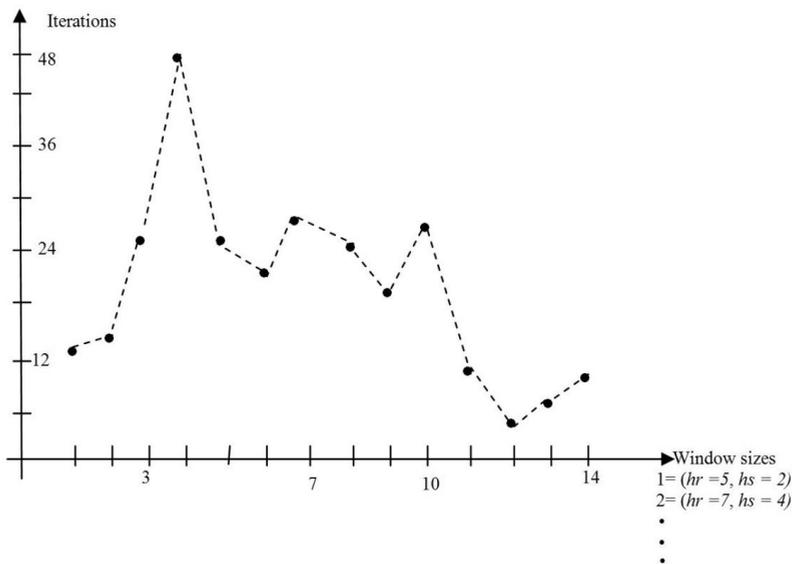


Figure 19. Graph that represents the number of iterations vs. the window sizes. Note the undulant behavior of the number of iterations with regard to the window sizes.

In this result, one can see that, starting from $hr = 13$ and $hs = 10$, the face of woman begins to be uniform and the stain below of the left eye disappeared (see arrow in Fig. 18 (e)). Moreover, starting from $hr = 19$ and $hs = 16$, the face of woman is totally uniform and

the shade that is under the nose also disappears (see arrow in Fig. 18 (h)). Here also in these results the behavior of the number of iterations, in relation to the increase of the window sizes hr and hs , were oscillating; these did not grow proportionally when increased the window sizes (see Fig. 19).

By way of summary, in Table 4, the window sizes and the iteration numbers appear for each of the segmented images.

Astro(hr, hs)		No. Iter.	Lena (hr, hs)		No. Iter.
5	2	6	5	2	14
7	4	7	7	4	16
9	6	12	9	6	25
11	8	65	11	8	48
13	10	70	13	10	25
15	12	11	15	12	20
17	14	39	17	14	26
19	16	55	19	16	24
21	18	7	21	18	21
23	20	7	23	20	26
25	22	37	25	22	13
27	24	4	27	24	5
29	26	23	29	26	8
31	28	17	31	28	10

Table 4. Window sizes (hr and hs) and the iteration numbers.

Table 4, it offers a comparative panoramic vision of all the segmented images. Also, one can see the behaviour of the iteration numbers with regard to the window sizes. This behaviour can be explained via Figure 20. In Figure 20, the value of hs is related to the spatial resolution of the analysis while the value hr defines the range resolution. The spatial resolution hs has a different effect on the output image when compared to the gray level resolution (hr , *spatial range*). Only features with large spatial support are represented in the segmented image with our algorithm when hs is increased. On the other hand, only features with high contrast survive when hr is large. Then, as $2xhs$ establishes the range of the movement spatially and in the range of $2xhr$ is carried out an averaged; then the characteristics of image in this range ($2xhr$) will influence on this average (bigger quantity of low or high space frequencies). This issue is what produced the oscillation, in the iteration number, when varying the window sizes (hr and hs). In addition, it is necessary to point out that the iteration number did not have relationship some with the quality of the segmented image. For example, with small windows can be bigger the number of iterations that with big windows; howev-

er, it was observed that with small window sizes the segmentation was rude. A quantitative comparison of these results can be found in [Rodríguez et. al., 2011b].

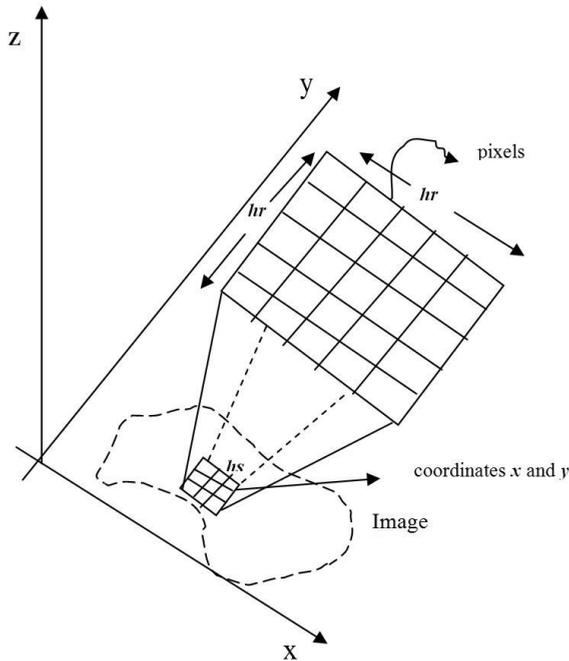


Figure 20. Graphic representation of the radius hr and hs . Movement through an image.

6.1. Some related philosophical issues with image segmentation

Segmentation is recognized to be one of the most important steps in most high-level image analysis systems. Its precise functioning highly determines the performance of the entire system. Image segmentation is today routinely used in a multitude of different applications, such as: medical diagnosis, treatment planning, robotics, pathology, geology, anatomical structure studies, meteorology and computer-integrated surgery, to mention a few. However, image segmentation remains a difficult task due to both the diversity of the object's shape and image's quality. In spite of the most elaborated algorithms developed until now, segmentation remains a very dependent procedure of the application. Until now any single method that can cope with all the problems that can be found does not exist and unfortunately, segmentation remains a complex problem with no exact solution.

For example, of the segmented images those appear in Figure16, which to select as the best segmented image? The answer to this question is not direct, because it depends on the aim of observer. If one wants, for example, a good segmentation of the sky and the earth, the

segmented images of Figure 16 (m), (n) and (o), these would be the most appropriate. However, one can observe in these images that the feet of the astronaut are practically lost. Therefore, if the aim of observer was a good segmentation of the astronaut, these images (Figs. 16 (m), (n)) would not be the best selection. For this aim, the images chosen in Figure 16 would be those (i), (j) or (k). This corroborates that the segmentation is heavily dependent on the application. All this analysis, the reader could carry out it to the segmented images that in Figure 18 appear.

In those segmentation applications where one wants a binary image and the background and objects only should appear; for example, to count regions, the problem could be a little less complicated. An example of this issue is in the count of blood vessels (see Figure 8). Here, one is speaking of segmented images where the final result is a bit/pixel. The problem of segmentation begins to get complicated when the segmented image has several regions, more than 5, 6, 7, 8, 9, bit/pixel, and mainly when one is working with unsupervised segmentation strategies. In the measure that is bigger the number of bit/pixel of the segmented image, the segmentation problem is much more complicated. For example, in the case of supervised segmentation the number of bit/pixel (number of regions) is controlled by the observer and in many practical applications this segmentation method is not very problematic.

On the other hand, one should have in mind the features of the original image. In general, before that the segmentation process is carried out, the original image should be filtered through a low pass filter. In many practical applications this step is very important and many times the final result of segmentation depends on this step. When one uses the iterative algorithm of mean shift this step is implicit, because the mean shift itself is a low pass filter.

The election of one or another segmentation method depends on several factors, namely: a) on the knowledge that the observer has on the method, b) of the application in itself, c) of features of the original image, among others. A universal method of segmentation has not been created, due to that the world of images is practically infinite. For example, the interpretation that a pathologist makes from a biopsy image; which necessarily goes by a segmentation process, it is different to the interpretation that a radiologist makes from a radiological image. It is evident that both are different applications. One could continue analyzing the segmentation issue, but this is an open theme which it will need of many more iterations.

7. Conclusions

In this chapter, we carried out an introduction on theoretical aspects of the mean shift. We also introduced the idea of working with l_∞ norm and we proved that, of this way, one can obtain bigger homogenization degree. We proved the convergence of the mean shift by using the l_∞ norm.

The iterative algorithm that was used in this chapter, where entropy was utilized as a stopping criterion, it was presented too. Through an extensive experimentation by using real and standard images, we showed and we discussed the obtained results with our iterative

algorithm. We proved that our algorithm can be used as an unsupervised suitable strategy to carry out complex problems of segmentation. Some application examples by using our algorithm were shown.

On the other hand, in order to prove the good performance of our algorithm, the same was compared with another segmentation algorithm established already. Through several experiments with real images, we proved that the segmented images by using our iterative algorithm were less noisy than those obtained by means of other methods.

Finally, some related philosophical themes with image segmentation were discussed.

Author details

Roberto Rodríguez Morales^{1*}, Didier Domínguez¹, Esley Torres¹ and Juan H. Sossa²

*Address all correspondence to: rrm@icimaf.cu

1 Institute of Cybernetics, Mathematics & Physics (ICIMAF), Digital Signal Processing Group, Cuba

2 National Polytechnic Institute (IPN), Computing Research Center, Mexico

References

- [1] Cheng, Y. (1995). Mean Shift, Mode Seeking, and Clustering. *IEEE Trans., Pattern Analysis and Machine Intelligence*, 17(8), 790-799.
- [2] Chenyang, X., Dzung, P., & Jerry, P. (2000). Image Segmentation Using Deformable Models. *SPIE Handbook on Medical Imaging, Medical Image Analysis*. Edited by J. M. Fitzpatrick and M. Sonka, III, Chapter 3, 129-174.
- [3] Cheriet, M., Said, J. N., & Suen, C. Y. (1998). A Recursive Thresholding Technique for Image Segmentation. *IEEE Transactions on Image Processing*, 7(6).
- [4] Chin-Hsing, C., Lee, J., Wang, J., & Mao, C. W. (1998, January). Colour image segmentation for bladder cancer diagnosis. *Mathematical and Computer Modelling*, 27(2), 103-120, 0895-7177.
- [5] Christoudias, C. M., Georgescu, B., & Meer, P. (August 2002). Synergism in Low Level Vision. Quebec City, Canada. *16th International Conference on Pattern Recognition*, IV, 150-155.
- [6] Comaniciu, D. I. (2000). Nonparametric Robust Method for Computer Vision. *Ph.D. Thesis*, New Brunswick, Rutgers, The State University of New Jersey.

- [7] Comaniciu, D., & Meer, P. (2002, May). Mean Shift: A Robust Approach toward Feature Space Analysis. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 24(5), 0162-8828.
- [8] Domínguez, D., & Rodríguez, R. (2009). Use of the L (infinity) norm for image segmentation through Mean Shift filtering. *International Journal of Imaging*, 2(S09), 81-93.
- [9] Domínguez, D., & Rodríguez, R. (2011). Convergence of the Mean Shift using L (infinity) norm in image segmentation. *International Journal of Pattern Recognition Research* [1], 32-42.
- [10] Fukunaga, K., & Hostetler, L. D. (1975, January). The Estimation of the Gradient of a Density Function. *IEEE Transactions on Information Theory* [1], IT-21, 32-40, 0018-9448.
- [11] Grenier, T., Revol-Muller, C., Davignon, F., & Gimenez, G. (2006, 8-11, (Oct. 2006)). Hybrid Approach for Multiparametric Mean Shift Filtering, Image Processing. *IEEE, International Conference, Atlanta, GA*, 1541-1544.
- [12] Kenong, W., Gauthier, D., & Levine, M. D. (1995, January). Live Cell Image Segmentation. *IEEE Transactions on Biomedical Engineering*, 42(1), 1-11, 0018-9294.
- [13] Koss, J., Newman, F., Johnson, D., & Kirch, D. (1999, July). Abdominal organ segmentation using texture transforms and a hopfield neural network. *IEEE Transactions Medical Imaging*, 18(7), 0278-0062.
- [14] Otsu, N. (1978). A threshold selection method from grey level histogram. *IEEE Trans. Systems Man Cybernet* [SMC-8], 62-66.
- [15] Rodríguez, R., Alarcón, T., & Sánchez, L. (2001). MADIP: Morphometrical Analysis by Digital Image Processing. Spain. *Proceedings of the IX Spanish Symposium on Pattern Recognition and Image Analysis*, I, 291-298, 84-8021-349-3.
- [16] Rodríguez, R., Alarcón, T., & Pacheco, O. (2005, October). A New Strategy to Obtain Robust Markers for Blood Vessels Segmentation by using the Watersheds Method. *Journal of Computers in Biology and Medicine*, 35(8), 665-686, 0010-4825.
- [17] Rodríguez, R., & Suarez, A. G. (2006). An Image Segmentation Algorithm Using Iteratively the Mean Shift. *Book Series Lecture Notes in Computer Science*, Publisher Springer, Berlin/Heidelberg, 4225, *Book Progress in Pattern Recognition, Image Analysis and Applications*, 326-335.
- [18] Rodríguez, R., & Pacheco, O. (2007). A Strategy for Atherosclerosis Image Segmentation by using Robust Markers. *Journal Intelligent & Robotic System*, 50(2), 121-140.
- [19] Rodriguez, R. (2008). Binarization of medical images based on the recursive application of mean shift filtering: Another algorithm. *Journal of Advanced and Applications in Bioinformatics and Chemistry*, I, 1-12, Dove Medical Press Ltd.
- [20] Rodríguez, R., & Tovar, R. (2010). An Unsupervised Strategy for Bio-medical Image Segmentation. *Journal of Advanced and Applications in Bioinformatics and Chemistry* [3], 67-73.

- [21] Rodriguez, R., Suarez, A. G., & Sossa, J. H. (2011a). A Segmentation Algorithm based on an Iterative Computation of the Mean Shift Filtering. *Journal Intelligent & Robotic System*, 63(3-4), 447-463.
- [22] Rodriguez, R., Torres, E., & Sossa, J. H. (2011b). Image Segmentation based on an Iterative Computation of the Mean Shift Filtering for different values of window sizes. *International Journal of Imaging and Robotics*, 6(A11), 1-19.
- [23] Rodríguez, R., Torres, E., & Sossa, J. H. (2012). Image Segmentation via an Iterative Algorithm of the Mean Shift Filtering for Different Values of the Stopping Threshold. *International Journal of Imaging and Robotics*, 7(1).
- [24] Salvatelli, A., Caropresi, J., Delrieux, C., Izaguirre, M. F., & Caso, V. (2007). Cellular Outline Segmentation using Fractal Estimators. *Journal of Computer Science and Technology* ., 7(1), 14-22.
- [25] Shen, C., & Brooks, M. J. (2007, May). Fast Global Kernel Density Mode Seeking: Applications to Localization and Tracking. *IEEE Transactions on Image Processing*, 16(5), 1457-1469, 1057-7149.
- [26] Suyash P., Awate, & Ross T., Whitaker. (2006, March). Higher-Order Image Statistics for Unsupervised, Information-Theoretic, Adaptive, Image Filtering. *IEEE Trans. Pattern Analysis and Machine Intelligence (PAMI)*, 364-376, 28(3).
- [27] Vicent, L., & Soille, P. (1991, June). Watersheds in digital spaces: An efficient algorithm based on immersion simulations. *IEEE Transactins on Pattern Analysis and Machine Intelligence*, 13(6), 583-593, 0162-8828.
- [28] Zhang, H., Fritts, J. E., & Goldman, S. A. (2003). Paper presented at Proceeding of The SPIE. *A Entropy-based Objective Evaluation Method for Image Segmentation, Storage and Retrieval Methods and Applications for Multimedia 2004*. Edited by Yeung, Minerva; Lienhart, M., Rainer W.; Li, Choung-Sheng, 5307, 38-49.