

Full Length Research Paper

A new algorithm for image segmentation by using iteratively the mean shift filtering

Roberto Rodríguez* and Ana G. Suarez

Institute of Cybernetics, Mathematics and Physics (ICIMAF), Digital Signal Processing Group Calle 15 No. 551 e/ C y D
CP 10400, La Habana, CUBA

Accepted 6 November, 2006

Image segmentation plays an important role in many systems of computer vision. The good performance of recognition algorithms depend on the quality of segmented image. According to the opinion of many authors the segmentation concludes when it satisfies the observer's objectives, the more effective methods being the iterative. However, a problem of these algorithms is the stopping criterion. In this work the entropy is used as stopping criterion in the segmentation process by using recursively the mean shift filtering. In such sense a new algorithm is introduced. The good performance of this algorithm is illustrated with extensive experimental results. The obtained results demonstrated that this algorithm is a straightforward extension of the filtering process. In this paper a comparison was carried out between the obtained results with our algorithm and with the *EDISON* System.

Key words: Entropy, image segmentation, mean shift, smoothing filter.

INTRODUCTION

Segmentation and contour extraction are important steps in many systems of high level. Segmented images are now used routinely in a multitude of different applications, such as, diagnosis, treatment planning, in the robotics, localization of pathology, geology, study of anatomical structure, meteorology, computer-integrated surgery, among others. However, image segmentation remains a difficult task due to both the variability of object shapes and the variation in image quality. In spite of the most complex algorithms developed until the present, segmentation continues being very dependent on the application and it does not exist on a single method that can solve all the problems that are presented in the universe. With the aim of obtaining segmentation methods more exact and more effective, several techniques have been proposed in the literature, where a great variety of them has been dedicated to biomedical images (Kenong et al., 1995; Sijbers, 1997; Chin-Hsing et al., 1998; Rodríguez, 2002; Schmid, 1999; Koss et al., 1999; Braiek, 2005). Unfortunately, segmentation is a

complex problem with no exact solution. Segmentation using traditional low-level image processing techniques, such as thresholding, histogram, 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. Furthermore, noise and other image artifacts can cause incorrect regions or boundary discontinuities in objects recovered from these methods.

At the present time the most robust algorithms in segmentation are the iterative methods, which cover a variety of techniques, from the mathematical morphology, deformable models until the thresholding methods. However, one of the problems of these iterative techniques is the stopping criterion, in which great quantities of methods have been proposed (Comaniciu and Meer, 2002; Chenyang et al., 2000; Vicent and Soille, 1991; Cheriet et al., 1998).

The mean shift is a nonparametric procedure and it is an extremely versatile tool for feature analysis and can provide reliable solutions for many computer vision tasks (Comaniciu, 2000). The mean shift was proposed in 1975 by Fukunaga and Hostetler (Fukunaga and Hostet-

*Corresponding Author's E-Mail: rrm@icmf.inf.cu.

ler, 1975) and largely forgotten until Cheng’s paper (Cheng, 1995) rekindled interest in it. The segmentation by using the mean shift as first step carries out a smoothing filter and later on the segmentation process (Comaniciu, 2000).

The term of entropy is not a new concept in the field based on information theory and it has been used in image restoration, edge detection and recently as an objective evaluation method for image segmentation (Zhang et al., 2003).

In this work a new segmentation strategy by using the mean shift is proposed. The new method uses the entropy as stopping criterion, where two steps to obtain the segmented image is not necessary to carry out. The obtained results with our algorithm are compared with the attained results by using EDISON System (Robust Image Understanding Laboratory, Rutgers University, www.caip.rutgers.edu/riul/). The results from this preliminary study indicate that the proposed strategy is effective and that at worst these are similar to those reported in (Comaniciu and Meer, 2002; Comaniciu, 2000).

The mean shift analysis

The iterative procedure of the mean shift is introduced as normalized density gradient estimate. 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 and Hostetler 1975) to guarantee asymptotic unbiasedness, mean-square consistency, and uniform consistency of the estimate in the expression (1)

For example, for *Epanechnikov* kernel

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

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 the volume $h^d c_d$, centered on 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,

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

The 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.

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 and 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 with the h_s and h_r parameters. 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,

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

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.

The novelty lies in applying the mean shift procedure for the data points in the joint spatial-range domain.

ENTROPY

From the point of view of digital image processing the entropy is defined according to the following expression,

$$E(x) = - \sum_{x=0}^{2^B-1} p(x) \log_2 p(x) \quad (6)$$

where B is the total quantity of bits of the digitized image and by agreement $\log_2 0 = 0$; $p(x)$ it is the probability of occurrence of a gray-level value. Within a totally uniform region the entropy reaches the minimum value, because one speaking theoretically, the probability of occurrence of the gray-level value is always one. In the practice when one works with real images, the entropy does not reach, in general, the zero value. This is due to the existent noise in the images. Therefore, if we consider the entropy as a measure of the disorder within a system, it could be used as a good stopping criterion into an iterative process by using the mean shift filtering. The entropy within each region diminishes in the measure that the regions become more homogeneous and at the same time in the whole image, until reaching a stable value. When the convergence is reached, a totally segmented image is obtained, because the mean shift filtering is not idempotent as it does not happen to some types of filters in the mathematical morphology (for example, with the opening). In addition, as by (Comaniciu and Meer, 2002) was pointed out, the mean shift procedure-based image segmentation 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. First, it is known that the addition of two independent random variables (for example, a signal and additive noise) increases the entropy (Suyash and Whitaker, 2006). Entropy reduction reduces the randomness in corrupted probability density function and tries to counteract noise. Then, following this same analysis as the segmented image is a simplified version of the original image, the entropy (segmented image) should be smaller. Recently, it was found empirically that the entropy of the noise diminishes faster than that of the signal (Suyash and Whitaker, 2006). Therefore, an effective criterion would be to stop when the relative rate of change of the entropy of iteration to the next one, falls below some threshold.

ALGORITHMS

Filtering algorithm by using the mean shift

Let X_i and Z_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 $z \in [0, 2^\beta - 1]$, β being the quantity of bits/pixel in the image. The filtering algorithm comprises the following steps (Comaniciu and Meer, 2002):

1. Initialize $j=1$ and $y_{i,1} = p_i$
2. Compute through the mean shift (see expression (3), $y_{i,j+1}$), 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.

Developed segmentation algorithm by using recursively the mean shift filtering

The proposed algorithm comprises the following steps

1. Initialize the control values, **ent1 = 1**, **errabs = 1**, **edsEnt**
2. While **errabs > edsEnt**, then
3. Do filtrate the image according to the steps of the previous algorithm; store in **Z** the filtering image.
4. Do calculate the entropy from the filtered image according to the expression (6); store in **ent**
5. Do calculate the absolute error with the entropy value obtained in the previous step; **errabs = abs(ent - ent1)**
6. Do update the value of the parameter; **ent1 = ent**

It is possible to observe that, in this case, the proposed segmentation algorithm is a straightforward extension of the filtering algorithm, which finishes when the entropy reaches the stability. Note the simplification of this algorithm compared with the one proposed in (Comaniciu and Meer, 2002). A detailed discussion on this issue will be made in the next section.

DISCUSSION

Image segmentation; that is, the decomposition of the gray-level values into homogeneous areas is maybe one of the most important steps in any system of computer vision. Homogeneity, in general, is defined as similarity in pixel values, where a piecewise constant model is enforced over the image (Comaniciu and Meer, 2002).

All the segmentation experiments were performed using a uniform kernel. The segmentation of the image lake obtained with the proposed algorithm is shown in Figure. 1b, while in Figure. 1c is presented the obtained results by using EDISON System (Robust Image Under-

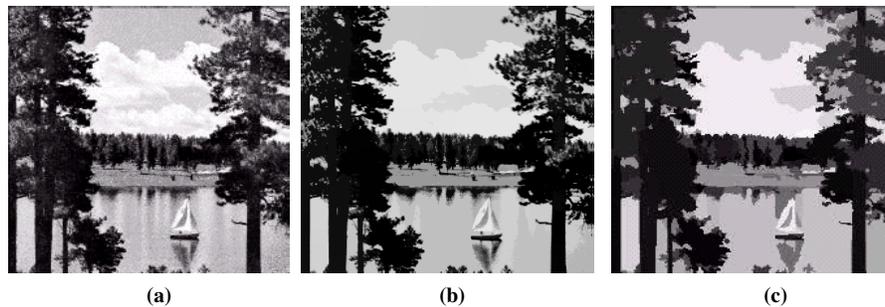


Figure 1. Lake: (a) Original image, (b) Segmented image by our strategy $(hs, hr) = (12, 15)$, (c) Segmented image according to algorithm proposed in (Comaniciu and Meer, 2002), $(hs, hr, M) = (16, 7, 40)$.

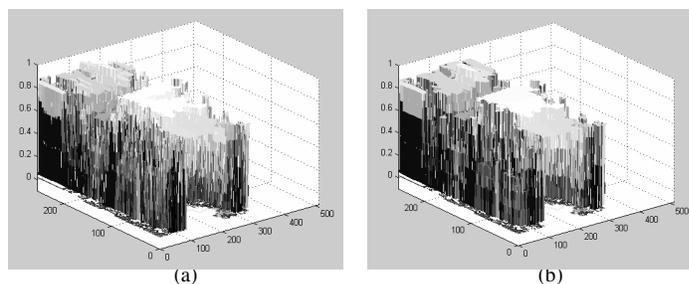


Figure 2. Visualization in 3D of the mean shift segmentation. (a) With our algorithm. (b) By using *EDISON* system.

standing Laboratory, Rutgers University, www.caip.rutgers.edu/riul/.

From the point of view of the visual analysis, it is not observed big differences between the image segmented by our algorithm and the one obtained by the *EDISON* System. However, to say of other observers the sky and the clouds were better segmented with our strategy. Note that all the other details in the segmented image are preserved. Comparing Figures 1b and 1c, in the part corresponding to the water with our algorithm an additional homogeneous area was obtained. The obtained result in Figure 1b took 9 iterations to reach the convergence; that is, until the stability was reached (no more change). From the point of view of the final result the image segmented with our algorithm has an aspect a little more natural with regard to the original image. In many occasions, given the application, segmentation imposes certain conditions (elimination of regions, prunes or integration of certain maxima, etc). This can originate a biased image with regard to the initial. With our algorithm the resolution is only imposed on the segmentation process; that is, the parameters hr and hs . For this reason, our algorithm did not make mistakes; that is, a segmented image very different to the original never was obtained.

To better visualize the segmentation process, it is represented in three dimensions in Figure 2. Observe that the

data was reflected over the horizontal axis, while the intensities over z . This gives a more informative display. In Figure 2a the segmented image with our algorithm is presented, while in Figure 2b the segmented image with the *EDISON* system. Note that in both representations significant differences are not observed, but in the segmented image with our algorithm one can see other homogeneous areas (do see the central plate of Figure 2a corresponding to the sky and the clouds). It is possible to observe in Figure 2a the integration of large homogeneous regions when the convergence is reached (see the plates). Also, one can see that these plates in both images reached the same level of intensity (see in axis z). In spite of with both algorithms very similar results are obtained, the difference underlies in that to obtain the image of Figure 1c it was necessary to carry out a filtering step and other of segmentation. In this last step, one can have certain complexity when adjacency graphs and hierarchical technique are used (Comaniciu and Meer, 2002).

We verified that the segmentation through our algorithm was not very sensitive to the choice of the parameters hr and hs . This was also tested in (Comaniciu and Meer, 2002). We selected for most of images the same parameters $hr = 15$ and $hs = 12$ and ***edsEnt* = 0.005** as threshold of stopping criterion. This value was empirically found, after carrying out several investigations



Figure 3. Cameraman: (a) Original image, (b) Segmented image according to our algorithm $hs=12$ and $hr=15$, (c) Segmented image by using *EDISON* system $hs=8$, $hr=4$.

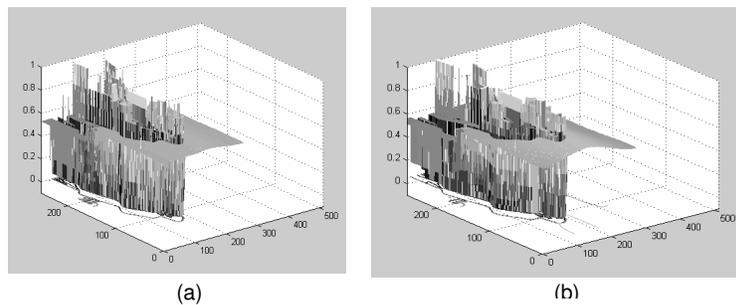


Figure 4. Visualization in 3D of the mean shift segmentation. (a) With our algorithm. (b) By using *EDISON* system.

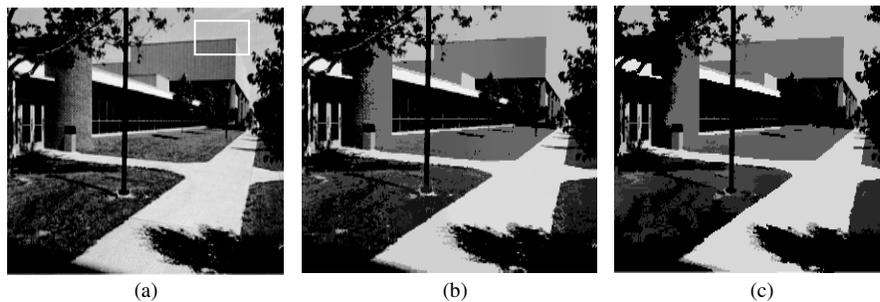


Figure 5. (a) Original image, (b) Segmented image according to our algorithm $hs=12$ and $hr=15$, (c) Segmented image using the *EDISON* system; $hs=8$, $hr=7$ y $M=20$.

with different images.

Another segmentation example through our algorithm and using the *EDISON* system is shown in Figure 3. Note that visually substantial differences are not appreciated between the segmented image with our algorithm and by using *EDISON* system. The convergence with our algorithm in 8 iterations was reached. It is possible to observe that with our strategy the grass is a little more homogeneous as compared with the image of Figure 3c. In addition, all the other regions were preserved the same as in the image of Figure 3c. Note that the tripod and the buildings were well segmented. However, with

our algorithm is not well appreciated the area of the cameraman's hand for an $hr=15$. In other images segmented with smaller value of hr was possible to segment the cameraman's hand. This was not put here for space problem. The range parameter hr controls the number of regions in the segmented image (Comaniciu and Meer, 2002). A three-dimensional representation of the segmented images with both methods is shown in Figure 4.

Note in Figure 4 that in both images is appreciated the plates corresponding to the sky and to the grass, with the slight difference that in the image of Figure 4b the plate of the grass is observed something curved. In the other

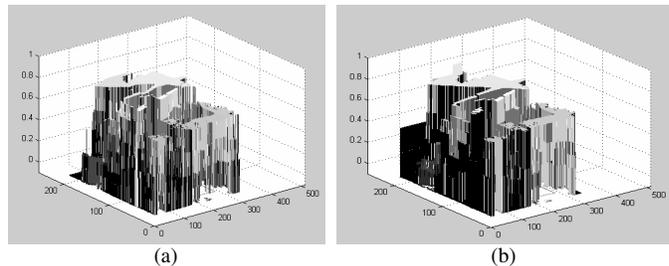


Figure 6. 3D visualization of the mean shift segmentation. (a) With our algorithm. (b) By using *EDISON* system.

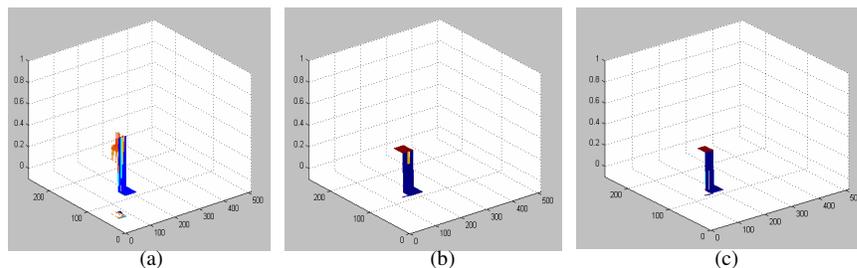


Figure 7. Visualization in 3D. (a) Input, (b) Our algorithm, (c) By using *EDISON* system.

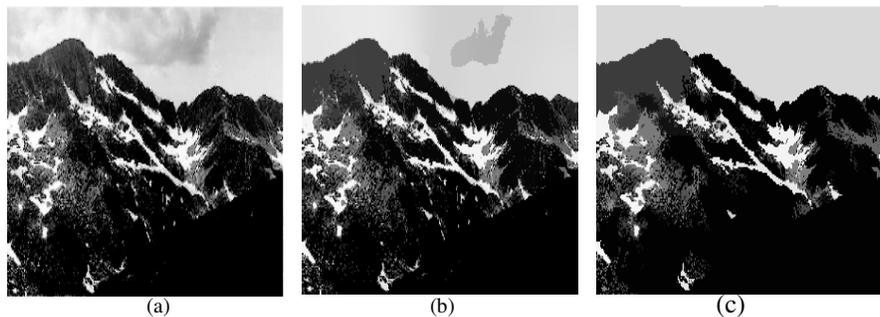


Figure 8. (a) Original image, (b) Segmented image according to our algorithm ($hs=12$, $hr=15$), (c) Segmented image by using *EDISON* system ($hs=8$, $hr=7$ and $M=20$).

details, significant differences are not appreciated. The advantage that presents our algorithm compared with the one proposed in (Comaniciu and Meer, 2002) is the simplicity of the same one. Our algorithm would be very useful in an automatic segmentation process.

In Figure 5 another segmentation examples are shown with both methods. Analyzing, both images visually, one can see that these were segmented well, where it is possible to observe differences in some details, which it are not significant. These differences can be given by the use of different parameters (hr , hs). In this case the convergence with our algorithm was reached in 10 iterations. To better visualization of the segmentation process in Figure 6 are represented in three dimensions. Note that in both images the plates corresponding to the

most homogeneous areas are appreciated with the same levels of intensities (in the axis z). It is important to point out that the same as what happened in the segmented images with the *EDISON* system, with those obtained with our algorithm any maximum was not obtained after the convergence was reached.

The abrupt changes that are appreciated in the fall or elevation from a plate to another in the 3D representation are indicative of different regions with different intensities. These correspond to the discontinuities (edges). For example, the window marked in Figure 5a is represented in three dimensions in Figure 7b. It is possible to see in this representation two homogeneous plates corresponding to the sky and the roof of the school. Note an abrupt change between a zone and the other one. This means

two totally different regions

According to the opinion of some observers with our algorithm the segmented images have a more natural aspect. An example is the result that appears in Figure 8. Note that the clouds and the sky were better isolated with our algorithm. It is necessary to point out that in this comparison an analysis from the point of view of the computation time was not carried out, due to that our algorithm was implemented in *MatLab* (ver. 6. 5), while the *EDISON* system was programmed in *C++*.

Conclusion

In this work was proposed a segmentation algorithm using an iterative process of the mean shift filtering. It was tested that the algorithm always converges and through the 3D representation was possible to demonstrate that this convergence always addressed to the homogeneous regions; that is, the different objects. It was verified through the obtained results that our algorithm was effective and that at worst, these (results) were similar to those reported in (Comaniciu and Meer, 2002; Comaniciu, 2000). From the point of view of the algorithmic complexity this strategy is simpler than the proposal in (Comaniciu and Meer, 2002), since adjacency graphs and hierarchical technique is not necessary to use. It was evidenced that this segmentation procedure is a straightforward extension of the filtering algorithm. For this reason, our algorithm did not make mistakes; that is, a segmented image very different to the original never was obtained. This strategy will be extended to the color image segmentation. The results from this preliminary study indicated that the proposed strategy was effective.

REFERENCES

- C Chin-Hsing, J Lee, J Wang, CW Mao (1998). "Color image segmentation for bladder cancer diagnosis", *Math. Comput. Model.* vol. 27(2): 103-120.
- Cheng Y (1995). Mean Shift, Mode Seeking, and Clustering. *IEEE Trans. Pattern Anal. Mach. Intell.* Vol. 17(8): 790-799.
- Comaniciu D, Meer P (2002). "Mean Shift: A Robust Approach toward Feature Space Analysis", *IEEE Transaction on Pattern Analysis and Machine Intell.* Vol. 24 (5): 1-18.
- Comaniciu DI (2000). *Nonparametric Robust Method for Computer Vision*, Ph.D. thesis, New Brunswick, Rutgers, The State University of New Jersey. pp. 38-39.
- EB Braiek, M Cheriet, V Dore (2005). SKCS-A Separable Kernel Family with Compact Support to improve visual segmentation of handwritten data. *Electronic Lett. Comp. Vision and Image Anal.* 5(1):14-29.
- Fukunaga K, Hostetler LD (1975). The Estimation of the Gradient of a Density Function. *IEEE Trans. Inform. Theor.* Vol. 21: 32-40.
- H Zhang, JE, Fritts SA Goldman (2003). A Entropy-based Objective Evaluation Method for Image Segmentation. *Storage and Retrieval Methods and Applications for Multimedia 2004*. Edited by Yeung, Minerva M, Lienhart Rainer W, Li Choung-Sheng. *Proceeding of The SPIE* Vol. 5307: 38-49.
- J Sijbers, P Scheunders, M Verhoye, A Van der Linden, D Van Dyck, E Raman (1997). Watershed-based segmentation of 3D MR data for volume quantization. *Magnet. Reson. Imag.* vol. 15 (6): 679-688.
- JE Koss, FD Newman, TK Johnson, DL Kirch (1999). "Abdominal organ segmentation using texture transforms and a hopfield neural network. *IEEE Trans. Med. Imag.*, vol. 18(7): 50-62.
- M Cheriet, JN Said, CY Suen (1998). A Recursive Thresholding Technique for Image Segmentation. *IEEE Trans. Image Process.* Vol. 7 (6): 918-921.
- P Schmid (1999). Segmentation of digitized dermatoscopic images by two-dimensional color clustering. *IEEE Trans. Med. Imag.*, vol. 18, no. 2, Feb.
- Rodríguez R, Alarcón T, Wong R, Sanchez L (2002). Color segmentation applied to study of the angiogenesis. Part I. *J. Intell. Robot. Sys.* Vol. 34(1): 83-97.
- Suyash P, Awate, Ross T (2006). Whitaker: Higher-Order Image Statistics for Unsupervised, Information-Theoretic, Adaptive, Image Filtering. *IEEE Trans. Pattern Anal. Mach. Intell. (PAMI)* Vol. 28(3): 364-376.
- Vicent L, Soille P (1991). "Watersheds in digital spaces: An efficient algorithm based on immersion simulations", *IEEE Transact. Pattern Anal. Machine Intell.* 13: 583-593.
- W Kenong, D Gauthier, MD Levine (1995). Live Cell Image Segmentation. *IEEE Trans. Biomed. Eng.* vol. 42 (1): 1-11.
- X Chenyang, P Dzung, P Jerry (2000). Image Segmentation Using Deformable Models, *SPIE Handbook on Medical Imaging, Medical Image Analysis*, Edited by JM Fitzpatrick and M Sonka, Vol. III, Chapter 3, pp. 129- 174.