

Statistical Hypothesis Testing and Wavelet Features for Region Segmentation

David Menoti¹, DÍbio Leandro Borges², and Arnaldo de Albuquerque Araújo¹

¹ UFMG - Universidade Federal de Minas Gerais

Grupo de Processamento Digital de Imagens - Departamento de Ciência da Computação

Av. Antônio Carlos, 6627, Pampulha - 31.270-010, Belo Horizonte-MG, Brazil

{menoti,arnaldo}@dcc.ufmg.br

² BIOSOLO, Goiânia - Go, Brazil

dibio@terra.com.br

Abstract. This paper introduces a novel approach for region segmentation. In order to represent the regions, we devise and test new features based on low and high frequency wavelet coefficients which allow to capture and judge regions using changes in brightness and texture. A fusion process through statistical hypothesis testing among regions is established in order to obtain the final segmentation. The proposed local features are extracted from image data driven by global statistical information. Preliminary experiments show that the approach can segment both texturized and regions cluttered with edges, demonstrating promising results. Hypothesis testing is shown to be effective in grouping even small patches in the process.

1 Introduction

Segmentation process is a major bottleneck in applications on Pattern Recognition and Computer Vision areas, and it is kept on the agenda of scientific community. There are several segmentation approaches on literature, however since most of the tests are designed for specific applications, many open problems remain. One which is addressed here in this paper is how to segment regions consistently having images either with highly texturized patches, or artefacts with brightness, or both. Usually the approaches work either for one situation, or another, and in this paper we propose an approach that could work more consistently in both situations.

Segmentation approaches found in the literature could be separated into two main groups: those regarding supervised learning, which take into account knowledge about the application (i.e., training); and those ones regarding unsupervised learning, which relies only in the input image data, or without *a priori* knowledge about the application.

The well known Canny edge detector [1] models boundaries as brightness step edges, which is the most common approach to detect local boundaries. The Canny detector fails wildly inside texturized regions where high contrast edges are present, but usually with no separation between regions. Moreover, it is unable to detect the boundary between textured regions when there is only a subtle change in average image brightness.

A review on image segmentation approaches of the 80's and previous can be found in [2], and an initial guide can be found in [3]. Recently, with availability of computation and memory store devices, more demanding and complex approaches have been proposed.

A highly cited approach was proposed by Shi *et al.* in [4]. The approach extracts the global impression of an image, considering the segmentation task as a graph partitioning problem proposing a global criterion, the normalized cut for segmenting the graph. This criterion can be optimized by an efficient computational technique ($O(N^{3/2})$) based on a generalized eigenvalue problem. In [5], Sharon *et al.* proposed a fast algorithm for image segmentation on multiscale framework based on graph partitioning, and it has a linear time complexity ($O(N)$) in number of pixels presenting results that are at least comparable to the results obtained by the spectral methods [4]. The algorithm is inspired on algebraic multigrid (AMG) solvers of minimization problems of heat and electric networks. In this approach, more measurements were combined in a multiscale framework. The results obtained for some images were better than the ones in [4].

In this paper, we propose a new segmentation algorithm focusing on local features and their consistencies in image data (i.e., an unsupervised approach), which is driven by global statistical information. Patches will be extracted from low and high frequency wavelet coefficients, and new features are devised based on those ones. Brightness and texture are modelled through patch images (e.g. square windows) from these extracted features. Statistic hypothesis testing is then proposed to perform the segmentation, using a fusion process in order to generate more consistent regions according to a control parameter given by the user.

The rest of this paper is organized as follows. In Section 2, the proposed approach is described. The experiments performed for testing and evaluating the algorithm are shown in Section 3. Finally, conclusions and future works for extending the approach in a multi-scale and multiresolution framework are pointed out in Section 4.

2 The Approach

Our proposed approach can be divided into 3 main steps, namely: 1) First, feature extraction by a Wavelet transform is performed. The image I is decomposed into a wavelet space, using Mallat's decomposition algorithm [6] with Haar basis function. This is a non-redundant transformation which leads to four output channels of features, being LL , LH , HL , and HH ; 2) New features are devised based on the wavelet coefficients. The output channels of the Wavelet transform are exploited in a windowed (e.g. patches of $k \times k$ elements in size) way in order to generate new features for characterizing brightness and texturized regions; 3) Region growing through statistical hypothesis testing is performed based on information extracted from windowed features, generating consistent region segmentation. A flowchart of the proposed approach can be seen in Figure 1. Details of each step are given in the following subsections.

2.1 Wavelet Transform

A Wavelet transform decomposes data into fundamental building blocks. Its basic difference from Fourier decomposition is that the wavelet functions are well localized in

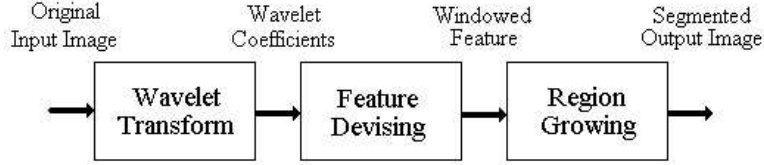


Fig. 1. Flowchart of the proposed approach.

time and space, whereas sinusoidal functions used in Fourier transform are not. Since it is possible to design wavelet decompositions with a great variety of basis functions, and also either emphasizing redundancy or eliminating it throughout the levels of decomposition, the literature for such is plenty. Moreover, the Wavelet transform has been used in several fields in Image Processing, Pattern Recognition, and Computer Vision for image denoising and coding, object segmentation, recognition, and characterization.

Here, for our purposes of wavelet feature extraction, a desired decomposition would have to help representing consistently both brightness and texturized regions, eliminating redundancies on scales. We use the well known Mallat's decomposition algorithm [6] with Haar (simpler) basis function (QMF filter pairs: $\tilde{H} = [\sqrt{2}/2; \sqrt{2}/2]$; $\tilde{G} = [\sqrt{2}/2; -\sqrt{2}/2]$) for performing the wavelet decomposition, namely Wavelet transform. Given an input image (with $n \times m$ in size), the result of this decomposition process are four output channels of low and high frequency content wavelet coefficients (with $n/2 \times m/2$ elements in size, since it is a downsampling process), e.g. LL , LH , HL , HH . LL is called approximation (low frequency) band, LH , HL , and HH are called horizontal, vertical, and diagonal details (high frequency) bands, respectively. Our first stage then consists of transforming an input image I into four output channels, e.g. LL , LH , HL , and HH .

2.2 Feature Devising

We propose to analyze a region considering what we call a degree of perturbation. In an approximate or complete homogeneous region, changes in brightness are very few or none. On the other hand, in a texturized region it is possible to notice an almost uniform degree of perturbation, or confusion, throughout that region. Therefore, in order to characterize a consistent region we will use this concept of degree of perturbation, namely a consistent region should have a homogeneous degree of perturbation. Thus, different, brightness or texturized, regions will have different degrees of perturbation.

For achieving such an aim, we shall have features allowing us to capture these characteristics from the regions data. Original patches of the image (e.g. square windows) could be the source of these local properties. In this step we devise new features based on the wavelet coefficients, extracted from the latter step. They will be able to represent and capture the degree of perturbation of the patches. In fact, we take square windows of $k \times k$ elements and extract statistics from those, which will be used in the next step for segmentation in a region growing process. Then, the resulting output of this step are

statistical matrix images with $n/2k \times m/2k$ elements, having the mean (μ) and variance (σ^2) of the patch windows.

From low and high frequency wavelet coefficients (only first level of decomposition is used) we extract enough information for characterizing consistent regions. Then, we first separate the wavelet coefficients in two kinds of features, e.g.,

$$I\mu_{Low} = \mu(|LL|) = \mu(LL), I\sigma_{Low}^2 = \sigma^2(|LL|) = \sigma^2(LL). \quad (1)$$

As the Haar basis function was used in the Wavelet transform, it produces only positive values for LL . Thus, we take them as absolute values. And for high frequency features we have,

$$I\mu_{High} = \mu(\sqrt{(LH)^2 + (HL)^2}), I\sigma_{High}^2 = \sigma^2(\sqrt{(LH)^2 + (HL)^2}), \quad (2)$$

HH coefficients are noisy, and usually not reliable for these purposes, and so they are left out.

Those are the proposed new features for the segmentation task, i.e., $I\mu_{Low}$, $I\mu_{High}$, $I\sigma_{Low}^2$, and $I\sigma_{High}^2$. In the next step, those measures extracted from the patches are used as cues for the segmentation. Experiments given here show that those features are enough for characterizing brightness and texturized regions.

2.3 Region Growing through Statistical Hypothesis Testing

The analysis proposed on mean and variance values makes an important assumption, namely that the features upon which segmentation is based is distributed normally (Gaussian distribution) [7]. The parameter σ^2 is called the variance (i.e., σ is the standard deviation), and it measures the flatness of the distribution. Discrimination between adjacent areas with differing means and standard deviations can be made according to Fischer's criterion [8]:

$$\frac{|\mu_1 - \mu_2|}{\sqrt{\sigma_1^2 + \sigma_2^2}} > \lambda, \quad (3)$$

where λ is a threshold, μ_1 , μ_2 , σ_1 , and σ_2 are averages and variances of respective regions.

In other words, if two regions have good separation in their means, and low variance, then it is possible to discriminate them. However, if the variance becomes high and the mean difference is low it is not possible to separate them.

In order to have a self tuned algorithm, we establish a measure for this purpose. For a better separation, the merging threshold, λ , for the mean intensity for two adjacent regions should be adjusted depending on the expected uniformity of the merged region. Less uniform regions will require a lower threshold to prevent under merging. The uniformity is a function of both intensity mean and variance of the region. A suitable heuristic law for combining both properties into one is [7]:

$$Uniformity = 1 - \sigma^2/\mu^2, \quad (4)$$

where μ and σ^2 are related with the full matrix feature image. Note that the uniformity will be in the range of 0 to 1 for cases where the samples are all positive. The threshold value, λ decreases with the decrease in uniformity as follows:

$$\lambda = (1 - \sigma^2/\mu^2)\lambda_0. \quad (5)$$

This has the advantage that the user need only to supply a single threshold, λ_0 . But for our algorithm, we use two features, and so discrimination functions are

$$\frac{|\mu_{Low1} - \mu_{Low2}|}{\sqrt{\sigma_{Low1}^2 + \sigma_{Low2}^2}} \geq \lambda_{Low}, \quad (6)$$

$$\frac{|\mu_{High1} - \mu_{High2}|}{\sqrt{\sigma_{High1}^2 + \sigma_{High2}^2}} \geq \lambda_{High}, \quad (7)$$

Note that we still need to supply only a single threshold, λ_0 for having λ_{Low} and λ_{High} , since they are dependent on the data by Equation 5. Therefore, for two regions 1 and 2 to be discriminated, they must hold either one of Equations 6 or 7. However, for them to be merged, they can not hold both Equations 6 and 7.

The algorithm is automatically started with chosen seed windows, those ones with more energy on low frequency wavelets coefficients. These seeds are enqueued, and then the region process is started dequeuing window by window. Each dequeued window (i.e., region) is tested with its neighboring windows, merging or generating new regions and putting them on a queue. At this moment the mean and the variance values are updated. The algorithm will evolve until there are no more windows left (i.e., regions) on queue.

3 Experiments

For the experiments, we have used three known images having texturized and brightness homogeneous regions. They are shown in Figure 2. These images are grayscale and 512×512 pixels in size. Parenthesized numbers on captions in Figure 3 and 4 are the percentual significance level in the Gaussian distribution used in statistical hypothesis testing. Resulting images of our proposed algorithm have granularity $2k$, namely each patch image corresponds to $2k \times 2k$ pixels from the original image, where k is the window size and the number 2 comes from the decimation process of the Wavelet transform, i.e., downsampling. Each of the segmented regions on the images in Figures 3 and 4 was fulfilled with its respective average grayscale.

We have performed experiments using different values for λ_0 , e.g., 1.28(80%), 1.64(90%), 1.96(95%), in order to verify the algorithm behavior and its extension for other scales. Results are shown in Figure 3, and input images used are those ones in Figure 2. Observing those images, we can notice that segmented regions are more consistent as the λ_0 value increases, namely the images are less oversegmentated. On the other hand, some small regions, which could be dependent on the application or type of scene, are merged together.

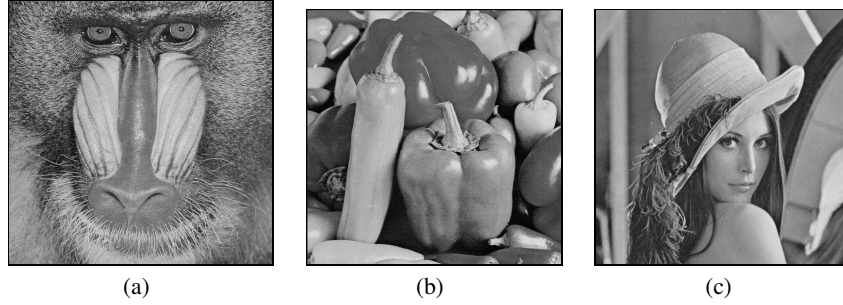


Fig. 2. Images (input) used in our experiments.

Figure 4 shows results from different windows size. For those images we have setup $\lambda_0 = 1.96$, which stands for significance level of 90% in the Gaussian distribution. We have tested two window size, 2, and 4. The images in Figure 4 give an idea of how usefull it could be to combine segmentation of several different windows size in order to obtain a more consistent segmentation.

Our algorithm implementation (C++) took less than 5 seconds for segmenting any of those images of 512×512 pixels in size in a Pentium III 1.0 GHz with 512 MB RAM. The time complexity of our proposed algorithm is $O(N)$ considering the number of windows, since region of windows will go to queue only with a new incorporated window, and so, the algorithm will process each window at a maximum of 8 times, i.e., number of neighbors.

4 Conclusions and Future Works

In this paper, we have proposed and tested a new approach for segmenting images into consistent regions. As the experiments show the approach has some advantages, which are to address region segmentation both in texturized regions and regions with brightness, using the same features. Even though one can observe oversegmentation, as for example in images in Figure 4 for small window size (i.e., $k = 2$), it is reasonable to say that these images were segmented in consistent regions, since we do not take into account any *a priori* knowledge or high level perception cue. Thus, the features devised by wavelet coefficients were shown to be a good characterization for both brightness and texturized regions. Our given experiments show interesting results that we think are promising for further investigation. For future works we want to address fusion among scales in order to obtain fine segmentation results, e.g., one pixel wide granularity. Also, a possibility would be to have different window sizes in different scales putting all together in a multi-scale and multiresolution approach. A more extensive comparison with other algorithms, such as the ones in [4], [5], would be necessary in order to confirm our hypothesis that algorithms based on graph cuts perform segmentation without any knowledge about the objects. We think this perception should be driven by other decision process, like intuition for example. Our proposed algorithm

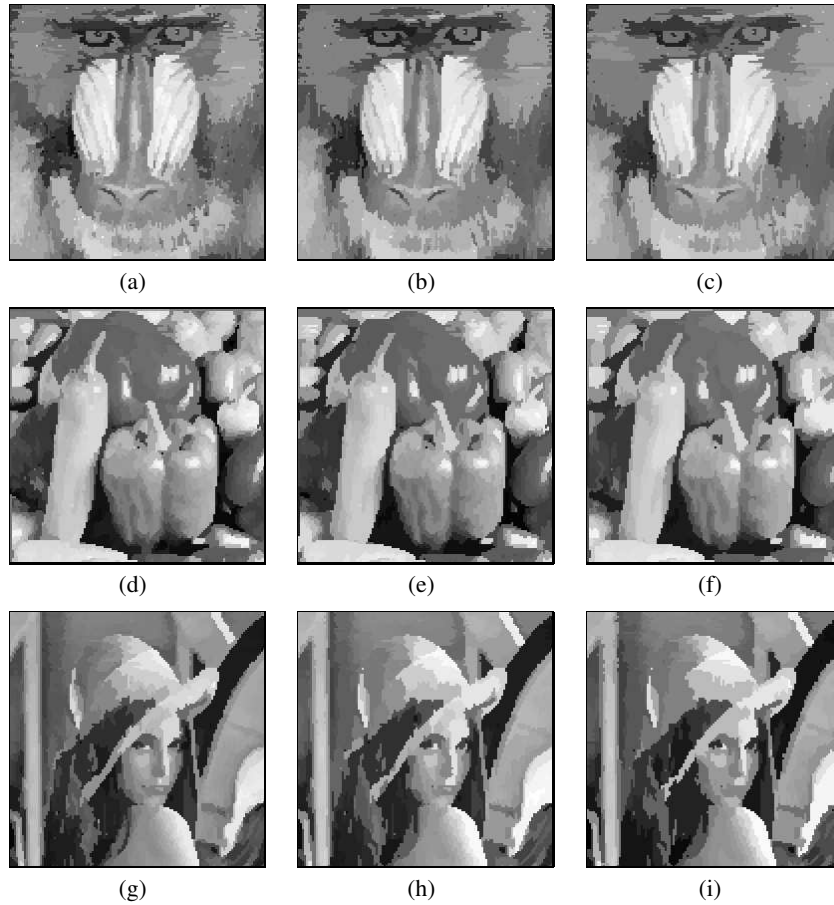


Fig. 3. Experiments used for verifying our proposed approach; using different values of λ_0 for input images shown in Figure 2. In first col we have (3(a), 3(d), and 3(g)) $\lambda_0 = 1.28$ (80%), second col (3(b), 3(e), and 3(h)) $\lambda_0 = 1.64$ (90%), and, in third col (3(c), 3(f), and 3(i)) $\lambda_0 = 1.96$ (95%).

can also be applied for Natural, Biomedical, SARs, and man-made structures image segmentation, since they are related with brightness and texturized regions.

Acknowledgments

We would like to acknowledge support for this research from UFMG, CAPES/MEC, CNPq/MCT.

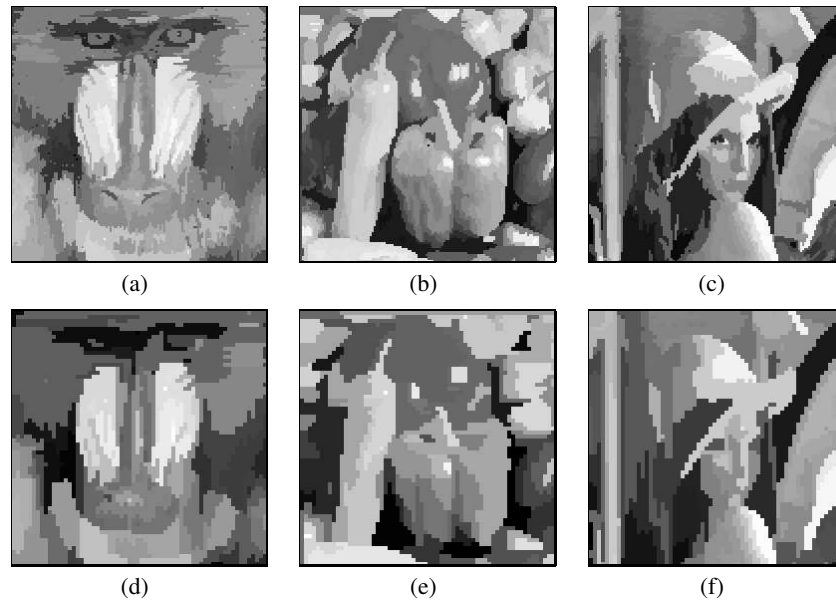


Fig. 4. Experiments used for verifying our proposed approach; using different values of window size for input images shown in Figure 2. Images on first (4(a),4(b) and 4(c)), second (4(d),4(e) and 4(f)) rows are resulting segmented images of our approach using $\lambda_0 = 1.96$, window sizes of $k = 2$ and 4, respectively, related with original ones in Figure 2.

References

1. Canny, J.F.: A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **8** (1986) 679–698
2. Pal, N.R., Pal, S.K.: A review on image segmentation techniques. *Pattern Recognition* **26** (1993) 1277–1294
3. Haralick, R., Shapiro, L.: *Computer and Robot Machine Vision*. Addison-Wesley, USA (1992 and 1993)
4. Shi, J., Malik, J.: Normalized cuts and image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **22** (2000) 888–905
5. Galun, M., Sharon, E., Basri, R., Brandt, A.: Texture segmentation by multiscale aggregation of filter responses and shape elements. In: *Proceedings of the IEEE International Conference on Computer Vision (ICCV2003)*, Nice, France (2003) 716–723
6. Mallat, S.: A theory of multiresolution signal decomposition: The wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **11** (1989) 674–693
7. Montgomery, D.C., Runger, G.C.: *Applied Statistics and Probability for Engineers*. John Wiley & Sons, Inc., New York - Chichester - Brisbane - Toronto - Singapura (1994)
8. Press, W., Teukolsky, S., Vetterling, W., Flannery, B.: *Numerical Recipes in C*. 2nd edn. Cambridge University Press, UK (1996)