

# Watershed Framework to Region-based Image Segmentation

Fernando C. Monteiro

*INEB - Instituto de Engenharia Biomédica  
Instituto Politécnico de Bragança - Portugal  
monteiro@ipb.pt*

Aurélio Campilho

*INEB - Instituto de Engenharia Biomédica  
FEUP - Faculdade de Engenharia, Universidade do Porto - Portugal  
campilho@fe.up.pt*

## Abstract

*This paper proposes a new framework to image segmentation which combines edge- and region-based information with spectral techniques through the morphological algorithm of watersheds. A pre-processing step is used to reduce the spatial resolution without losing important image information. An initial partitioning of the image into primitive regions is set by applying a rain-falling watershed algorithm on the image gradient magnitude. This initial partition is the input to a computationally efficient region segmentation process which produces the final segmentation. The latter process uses a region-based similarity graph representation of the image regions. The experimental results clearly demonstrate the effectiveness of the proposed approach to produce simpler segmentations and to compare favourably with state-of-the-art methods.*

## 1. Introduction

Image segmentation is one of the largest domains in image analysis, and aims at identifying regions that have a specific meaning within images. Another definition of image segmentation is the identification of regions that are uniform with respect to some parameter, such as image intensity or texture. While the latter definition is often used for technical reasons, the former definition should be preferred from an application point of view. Although the effort made in the computer vision community there is no algorithm that is known to be optimum in image segmentation. Much research is being done to discover new methods building up on previous ideas [1, 4, 7].

The main goal of this paper is to develop an algorithm for efficient segmentation of a grey level image that a) identifies perceptually homogeneous regions in the images, b) makes minimal assumptions about the scene, and c) avoids merging of multiple objects into single segments and vice-versa. The definition of a new structure for region similarity graph, the presentation of a new similarity function, and the application of multiclass normalized cuts to group atomic regions are the main contributions of this paper.

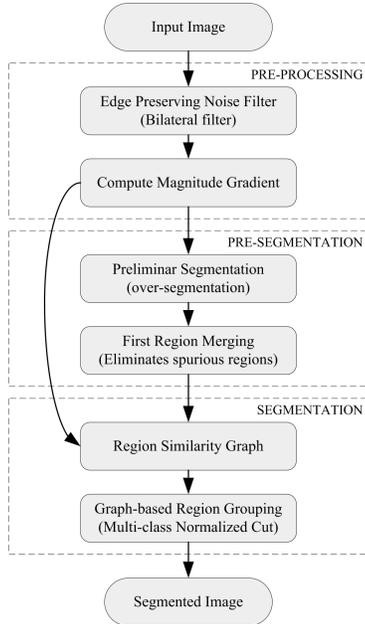
Spectral methods are promising approaches to perceptual grouping or image segmentation that take into account global image properties as well as local spatial relationships. They treat image segmentation as a graph partitioning problem. These methods use the eigenvectors and the eigenvalues of a matrix representation of a graph derived from the pairwise similarities, as measure by one or more cues to partition an image into disjoint regions. One important issue of these approaches is the size of the corresponding similarity matrix.

The combination of watershed and spectral methods solves the weaknesses of each method by using the watershed to provide small prototype regions from which similarity matrix can be obtained. Rather than clustering single points we will cluster micro-regions, confident that the underlying primitive regions are reliable.

The algorithm described in this paper can be well classified into the category of hybrid techniques, since it combines the edge-based, region-based and the morphological techniques together through the spectral-based clustering approach. We propose that our method can be considered as an image segmentation framework within which existing image segmentation algorithms that produce over-segmentation may be used in the preliminary segmentation step.

## 2 Overview of the proposed method

The proposed methodology has four major stages. First, we reduce image noise, as a pre-processing stage, using an anisotropic filter. Next, we create an over-segmented image based on the gradient image and watershed transform. In the third stage, the over-segmented image will be the input for the image Region Similarity Graph (RSG) construction. Finally, we apply a spectral approach on the RSG. This framework integrates edges and region-based segmentation with spectral-based clustering as depicted in Figure 1.



**Figure 1. Block diagram of the proposed method (WNCUT).**

Our Watershed Normalized Cut (WNCUT) approach overcomes the problem of over-segmentation in the preliminary segmentation stage by using the spectral methods to intelligently re-assemble the sub-set of atomic regions into the final segmented object based on a similarity function among the regions.

## 3 Region similarity graph

Spectral methods use the eigenvectors and eigenvalues of a matrix derived from the pairwise similarities of features. This effect is achieved by creating a fully connected graph. Based on the graph construction, there are two main groups of methods for image segmentation: region-based methods where each node repre-

sents a set of connected pixels, and pixel-based methods where each node corresponds to a pixel of the image.

The RSG structure is an undirected weighted graph where the set of nodes corresponds to the atomic regions. For each pair of regions, the set of links represent relationships and the link weights  $\mathbf{W}$  represent similarity measures between the regions.

Some characteristics of the RSG model that yield to relevant advantages with regard to the region adjacency graph model are: i) the RSG allows the existence of links between pairs of non-adjacent regions. ii) it is defined once and it does not need any dynamic updating when merging regions.

### 3.1 Pairwise similarity

In the RSG model nodes are represented by the centroid of each micro-region. Links together with their associated weights are defined using the spatial similarity between nodes, their connectivity and the strength of intervening contours between region centroids. The resulting graph is a structure where region nodes represent complete image regions.

For each pair of nodes, node similarity is inversely correlated with the maximum contour energy encountered along the line connecting the centroids of the regions. If there are strong contours along a line connecting two centroids, these atomic regions probably belong to different segments and should be labelled as dissimilar. Let  $i$  and  $j$  be two atomic regions and the orientation energy  $OE^*$  between them, then the intervening contours contribution to the link weight is given by:

$$w_{ic}(i, j) = \exp \left[ -\frac{\max_{t \in \text{line}(i, j)} \|OE^*(\bar{x}_i, \bar{x}_j)\|^2}{\sigma_{ic}^2} \right] \quad (1)$$

where  $\text{line}(i, j)$  is the straight line between centroids  $\bar{x}_i$  and  $\bar{x}_j$  formed by  $t$  pixels.

The intensity distance between nodes contributes for the link weight according to the following function:

$$w_I(i, j) = \exp \left( -\frac{(I_{\bar{x}_i} - I_{\bar{x}_j})^2}{\sigma_I^2} \right) \quad (2)$$

These cues are combined in a final link weight similarity function, with the values  $\sigma_{ic}$  and  $\sigma_I$  selected in order to maximize the dynamic range of  $\mathbf{W}$ :

$$\mathbf{W}(i, j) = w_{ic}(i, j) \cdot w_I(i, j) \quad (3)$$

For each region  $r_i$ , spatial location  $\bar{x}_i$  is computed as centroid of their pixels. Two dynamic data structures are used through which it is very convenient to add or remove regions: 1) A label map in which each

pixel value corresponds to the label of the region that this pixel belongs to; 2) An array of regions where each region is represented by a linked-list of pixels which correspond to the pixels that belong to the region. This dual representation of a partitioned image allows for an efficient implementation. The label map grants us immediate access to the label of every pixel in the image. The array of lists gives us immediate access to the set of pixels that belong to each region. Using this representation two different regions can be merged into one by iterating through the corresponding linked-lists and updating the label map. Even more, we can easily obtain the centroid and the mean value of the new region.

## 4 Results

For image segmentation we used images from the Berkeley Segmentation Dataset [6]. This database comprises a ground truth of 300 hand-segmented images by a minimum of 5 subjects, to compare the segmentation outputs. We identify each image with the number presented in [6]. Although some optimisation could be made, in our experiments we use the same threshold values for every images. The standard deviation of the similarity measures proposed in Equations (1) and (2) are  $\sigma_{ic} = 0.02$  and  $\sigma_I = 0.02$  times the maximum intensity value of the image.

### 4.1 Evaluation

We have compared our method (WNCUT) with three state-of-the-art segmentation algorithms: (i) mean shift (EDISON) [2], (ii) a multiscale graph based segmentation method (MNCUT) [3], and (iii) JSEG [5]. For this comparison we use the set of natural images shown in Figure 2. To provide a numerical evaluation measure and thus allow comparisons, the experiments for the evaluation were conducted on the manual segmentations of the Berkeley Segmentation Dataset [6]. The task is cast as a boundary detection problem, with results presented in terms of Precision (P) and Recall (R) measures. The algorithm provides a binary boundary map which is scored against each one of the hand-segmented results of Berkeley Dataset, producing a  $(R, P, F)$  value. The final score is given by the average of those comparisons. The quantitative evaluation results (F-measure) are summarized in Table 1.

The resulting segmentation after the application of the examined algorithms is shown in Figure 2. Since the F-measure is a boundary-based measure the segmentation results are presented as boundaries over the original images. The proposed approach produces segmentations of high quality and with better results than

the other methods for all tested images. This new approach overcomes some limitations usually associated with spectral clustering approaches.

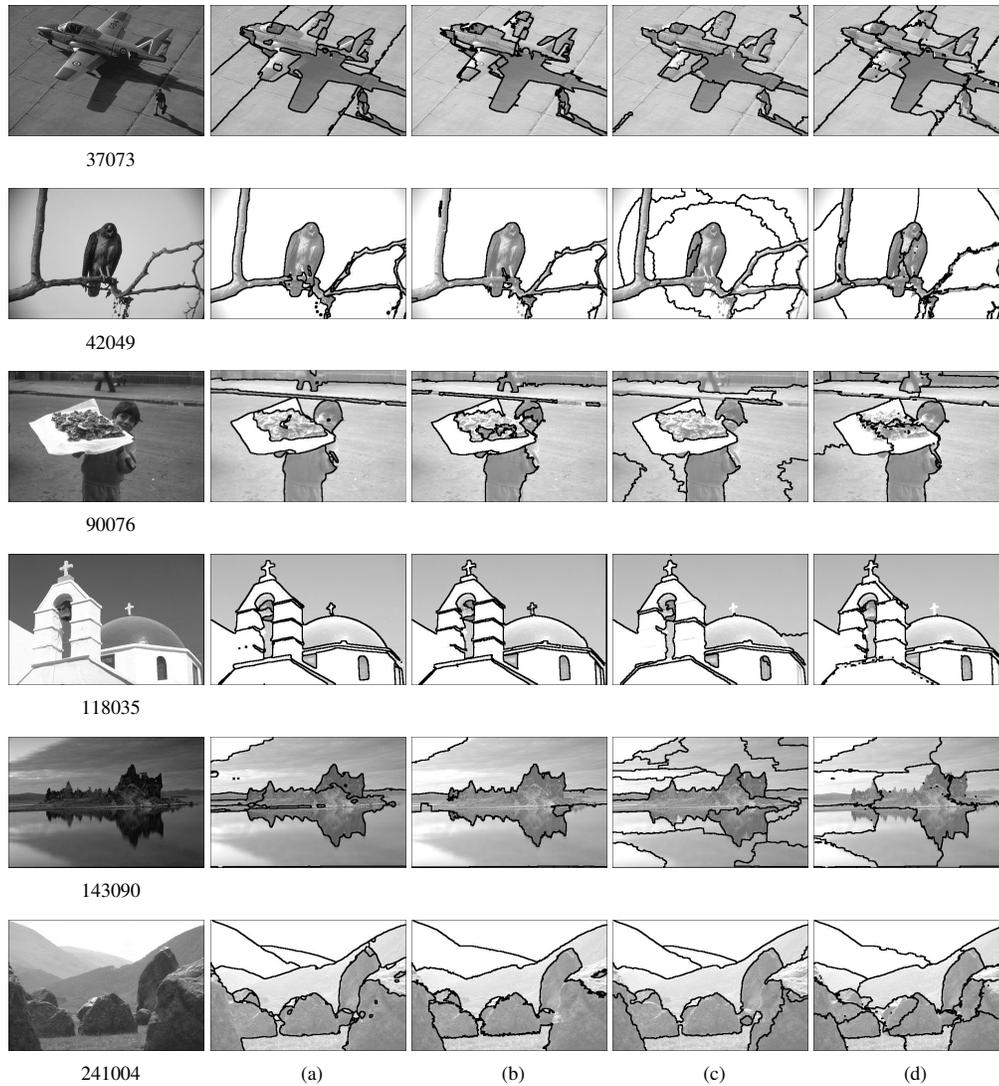
## 5 Conclusions

In this paper we have proposed an image segmentation method which combines edge- and region-based information with spectral techniques through the morphological algorithm of watersheds. An initial partitioning of the image into primitive regions is set by applying a rainfalling watershed simulation on the image gradient magnitude. This step presents a new approach to overcome the problems with flat regions. This initial partition is the input to a computationally efficient graph partition process that produces the final segmentation. The latter process uses a region similarity graph representation of the image regions.

Using small atomic regions instead of pixels leads to a more natural image representation - the pixels are merely the result of the digital image discretization process and they do not occur in the real world. Besides producing smoother segmentations than pixel-based partitioning methods, it also reduces the computational cost in several orders of magnitude.

## References

- [1] R. O. Callaghan and D. Bull. Combined morphological-spectral unsupervised image segmentation. *IEEE Transactions on Image Processing*, 14(1):49–62, 2005.
- [2] D. Comaniciu and P. Meer. Mean shift: a robust approach toward feature space analysis. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 24(5):603–619, 2002.
- [3] T. Cour, F. Benezit, and J. Shi. Spectral segmentation with multiscale graph decomposition. In *Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition*, volume II, pages 1124–1131, June 2005.
- [4] J. De Bock, P. D. Smet, and W. Philips. Image segmentation using watersheds and normalized cuts. In *Proceedings of the IS&T/SPIE Electronic Imaging 2005*, volume 5675, pages 164–173, January 2005.
- [5] Y. Deng and B. Manjunath. Unsupervised segmentation of color-texture regions in images and video. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(8):800–810, 2001.
- [6] D. Martin, C. Fowlkes, D. Tal, and J. Malik. A database of human segmented natural images and its application to evaluating segmentation algorithms and measuring ecological statistics. In *Proc. of IEEE Int. Conference on Computer Vision*, volume II, pages 416–423, 2001. online at: <http://www.cs.berkeley.edu/projects/vision/grouping/segbench/>.
- [7] F. Monteiro. *Region-based spatial and temporal image segmentation*. PhD thesis, Faculdade de Engenharia da Universidade do Porto, Portugal, 2007.



**Figure 2. Segmentation results: First column: original image. Each image is identified with the Id number used in the dataset, (a) proposed method (WNCUT), (b) Mean shift (EDISON) [2], (c) JSEG [5], and (d) the multiscale segmentation MNCUT [3].**

**Table 1. Results of quantitative evaluation in terms of F-measure for the comparison between the proposed method (WNCUT), Mean shift (EDISON), JSEG and the multiscale segmentation MNCUT. The last row shows the evaluation among hand-segmented results.**

Method	37073	42049	90076	118035	143090	241004
WNCUT	0.65	0.89	0.85	0.74	0.71	0.81
EDISON	0.62	0.85	0.73	0.70	0.62	0.72
JSEG	0.61	0.64	0.66	0.71	0.49	0.78
MNCUT	0.58	0.75	0.64	0.70	0.33	0.69