# Stochastic watershed segmentation

JESÚS ANGULO and DOMINIQUE JEULIN

Centre de Morphologie Mathématique (CMM), Ecole des Mines de Paris, Fontainebleau, France {jesus.angulo,dominique.jeulin}@ensmp.fr

- Abstract This paper introduces a watershed-based stochastic segmentation methodology. The approach is based on using M realizations of N random markers to build a probability density function (pdf) of contours which is then segmented by volumic watershed for defining the R most significant regions. It improves the standard watershed algorithms when the aim is to segment complex images into a few regions. Three variants of the random germs framework are discussed, according to the algorithm used to build the pdf: 1) uniform random germs on the same gradient, 2) regionalised random germs on the same gradient, and 3) uniform random germs on levelled-based gradient. The last algorithm is more complex but it yields the best results.
- **Keywords:** watershed transform, leveling, Poisson points, density of contours, random germs segmentation.

## 1. Introduction

Watershed transformation is one of the most powerful tools for image segmentation. Starting from a gradient, the classical paradigm of watershed segmentation consists in determining markers for each region of interest. The markers avoid the over-segmentation (a region is associated to each minimum of the function) and moreover, the watershed is relatively robust to marker position [2]. The markers-based watershed is appropriate for interactive segmentation. Several watershed-based hierarchical approaches allow addressing fields where the markers cannot be easily defined (e.g., multimedia applications). Mainly, two hierarchical techniques can be distinguished: 1) non-parametric waterfalls algorithm [3] and 2) hierarchies based on extinction values, which allows to select the minima used in the watershed according to morphological criteria (dynamics, surface area and volume) [10, 15].

The volume-based hierarchical segmentation is particularly useful in many applications aiming at segmenting natural images since the volume, which combines the criteria of dynamics and area, selects the most significant regions from a visual viewpoint. However, the performance of the approach decreases drastically when the image is segmented in very few regions, which is just the goal of several applications (e.g., segmentation-based image indexing). Figure 1 gives four colour images segmented by volumic watershed into R = 10, 20 and 50 regions and we can observe that many important regions are not well determined (even when R = 50). The classical solution involves to filter out the image in order to simplify the details and to enhance the main regions (typically using morphological filters such as levelings [11]).



Figure 1. Examples of colour images segmented by means of volumic watershed into the *R* most significant regions. First column, original images **f**; second column, colour gradient  $\varrho^{LS+H}(\mathbf{f})$ ; third column,  $sg^{R-vol}(\varrho^{LS+H}(\mathbf{f}), 10)$ ; fourth column  $sg^{R-vol}(\varrho^{LS+H}(\mathbf{f}), 20)$ , and last column  $sg^{R-vol}(\varrho^{LS+H}(\mathbf{f}), 50)$ .

In fact, the problem lies in the deterministic criterion of volume, computed for each minimum of the function to flood, which depends on the local image information; and nevertheless, the final flooding watershed is a competition between the different minima to determine the optimal partition (in fact, it is the solution of shortest path problem when the path cost is given by the maximum of the arc weights in the path [10]). The aim of this paper is to introduce a watershed-based probabilistic framework to detect the contours which are robust with respect to variations in the segmentation conditions. More precisely, we explore here a stochastic approach based on using random markers to build a probability density function of contours which is then segmented by volumic watershed for defining the most significant regions. Keeping in mind that the goal is the unsupervised segmentation of natural images in very few regions. The probabilistic segmentation has been already studied in the literature, for instance using cooccurrence probability on graphs [5], Bayesian framework [16], Markov Random Fields [7] (combined with watershed segmentation [8]), Markov Chain Monte Carlo [14]. But to our knowledge, this is first study of probabilistic segmentation based on random markers simulations for watershed transformation. The closest previous work to our study is [13], where the sum of watersheds from a series of polarimetric images was used to define a final distribution of contours.

## 2. Basic notions and operators

Watershed segmentation. The function used in the watershed transformation is the image gradient. In this paper, the aim is to segment colour images and hence, according to our previous works [1], we propose to compute a colour gradient in a luminance/saturation/hue (LSH) representation, presenting better performances that other colour gradients. But any other colour gradient can be also used, including for instance marginal gradients in RGB. Let  $\mathbf{f}(x) = (f_L(x), f_S(x), f_H(x))$  be a colour image in the LSH representation, the colour gradient is given by  $\varrho^{LS+H}(\mathbf{f}(\mathbf{x})) =$  $f_S(x) \times \varrho^{\circ}(f_H(x)) + (1 - f_S(x)) \times \varrho(f_L(x)) + \varrho(f_S(x))$ , where  $\varrho(g(x))$  is the morphological gradient of the scalar function g(x) and  $\varrho^{\circ}(a(x))$  is the circular centred gradient of the angular function a(x).

Two watershed algorithms are used in this study. Let mrk(x) be the image of markers, the binary image of segmentation contours associated to these markers, and according the colour gradient  $\varrho^{LS+H}(\mathbf{f}(x))$ , is denoted by  $sg^{mrk}(\varrho^{LS+H}(x))$ . Using the same gradient, the volumic-based segmentation into R regions is named  $sg^{R-vol}(\varrho^{LS+H}(x))$ .

**Leveling.** The leveling  $\lambda(mrk, f)$  of a reference function f and a marker function mrk (f(x) and mrk(x) are two grey level images) can be computed by means of an iterative algorithm with geodesic dilations/erosions [11]. Several extensions to colour images have been proposed for levelings. We propose for this study to apply a marginal approach in RGB, which consists in computing a separated leveling for each red/green/blue colour component, i.e. the colour leveling of image  $\mathbf{f}(x) = (f_R(x), f_G(x), f_B(x))$  according to the markers mrk(x) is the colour image  $\lambda(\mathbf{f}, mrk) = (\lambda(mrk, f_R), \lambda(mrk, f_G))$ . The marginal approach introduces false colour but this is not critical for segmentation purposes.

Generation of random germs. The paradigm of watershed segmentation lays on the appropriate choice of markers, which are the seeds to generate basins of attraction [2,3]. It is claimed, and known from practice, that the most intelligent part of this technique of segmentation resides in the development of criteria used to select the required markers. In the present approach, we follow an opposite direction, by selecting random germs for markers. This arbitrary choice will be balanced by the use of a given number M of realizations, in order to filter out non significant fluctuations.

A rather natural way to introduce random germs [9] is to generate realizations of a Poisson point process with a constant intensity (namely average number of points per unit area)  $\theta$ . It is well known that the random number of points N falling in a domain D with area |D| follows a Poisson distribution with parameter  $\theta |D|$ . In addition, conditionally to the fact that N = n, the *n* points are independently and uniformly distributed over *D*. In what follows, we will fix the value N of the number of random germs (instead of using a random number as for the Poisson point process), and we will generate independent realizations of the location of the germs in D. In some cases, as will be illustrated below, it may be interesting to generate a non-uniform distribution of germs, with a regionalised intensity (or measure)  $\theta(x)$ . In the Poisson case, N follows a Poisson distribution with parameter  $\theta(D)$ , and conditionally to the fact that N = n, the n points are independently distributed over D with the probability density function  $\theta(x)/\theta(D)$ . In what follows, the intensity  $\theta(x)$  will be generated from the image, and we will use a fixed number of germs N, as for the homogeneous case.

**Parzen method to calculate a pdf.** The kernel density estimation, or Parzen window method [6], is a way of estimating the probability density function (pdf) of a random variable. Let  $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_M \in \mathbb{R}^n$  be M samples of a random variable, the kernel density approximation of its pdf is:  $\hat{f}_h(\mathbf{x}) = \frac{1}{Nh} \sum_{i=1}^N K(\frac{\mathbf{x}-\mathbf{x}_i}{h})$ , where  $K(\mathbf{x})$  is some kernel and the bandwidth h a smoothing parameter. Usually, K(x) is taken to be a Gaussian function with mean zero and variance  $\sigma^2$ , which determines the smoothing effect.

## 3. Uniform random germs segmentation

Let  $\{mrk_i(x)\}_{i=1}^{M}$  be a series of M realizations of N uniform random markers. Each one of these binary images of points is considered as the markers for a watershed segmentation of colour gradient  $sg^{mrk}(\varrho^{LS+H}(x))$  and consequently, a series of segmentations is obtained, i.e.,  $\{sg_i^{mrk}(x)\}_{i=1}^{M}$ , see Figure 2. Note that the number of points determines the number of regions obtained (i.e., essential property of watershed transformation). As we can observe from the example, the main contours appear regardless of the position of germs.

Starting from the M realizations of contours, the probability density function of contours is computed by Parzen window method. The smoothing effect of the Gaussian kernel (typically  $\sigma = 3$ ) is important to obtain a function where closed contours (e.g., in textured regions or associated to small regions) are added together. The pdf(x) could be thresholded in order to obtain the most prominent contours, however the result are only pieces of contours (not enclosing regions). In addition, we have studied the histograms for several examples and there is not an optimal threshold to separate the classes of contours.



Figure 2. Left, M realizations of N uniform random germs,  $mrk_i(x)$ , and corresponding marker-based watershed contours,  $sg_i^{mrk}(x)$ . Right, probability density function of contours computed by Parzen window method for N = 10 and M = 50.

The main drawback of using a uniform distribution of the random markers is to induce an over-segmentation of the largest watersheds, since the average number of germs falling in a given region is proportional to its area. This is avoided by means of a volume-watershed segmentation, or by using a regionalised intensity of germs, as illustrated later. We propose to partition the pdf(x) of contours with the volume-based watershed to obtain the R most significant regions, i.e.,  $sq^{R-vol}(pdf(x), R)$ . Each catchment basin (each minima) of pdf(x) corresponds to one the regions of the sum (or union) of the different  $sg_i^{mrk}(x)$  and the integral of each catchment basin corresponds to the probability to be region of the segmentation. Consequently, the volumic watershed of pdf(x) yields the regions according to their probabilities. In Figure 3 is given a comparison of segmentation into R = 10, 20 and 50 regions for two different pdf(x). The results should be compared with those associated to  $sg^{R-vol}(\varrho^{LS+H}(x), R)$  (see Figure 1). A property of the Gaussian filter, observed from the examples, is the regularisation of pdf(x) which involves relatively rounded watershed contours.

#### **3.1** Influence of parameters N and M

From the examples of Figure 1 and other similar results, we state that the method hardly depends on the number of realizations M, which is a good characteristic to guarantee its robustness. In practice, we have verified that the pdf(x) converges to a stable distribution of contours even for low values of M (20 or 50). We propose in any case, to take a higher value, typically M = 100 or 200 in order to obtain more regular contours.

The random points explore uniformly the image space and the choice of N is important to fix the degree of stochastic sampling (note that the probability depends on the ratio between N and the image size or number of pixels). Moreover, if the value of N is low, a segmentation into large regions



Figure 3. Left, probability density function of contours, pdf(x) for N unifom germs and M realizations. Right, volumic watershed-based segmentation of pdf(x) into the R most significant regions,  $sg^{R-vol}(pdf), R)$ .

is privileged; instead of a high value of N will produce smaller regions. If N is too high, the over-segmentation of  $sg_i^{mrk}$  leads to a very smooth pdf(x), which loss its property for selecting the R contours. In fact, we can conclude that the uniform germs segmentation is mainly depending on parameter N which is related to R (number of regions to be determined) and it is logical to take N > R. But again, the method is quite robust to the choice of N: from the examples of images of size  $256 \times 256$  to be segmented into R = 10, 20 or 50 the choice of N = 50 or 100 produces exactly the same results.



Figure 4. First row, probabilistic gradient  $\rho(x)$  (i.e., linear combination of colour gradient and pdf) for different values of  $\lambda$  and second row, associated volumic watershed-based segmentation into R = 20 regions.

#### 3.2 Probabilistic gradient

The function pdf(x) can be combined with the initial gradient in order to reinforce the gradient contours which have a high probability:  $\rho(x) = \omega_1 \rho^{LS+H}(\mathbf{f}(x)) + \omega_2 pdf(x)$ , considering a typical barycentric combination (both functions defined in [0, 1]), i.e.,  $\omega_1 = (1 - \lambda)$  and  $\omega_2 = \lambda$ .

We have studied the behaviour of  $\rho(x)$  for volumic segmentation, i.e.,  $sg^{R-vol}(\rho(x))$ , with respect to the value of control  $\lambda$  (note that for  $\lambda = 0$  the gradient is obtained and for  $\lambda = 1$ , exclusively the probability density function of contours). In Figure 4 is shown an example of segmentation into 20 regions for different  $\lambda$ . It is observed that, even for low values of  $\lambda$ , the results of segmentation are notably improved. This is coherent with the fact that the pdf(x), derived from the gradient, contains all the useful information for the segmentation. In any case, we have confirmed on the basis of many other examples that when  $\lambda = 0, 5$  (averaged combination) the results are in general more satisfactory.

#### 4. Regionalised random germs segmentation

In the previous algorithm, the random germs are uniformly distributed in the image domain. We have also studied how a regionalised distribution of germs could be used to build the distribution of contours. The first question to deal with is the choice of the regionalisation function  $\theta(x)$ .



Figure 5. Left, M realizations of regionalised random germs (the function of regionalisation is the colour gradient  $\theta = \rho^{LS+H}$ ),  $mrk_i(x)$ , and corresponding marker-based watershed contours,  $sg_i^{mrk^{\theta}}(x)$ . Right, probability density function of contours computed by Parzen window method for M = 50, and segmentation of  $pdf^{\theta}(x)$  into R = 20 volumic regions.

Several alternatives are possible. We can for instance use the component of luminance  $\theta(x) = f_L(x)$  (respectively, the negative of luminance  $\theta(x) =$  $f_L^c(x)$ ), in such a way that the bright regions (respectively, the dark regions) will produce random germs. It is evident that this kind of regionalisation is not very useful for segmentation. It seems more natural to work on the colour gradient,  $\theta(x) = \rho^{LS+H}(x)$ . In this case, the germs of  $mrk_i^{\theta}(x)$  are located around the zones of high gradient value, that is the zones closed to the contours. Once the series of M contours  $sg_i^{mrk^{\theta}}(x)$  is computed, the corresponding probability density of contours  $pdf^{\theta}(x)$  is obtained by the Parzen window method. As previously, this function is finally segmented by volume-based watershed, see the example of Figure 5. The regionalised segmentation depends on the properties of dynamics of colour gradient. Moreover, the different random point realizations using the same  $\theta(x)$  are quite similar and consequently, the realizations of contours too. By this regionalised sampling, another characteristic of the obtained  $pdf^{\theta}(x)$  is that the distribution is very similar to the gradient, but where all the contours are enhanced. The final results of segmentation for  $sg^{R-vol}(pdf^{\theta}, R)$  are in any case better than for  $sq^{R-vol}(\rho^{LS+H}, R)$ . We have also evaluated the interest of  $\theta(x)$  equal to the negative of the gradient (i.e., locating germs in low gradient zones); however in this case too many germs are introduced in each realization and the over-segmentation involves useless pdf's.

#### 5. Uniform random germs leveling and segmentation

The morphological connected filters suppress details but preserve the contours of the remaining objects. Levelings are a subclass of symmetric connected operators which are very useful to simplify an image before segmentation by watershed transformation [11]. In fact, the image marker for the leveling is a rough simplification of original image. Pushing our approach to the limit, the rationale behind the last variant of the proposed stochastic segmentation is based on using the random germs as markers before for the leveling, in order to obtain a very simplified gradient on which is computed the watershed with the same markers.

The steps of this algorithm are summarised as follows (see Figure 6).

- To throw the *M* realizations of *N* uniform random germs:  $\{mrk_i(x)\}_{i=1}^M$ .
- To compute the leveling for the colour image associated to each image of germs: lev<sub>i</sub>(x) = λ(f, mrk<sub>i</sub>).
- To calculate the series of colour gradients associated to the leveled colour image:  $\rho_i(x) = \rho_i^{LS+H}(\mathbf{lev}_i)$ .
- Each colour gradient  $\rho_i$  is segmented with the markers  $mrk_i$ :  $sg_i^{lev-mrk}(x) = sg^{mrk_i}(\rho_i).$



Figure 6. Left, M realizations of N uniform random germs,  $mrk_i(x)$ , marginal colour levelings using the random germs as markers,  $lev_i(x)$ , associated colour gradients,  $\varrho_i(x)$  and corresponding marker-based watershed contours,  $sg_i^{lev-mrk}(x)$ . Right, probability density function of contours computed by Parzen window method for N = 10 and M = 50, and segmentation of  $pdf^{lev}(x)$  into R = 10 volumic regions.

- To obtain the probability density function of contours:  $pdf^{lev}(x) = \frac{1}{M} \sum_{i=1}^{M} sg_i^{lev-mrk} * G_{\sigma}.$
- Let  $\widehat{\varrho}(x) = \frac{1}{M} \sum_{i=1}^{M} \varrho_i(x)$  be the averaged colour gradient for the M realizations, to compute the leveling-based probabilistic gradient which is defined as follows:  $\rho^{lev}(x) = (1 \lambda)\widehat{\varrho}(x) + \lambda pdf^{lev}(x)$  (typ.  $\lambda = 0.5$ ).
- To segment by volumic watershed into R regions the function of contours  $sg^{R-vol}(pdf(x))$  (or the probabilistic gradient  $sg^{R-vol}(\rho^{lev}(x))$ ).

Figure 7 shows a final comparison with examples of colour images segmented by volumic watershed (R = 10, 20 and 50) on probabilistic gradient  $\rho(x)$  and on leveling-based probabilistic gradient  $\rho^{lev}(x)$  (both with  $\lambda = 0.5$ ). These results should be compared with those of Figure 1. It is evident that the stochastic algorithms proposed in this study yields to better segmentations than the standard watershed. It is observed also that the last method including levelings results in very good image partitions.

Proceedings of the 8<sup>th</sup> International Symposium on Mathematical Morphology, Rio de Janeiro, Brazil, Oct. 10–13, 2007, MCT/INPE, v. 1, p. 265–276. http://urlib.net/dpi.inpe.br/ismm@80/2007/03.05.12.37



Figure 7. Examples of colour images segmented by means of volumic watershed into the R most significant regions (R = 10, 20 and 50) on probabilistic gradients  $\rho(x)$  and  $\rho^{lev}(x)$  (both with  $\lambda = 0.5$ ) and derived from pdf(x) and  $pdf^{lev}(x)$  respectively (both for N = 100 and M = 200).

# 6. Implementation issues

The M realizations of uniform/regionalised random germs contours are obtained from the same function (i.e., colour gradient) using different markers. Consequently, working on the neighbourhood graph of catchment basins and its minimum spanning tree (MST) [12], the N random markers can be considered as N random nodes of the MST instead of N image points. Two main advantages are associated to the graph implementation: firstly, a fast computation of M segmentations from different markers on the same MST; and secondly, the control of watershed bias which could be associated to the random positions of markers [4].

The algorithm using the uniform random germs as markers, first for the levelling and then for the watershed has an upper computational load (time of computation). Moreover, in each realization, the gradient is different (i.e., a different graph) and therefore the MST cannot be reused. In any case, nowadays using the fast implementations of watershed algorithms (100 ms for a  $256 \times 256$  images running on a current standard Laptop), the time of execution to segment a colour image according our stochastic framework is around 10 s.

# 7. Discussion and conclusions

We have introduced in this paper a new morphological stochastic segmentation approach which improves the standard watershed algorithms when the aim is to segment complex images into a few regions. The improvement in the segmentation is less important for images presenting specific objects on a homogenous background. We have illustrated three variants of the random germs framework, according to the algorithm used to build the probability density of contours: uniform random germs on the same gradient, regionalised random germs on the same gradient and uniform random germs on levelled-based gradient. The last algorithm is more complex but it yields the best results.

In ongoing research, we consider to explore other variants using evolved random point simulations (structural grids, conditional models, etc.), working on a multi-scale framework (image pyramids and image decompositions). We are also working on probabilistic approaches combining colour gradients and texture information. In the last case, probabilistic rules of aggregation in the construction of the watersheds from the random seeds would introduce a second level of randomness in the segmentation process.

From a fundamental viewpoint, the idea behind our approach is that there are two types of contours associated to the watershed of a gradient:  $1^{st}$  order contours, which correspond to "significant" regions and which are relatively independent from markers; and  $2^{nd}$  order contours, associated to "small", "low contrasted" or "textured" regions and which depend strongly on the place of markers. Our probabilistic framework aims at enhancing the  $1^{st}$  order contours from a sampling effect, to improve the result of watershed. It should be interesting to study if it is possible to determine by deterministic methods the type of each contour present in an image.

#### References

- J. Angulo and J. Serra, Modelling and Segmentation of Colour Images in Polar Representations, Image Vision and Computing 25 (2007), no. 4, 475–495.
- [2] S. Beucher and F. Meyer, The Morphological Approach to Segmentation: The Watershed Transformation, New York (1992), Mathematical Morphology in Image Processing, (E. Dougherty Ed.), pp. 433–481.
- [3] S. Beucher, Watershed, hierarchical segmentation and waterfall algorithm, ISMM'94 (1994), Mathematical Morphology and its Applications to Image and Signal Processing, pp. 69–76.
- [4] S. Beucher, Algorithmes sans biais de Ligne de Partage des Eaux, Ecole des Mines de Paris, Internal Note CMM, 34 p., Februry 2002.
- [5] K. Cho and P. Meer, *Image Segmentation from Consensus Information*, Computer Vision and Image Understanding 68 (1996), no. 1, 72–89.
- [6] R. O. Duda and P. E. Hart, Pattern Classification and Scene Analysis, Wiley, New York, 1973.
- [7] S. Geman and D. Geman, Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images, IEEE Trans. on Pattern Analysis and Machine Intelligence 6 (1984), 721–741.
- [8] V. Grau, A. U. J. Mewes, M. Alcañiz, R. Kikinis, and S. K. Warfield, Improved Watershed Transform for Medical Image Segmentation Using Prior Information, IEEE Trans. on Medical Imaging 23 (2004), 447–458.
- [9] D. Jeulin, Modèles morphologiques de structures aléatoires et de changement d'échelle, Université Caen, Thèse Doctorat ès Sciences Physiques, 1991.
- [10] F. Meyer, An Overview of Morphological Segmentation, International Journal of Pattern Recognition and Artificial Intelligence 15 (2001), no. 7, 1089–1118.
- [11] \_\_\_\_\_, Levelings, Image Simplification Filters for Segmentation, Journal of Mathematical Imaging and Vision 20 (2004), 59–72.
- [12] \_\_\_\_\_, Grey-weighted, ultrametric and lexicographic distances, ISMM'05 (2005), Mathematical Morphology and its Applications to Image and Signal Processing, pp. 289–298.
- [13] J. Serra and M. Mlynarczuk, Morphological merging of multidimensional data, STERMAT'00 (2000), 6th International Conference Stereology and Image Analysis in Materials Science, pp. 385–390.
- [14] Z. Tu and S.-C. Zhu, Image Segmentation by Data-Driven Markov Chain Monte Carlo, IEEE Trans. on Pattern Analysis and Machine Intelligence 24 (2002), 657– 673.
- [15] C. Vachier and F. Meyer, Extinction value: a new measurement of persistence, (1995), IEEE Workshop on Nonlinear Signal and Image Processing, pp. 254–257.
- [16] K. L. Vincken, A. S. E. Koster, and M. A. Viergever, *Probabilistic Multiscale Image Segmentation*, IEEE Trans. on Pattern Analysis and Machine Intelligence **19** (1997), 109–120.