Validation of Watershed Regions by Scale-Space Statistics

Tomoya Sakai and Atsushi Imiya

Institute of Media and Information Technology, Chiba University, Japan
{tsakai,imiya}@faculty.chiba-u.jp

Abstract. This paper shows a potential use of scale space for statistical validation of watershed regions of a greyscale image. The watershed segmentation has difficulty in distinguishing valid watershed regions associated with real structures of the image from invalid random regions due to background noise. In this paper, a hierarchy of watershed regions is established by following merging process of the regions in a Gaussian scale space. The distribution of annihilation scales (lives) of the regional minima is investigated to statistically judge the regions as being valid or not. Recursive validation using the hierarchy prevents oversegmentation due to the randomness.

1 Introduction

The aim of this study is to develop a statistical validation scheme for segmentation of a greyscale image. If we do not have a priori knowledge on the shapes or structures of objects in the image, topographic features of the greyscale image, and the watersheds in particular, are useful for unsupervised image segmentation. A well-known phenomenon in the watershed segmentation is oversegmentation, that is, producing a large number of undesired tiny regions. Since the undesired watershed regions are mainly caused by noise in the image, it is desirable to settle the oversegmentation problem by taking account of statistical properties of the randomness.

There is a body of literature dealing with the oversegmentation problem of watersheds [1,2,3,4,5,6,7,8]. In the antecedent work, most schemes for preventing the oversegmentation attempt to hierarchically merge the oversegmented regions on the basis of similarity between adjacent regions measured by the MDL [3], colour distance [8], and so on. Diffusion-based multiscale image representations are also used for merging the regions [5,6,8], since the scale space theory [9,10,11,12,13,14,15] mathematically underpins topological relationships among the topographic features without a priori knowledge about them. The oversegmentation can be reduced by selecting levels in the hierarchy of regions, or by setting lower bounds to the scale above and below which the watersheds are valid and invalid, respectively.

In this paper, we show that the scale-space treatment of the image is also useful for the statistical analysis of the random watershed regions. The validity

X.-C. Tai et al. (Eds.): SSVM 2009, LNCS 5567, pp. 175–186, 2009.

of a watershed region can be quantified in terms of the statistical confidence of distinguishing it from invalid watershed regions due to randomness. We present a fully unsupervised watershed segmentation algorithm, in which the watershed regions are recursively validated according to their hierarchical relationships in the scale space.

2 Watershed Segmentation with Variable Scale

2.1 Gaussian Scale Space

In the Gausian scale-space theory [9,10,11,12,14,15,16], a one-parameter family of nonnegative functions is derived from a d-dimensional greyscale image $f(\boldsymbol{x})$, $\boldsymbol{x} \in \mathbb{R}^d$.

$$f(\boldsymbol{x}, \sigma) = G(\boldsymbol{x}, \sigma) * f(\boldsymbol{x}) \tag{1}$$

Here, "$*$" expresses d-dimensional convolution, and $G(\boldsymbol{x}, \sigma)$ is an isotropic Gaussian function with the *scale* σ.

$$G(\boldsymbol{x}, \sigma) = \frac{1}{\sqrt{2\pi}^d \sigma^d} \exp\left(-\frac{|\boldsymbol{x}|^2}{2\sigma^2}\right) \tag{2}$$

We redefine the d-dimensional greyscale image and its scale-space representation in the extended real scale and space as follows.

Definition 1. *A d-dimensional greyscale image is defined as a nonnegative scalar function $f(\boldsymbol{x})$, $\boldsymbol{x} \in \bar{\mathbb{R}}^d$ with a finite net image intensity $\int_{\boldsymbol{x} \in \bar{\mathbb{R}}^d} f(\boldsymbol{x}) dx^d$.*

Definition 2. *The scale-space image $f(\boldsymbol{x}, \sigma)$, $(\boldsymbol{x}, \sigma) \in (\bar{\mathbb{R}}^d, \bar{\mathbb{R}}^+)$, is the convolution of the greyscale image $f(\boldsymbol{x})$ with the isotropic Gaussian kernel $G(\boldsymbol{x}, \sigma)$.*

Here, $\bar{\mathbb{R}}^d$ and $\bar{\mathbb{R}}^+$ denote the d-dimensional extended real space including a point at infinity and the extended real scale including an infinite scale, respectively. Although the domain of a greyscale image in practice is bounded within a limited area or volume, we embed such an image in the extended real scale space. The point at infinity will be theoretically used as a representative point of the background of the image in the watershed segmentation later.

2.2 Watershed Segmentation and Hierarchy of Regions

The watershed segmentation was derived from spatial partitioning on the basis of the drainage patterns of rainfall. As the topographic height map defines the boundaries of the catchment basins draining to the same lowest points, a two-dimensional greyscale image defines the watershed boundary curves enclosing regions with local minima when we regard the image intensity as the topographic height. For a d-dimensional image, the entire space is partitioned by $(d-1)$-dimensional hypersurfaces into d-dimensional watershed regions. Each watershed region defined by a smooth function $f(\boldsymbol{x})$ contains a unique local minimum, to which any point in the watershed region is connected by a gradient curve of $f(\boldsymbol{x})$. In practice, the watershed segmentation of the gradient image $|\nabla f(\boldsymbol{x})|$ is known to provide better intuitive partitions than that of the image $f(\boldsymbol{x})$ itself [2,5,6,8] because object boundaries in a scene may cause large spatial changes in the image intensity.

Simple computation of the watersheds of the images results in oversegmentation caused by tiny and insignificant catchment basins. As suggested in the antecedent work [3, 5, 6, 8], hierarchical relationships among the watershed regions are of great help for merging the oversegmented regions. We employ the scale-space framework to derive the hierarchy because the scale-space axioms are acceptable in general cases where any prior information about the similarities among the unexpected watershed regions are not given. If we apply the gradient watershed segmentation to the image $f(\boldsymbol{x}, \sigma)$ with the variable scale σ, we can observe the evolution of the watersheds with respect to scale. The catastrophy theory applied to the gradient watershed segmentation in the Gaussian scale space [5] shows that the gradient watershed regions of $f(\boldsymbol{x}, \sigma)$ may be generically annihilated, merged, created and splitted with increasing scale σ. Therefore, hierarchical watershed segmentation using multiscale representation of the image [2,6,8] is essentially the extraction of the hierarchical relationships among the watershed regions in the scale space through the generic events.

Since every watershed region is represented by its local minimum, the trajectories of the regional minima in scale space describe the relationships among the regions. For the purpose of validation of the regions, we derive the hierarchy from all the traceable regional minima from the finest scale along their trajectories in scale space. We trace the trajectories by local minimisation at every level of scale [16]. In an annihilation or merging event, two regional minima and a saddle between them are involved. We regard one of these two regional minima as a child of the resulting regional minimum after the event. We trace only one of two local minima after a creation or splitting event because we are interested in the hierarchy of the regions at the finest scale. Remark that the point at infinity is a local minimum which exists at any scale. The local minimum at infinity is the regional minimum of the image background because the rainfall in the background region is drained to this ideal point.

The following algorithm RegionHierarchy traces every trajectory of the regional minimum from every pixel $\boldsymbol{p} \in P$ at $\sigma = 0$ until the regional minimum disappears or goes outside the image boundary toward the local minimum at infinity with increasing scale.

RegionHierarchy(set of pixel centres P, image $f(\boldsymbol{p} \in P)$)

let G be a graph with $\mathrm{card}(P) + 1$ nodes with the labels $l = 0, \ldots, N$ where $l = 0$ represents the point at infinity;
store $\sigma_l^{\mathrm{t}} = \infty$ in all nodes of G;
set $\sigma_{\max}$ to be the size of the convex hull of P;
$\sigma := 0$;
$Q := P$;
while $\mathrm{card}(Q) \neq 1$ or $\sigma < \sigma_{\max}$ **do**

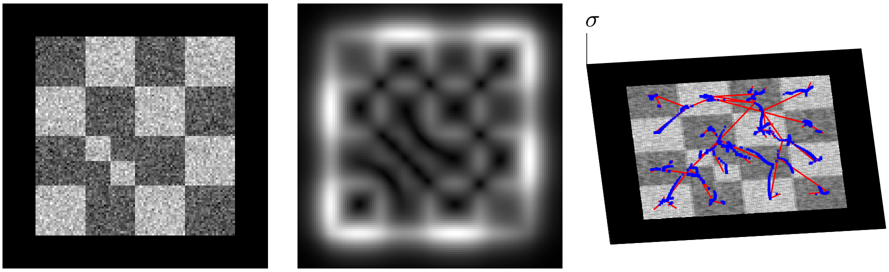
7 $Q' := Q$;
8 $\sigma := \sigma + \Delta\sigma$, where $\Delta\sigma$ is a small value compared with the space intervals of the points Q;
9 **for** each $\boldsymbol{q}_l \in Q$ **do**
10   update $\boldsymbol{q}_l$ by minimising $|\nabla f(\boldsymbol{x}, \sigma)|^2$ with $\boldsymbol{q}_l$ as the initial position [1];
11   **if** $\boldsymbol{q}_l$ is outside the convex hull of P **then**
12   connect the two nodes of G labelled 0 and l;
13   **end if**
14 **end for**
15 let L be a list of labels corresponding to the points in Q;
16 **while** $\mathrm{card}(L) \neq 1$ **do**
17   pop a label l from L;
18   $n :=$ NearestNeighbour(L, l);
19   **if** $|\boldsymbol{q}_l - \boldsymbol{q}_n| < \varepsilon\sigma$, where $\boldsymbol{q}_l, \boldsymbol{q}_n \in Q$, and ε is the tolerance of minimisation **then**
20   **if** $|\boldsymbol{q}_l - \boldsymbol{q}'_l| > |\boldsymbol{q}_n - \boldsymbol{q}'_n|$, where $\boldsymbol{q}'_l, \boldsymbol{q}'_n \in Q'$ **then**
21       $child := l$ and $parent := n$;
22   **else**
23       $child := n$ and $parent := l$;
24   **end if**
25   connect the two nodes of G labelled $parent$ and $child$;
26   set $\sigma^{\mathrm{t}}_{child} := \sigma$;
27   remove $\boldsymbol{q}_{child}$ from Q;
28   **end if**
29 **end while**
30 **end while**
31 **return** G.

The resulting graph G is a set of trees representing the hierarchy of the watershed regions of the gradient image. Any node in G represents a watershed region consisting of the pixels indicated by its subtree nodes. The annihilation or merging scale σ^{t} is stored at the node in G corresponding to $\boldsymbol{p}$.

We utilise the bicubic spline interpolation [17] to seach for the local minimum with subpixel precision in Step 10. The function NearestNeighbour in Step 18 searches for the nearest point to $\boldsymbol{p}_l$ in the set of points listed in L and returns its label. The annihilation or merging event is detected in Step 19, and one of the two regional minima with larger displacement is identified as the child in Step 20.

Figure 1 shows an example of the trajectories of regional minima and the region hierarchy obtained by RegionHierarchy. Since the set of tree, G, expresses hierarchical relationships among the image pixels, any tree node with a scale $\sigma > 0$ represents a set of pixels consisting a watershed region.

[1] It is trivial that the watershed regions of the gradient magnitude squared $|\nabla f|^2$ are identical to those of the gradient magnitude $|\nabla f|$.



(a) (b) (c)

Fig. 1. Trajectories of regional minima and region hierarchy. (a) A noisy 96×96 image $f(\boldsymbol{x})$ embeded in a dark background. (b) Gradient magnitude squared $|\nabla f(\boldsymbol{x}, \sigma = 20)|^2$. The brighter the larger magnitude. (c) Trajectories of the regional minima in scale space. The thick curves (blue) are the parts of the trajectories for $\sigma > 5$. The thin straight lines (red) are the edges of G between the nodes with $\sigma > 5$.

2.3 Scale Selection Problem

We need a criterion to select the scales or the tree levels in hierarchy. One may expect that the watersheds of the image $f(\boldsymbol{x}, \sigma)$ at a small scale σ well approximates the boundary of true image regions. However, if noise spoils the fine structure of the image, the estimated watersheds at small scales are stochastic and experimentally less reproducible. The noise is suppressed at a large scale, but the watershed segmentation is poor in terms of detection ability and localisation: the edges of small watershed regions are smoothed out, and the boundary shapes of large regions are simplified. Since the randomness is the major cause of the oversegmentation problem in the watershed methods [1,4,5], the oversegmentation problem should be resolved in a statistical manner.

3 Validation of Watershed Regions

3.1 Valid Watershed Regions

Generally, a greyscale image expresses spatial distribution of a measured physical quantity. The true image $f^{\mathrm{true}}(\boldsymbol{x})$, which we want to measure and apply the watershed segmentation to, is inevitably spoiled by random noise through the measurement. Therefore, the actual image $f(\boldsymbol{x})$ presents valid watersheds related to those of the true image $f^{\mathrm{true}}(\boldsymbol{x})$ and invalid watersheds due to the randomness.

Assertion 1. *A valid watershed region of an observed gradient image $|\nabla f(\boldsymbol{x})|$ is related to one of the watershed regions of the true gradient image $|\nabla f^{true}(\boldsymbol{x})|$.*

Since the watershed regions are represented by the region minima, the image $f(\boldsymbol{x})$ has the valid watershed regions of the gradient image $|\nabla f(\boldsymbol{x})|$ iff the true gradient image $|\nabla f^{\mathrm{true}}(\boldsymbol{x})|$ has corresponding local minima. Contrapositively, iff $|\nabla f^{\rm true}(\boldsymbol{x})|$ is a featureless image without any local minimum, then no valid watershed exists for any observation $f(\boldsymbol{x})$, which should be considered as an image of the background only. This condition means that $f^{\rm true}(\boldsymbol{x}) = 0$ everywhere in $\bar{\mathbb{R}}^d$ because of the Definition 1. Therefore, $f(\boldsymbol{x})$ for $f^{\rm true}(\boldsymbol{x}) = 0$, i.e., the noise image, produces only the invalid watershed regions. The valid watershed region must be statistically distinguishable from such invalid region. From this viewpoint, the validity of the watershed region is interpreted as the statistical confidence in rejecting the following null hypothesis.

Null hypothesis H_0: The watershed region is that of the noise image.
Alternative hypothesis H_1: The watershed region is not that of the noise image.

The null hypothesis H_0 is rejected if the regional minimum is distinguishable from that of the noise image using test statistics.

3.2 Life Distribution

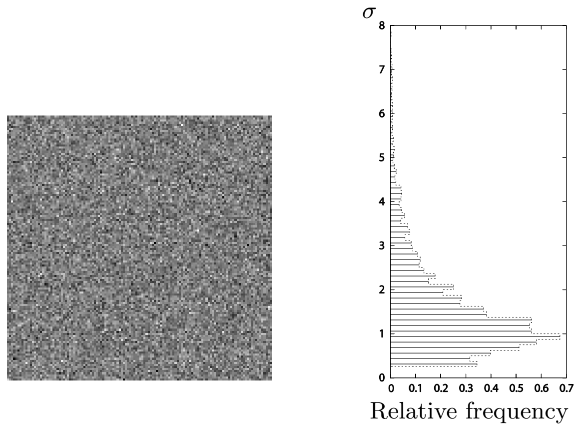
An important fact is that the randomness of the image $f(\boldsymbol{x}, \sigma)$ is filtered out as the scale σ increases, and deterministic features of the image $f(\boldsymbol{x})$ emerge at large scales. In other words, the deterministic features such as the valid watershed regions are established from coarse to fine. There presumably exists a critical lower bound of scale, above and below which the watersheds of $f(\boldsymbol{x}, \sigma)$ are valid and invalid, respectively.

In order to observe how the valid regions survive until large scales against the scale-space filtering, we define the *life* of the watershed region.

Definition 3. *The life of the watershed region is defined as the annihilation scale σ^t of the regional minimum.*

Let $\mathcal{W}$ be a distribution of the lives of the watershed regions of $|\nabla f(\boldsymbol{x}, \sigma)|$ for the image of random noise. If $\mathcal{W}$ can be parametrically modelled, a goodness-of-fit test can be performed under the null hypothesis H_0. That is, if an image $f(\boldsymbol{x})$ is an observation of a true uniform image with noise, then the model of $\mathcal{W}$ fits the distribution of lives $\{\sigma^{\rm t}\}$ of its watershed regions, and H_0 for any watershed regions of $f(\boldsymbol{x})$ is accepted.

We investigate experimentally the life distribution $\mathcal{W}$ for the gradient watershed regions of a Gaussian white noise image as shown in Fig. 2(a). We averaged the frequencies of lives over one hundred noise images. We discard the lives of pixel points whose annihilations are detected in $0 < \sigma \leq \Delta\sigma$ by RegionHierarchy because not all the pixel centres are the local minima. Figure 2(b) is the averaged histogram of life. The obtained life histogram shows an unimodal shape. This implies that there exists a scale where the merging of the regions most frequently occurs. The regional minima of the noise image are uniformly distributed random points, and the regions tend to merge with nearest regions. Therefore, we deduce that this unimodal property is associated with distribution of the nearest neighbour distances of random points. In fact, the nearest neighbour distance distribution has a unimodal shape (See appendix A). The scale of



(a) (b)

Fig. 2. Noise image and the averaged life histogram for its gradient watershed regions. (a) The noise image has uncorrelated random pixel values. (b) The life histogram shows relative frequency of scale at which the regional minima of the gradient image are annihilated as Gaussian blurring of the noise image proceeds.

the mode can be used as a gauge of the density of invalid regions. The regional minima with significantly large values of life out of the unimodal distribution $\mathcal{W}$ can be identified to be valid, because such regional minima are distinguishable from the invalid regional minima of the noise image.

3.3 Recursive Validation

We can set a critical value of the scale to judge the watershed regions valid or invalid. Although the computation of such a critical scale requires the parametric model of the life distribution in the strict sense of statistics, the critical scale can be roughly evaluated by the peak and decaying form of the life histogram. If the image contains valid regions, the life histogram may be multimodal or may have a peak at a small scale relative to the outlying lives representative of the valid regions. According to our experimental result in Section 3.2, a regional minimum with a life which is more than six times greater than the peak can be considered to be valid with the statistical confidence level $\alpha > 99\%$ under the assumption of uncorrelated Gaussian random pixel values of a two-dimensional image as the noise.

We present an algorithm RegionDiscovery for discovery of the valid watershed regions. This algorithm recursively validates the regions in a top-down fashion using each tree T in G by RegionHierarchy. According to the hierarchy, any discovered region is split into subregions as long as they are valid. Each subregion is validated using the life histograms constructed from the lives stored in the subtrees of T corresponding to the subregion.

RegionDiscovery(tree T, set of valid regions V, significance level α)

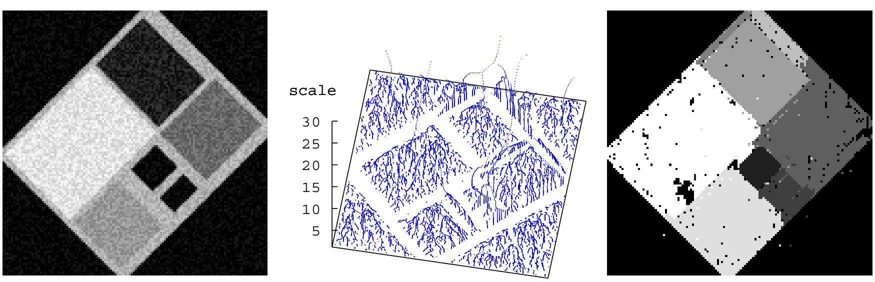
let Σ be a set of life values stored in T except the root;
let s be the subroot node of T with the largest life value $\sigma_{\max} \in \Sigma$;
if IsMultimodal(Σ) or IsOutlier($\sigma_{\max}$, Σ, α) **then**
  RegionDiscovery(Subtree(T, s), V, α);
  RegionDiscovery($T \setminus$Subtree(T, s), V, α);
else
  push the region $R :=$ Pixels(T) into V;
end if.

Here, the function IsMultimodal returns **true** if the histogram of Σ is not unimodal. IsOutlier returns **true** if the life σ^{t} is greater than the critical α-level of scale computed from the given set of lives Σ. Note that these functions discard the lives in $0 < \sigma \leq \Delta\sigma$. Subtree extracts the subtree with subroot node s from the tree T. Pixels returns a set of pixels whose labels are recorded at the nodes in the given tree.

The following function, Watershed, executes our watershed segmentation algorithm for a given image f with a set of pixels P and a significance level α. It returns the set of valid watershed regions consisting of subsets of P.

Watershed(set of pixel centres P, image f, significance level α)

set $V := \emptyset$;
$G :=$ RegionHierarchy(P, f);
for each tree T in G **do**
  RegionDiscovery(T, V, α);
end for
return V.



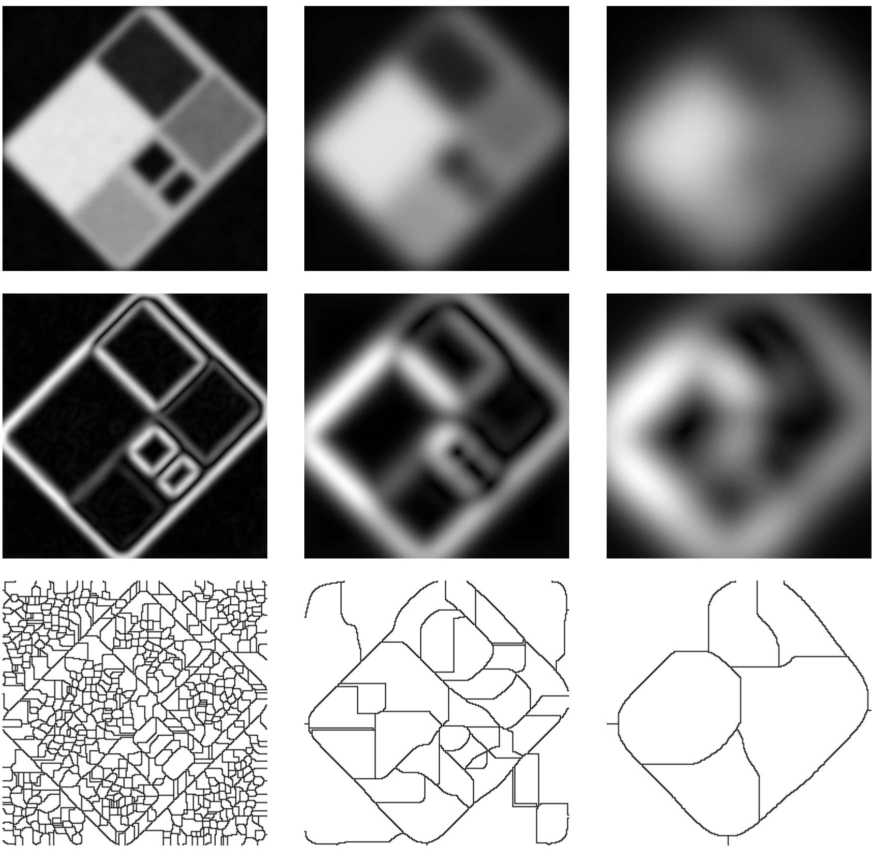
(a) (b) (c)

Fig. 3. An example of our watershed segmentation of noisy image. (a) Original image with 20% noise. (b) Trajectories of local minima of the gradient magnitude of (a) in scale space. The trajectories reaching out of the spatial domain are subordinate to a local minimum at infinity. (c) Watershed regions of the gradient magnitude by the algorithm Watershed. The brightness indicates the order of lives.

4 Test Example

We demonstrate our gradient watershed segmentation WATERSHED for a noisy greyscale image. The purpose of this section is not to test the performance of the algorithm, but to show that the statistics in scale space has potential to discover the valid watershed regions without any prior information about them.

Figure 3(a) shows a 128×128 test image $f(\boldsymbol{x})$ with 20% additive noise [18]. The trajectories of local minima of $|f(\boldsymbol{x}, \sigma)|$ traced from $\sigma = 0$ in scale space are shown in Fig. 3(b). We see a large number of local minima created by the noise at small scales. As the scale increases, the local minima are hierarchically grouped and representative local minima survive at larger scales. Figure 3(c) shows the segmentation result with a confidence level $\alpha = 99\%$ for $f(\boldsymbol{x})$. There are nine



$\sigma = 2$ $\sigma = 6$ $\sigma = 12$

Fig. 4. Watershed segmentation of Fig. 3(a) at different scales. First row: the scale-space image $f(\boldsymbol{x}, \sigma)$. Second row: the gradient magnitude $|\nabla f(\boldsymbol{x}, \sigma)|$. Third row: the watersheds of $|\nabla f(\boldsymbol{x}, \sigma)|$. Each column corresponds to the same scale indicated below.

discovered regions clearly corresponding to the major regions of the original image. The tiny faults in the regions were caused by failure in the minimisation. They were wrongly assigned to the image background, which should be fixed in the future work.

For the comparision purpose, we show in Fig. 4 the simple watershed segmentation results at a few levels of scale without using the region hierarchy or statistics in scale space. We see invalid small regions at small scales while the shapes of valid regions at large scales are distorted. It is remarkable that structural and statistical analyses using scale space can reconstruct the precise edges of statistically valid watershed regions despite the significant noise.

5 Concluding Remarks

The scale-space treatment of the image clarifies not only the hierarchical relationships among the watershed regions but also their statistical properties. We can observe in the Gaussian scale space how the random features are suppressed and deterministic features emerge as the scale grows.

A valid watershed region must be statistically distinguishable from unreproducible regions caused by the random features. The reproducibility is a desirable ability of image recognition techniques. On the basis of this simple requirement we described the null hypothesis H_0, which is to be rejected if the watershed region is valid. A watershed region is recognised as valid at a statistical confidence level in rejecting H_0.

We presented a validation scheme for watershed segmentation using statistics in scale space. We defined the life of a watershed region, whose distribution is useful for testing H_0. We showed that the life distribution for the noise image is unimodal, and the valid regions can be identified by the regional minima with significantly large values of lives out of the unimodal distribution. The statistical properties of the life and the region hierarchy enable the recursive validation of the watershed regions.

A distinctive feature of our scheme is that it does not require any definition of similarity or dissimilarity measures between watershed regions, which is used in many methods for preventing oversegmentation. Instead, we focused on the statistical differences between the valid and invalid regions in scale space. In order to take advantage of the potential of scale-space statistics, our scheme requires further investigation, especially in relation to the model of the life distribution, and improvement and acceleration of the algorithms to obtain feasible segmentation results for larger size real images.

References

1. Vincent, L., Soille, P.: Watersheds in digital spaces: An efficient algorithm based on immersion simulations. IEEE Trans. on Pattern Analysis and Machine Intelligence 13(6), 583–598 (1991)

2. Beucher, S.: Watershed, hierarchical segmentation and waterfall algorithm. In: Proc. Math. Morphology and Its Appl. to Image Processing, pp. 69–76 (1994)
3. Maes, F., Vandermeulen, D., Suetens, P., Marchal, G.: Computer-aided interactive object delineation using an intelligent paintbrush technique. In: Ayache, N. (ed.) CVRMed 1995. LNCS, vol. 905, pp. 77–83. Springer, Heidelberg (1995)
4. Hagyard, D., Razaz, M., Atkin, P.: Analysis of watershed algorithms for greyscale images. In: Proc. of IEEE Intl. Conf. Image Procesing, vol. 3, pp. 41–44 (1996)
5. Olsen, O.F., Nielsen, M.: Generic events for the gradient squared with application to multi-scale segmentation. In: ter Haar Romeny, B.M., Florack, L.M.J., Viergever, M.A. (eds.) Scale-Space 1997. LNCS, vol. 1252, pp. 101–112. Springer, Heidelberg (1997)
6. Gauch, J.M.: Image segmentation and analysis via multiscale gradient watershed hierarchies. IEEE Trans. on Image Processing 8(1), 69–79 (1999)
7. Roerdink, J.B.T.M., Meijster, A.: The watershed transform: definitions, algorithms, and parallelization strategies. Fundamenta Informaticae 41, 187–228 (2001)
8. Vanhamel, I., Pratikakis, I., Sahli, H.: Multiscale gradient watersheds of color Images. IEEE Trans. on Image Processing 12(6), 617–626 (2003)
9. Witkin, A.P.: Scale space filtering. In: Proc. of 8th IJCAI, pp. 1019–1022 (1986)
10. Koenderink, J.J.: The structure of images. Biological Cybernetics 50, 363–370 (1984)
11. Lindeberg, T.: Scale-Space Theory in Computer Vision. Kluwer, Boston (1994)
12. Weickert, J., Ishikawa, S., Imiya, A.: Linear Scale-Space has First been Proposed in Japan. Journal of Mathematical Imaging and Vision 10, 237–252 (1999)
13. Lifshitz, L.M., Pizer, S.M.: A multiresolution hierarchical approach to image segmentation based on intensity extrema. IEEE Trans. on Pattern Analysis and Machine Intelligence 12(6), 529–540 (1990)
14. Florack, L.M.J., Kuijper, A.: The topological structure of scale-space images. Journal of Mathematical Imaging and Vision 12(1), 65–79 (2000)
15. Kuijper, A.: The deep structure of Gaussian scale-space images. PhD thesis, Utrecht University (2002)
16. Sakai, T., Imiya, A.: Gradient structure of image in scale space. Journal of Mathematical Imaging and Vision 28(3), 243–257 (2007)
17. Keys, R.: Cubic convolution interpolation for digital image processing. IEEE Trans. on Acoustics, Speech, and Signal Processing 29(6), 1153–1160 (1981)
18. SAMPL database, `http://sampl.ece.ohio-state.edu/database.htm`
19. Suwa, N.: Quantitative morphology: stereology for biologists. Iwanami Shoten (1977) (in Japanese)

A Distribution of Nearest Neighbour Distances

We present a proof that the nearest neighbour distances obey the Weibull distribution if the points in $\mathbb{R}^d$ are uniformly distributed in a Poisson arrangement [19]. The Poisson arrangement is defined as the uniformly random distribution of points with constant density ρ such that the number of points x in a fixed volume V follows the Poission distribution.

$$\mathrm{Po}(x;\lambda) = \frac{\lambda^x \exp(-\lambda)}{x!} \tag{3}$$

Here, $\lambda = \rho V$ is the expected number of points in the volume V. Let r be the distance from an arbitrary point. The distribution of the nearest neighbour distances, $p(r)$, can be regarded as the probability that the nearest neighbour is found in an infinitesimal gap between r and $r + \delta r$. This is the case that no points are found within the distance r, and at least one point is found between r and $r + \delta r$. Since the volume V_d of a unit d-ball and its surface area S_{d-1} has a relationship $V_d d = S_{d-1}$, we have

$$\begin{aligned} p(r)\delta r &= \mathrm{Po}(0; \rho V_d r^d) \left\{1 - \mathrm{Po}(0; \rho S_{d-1} r^{d-1} \delta r)\right\} \\ &\approx \exp(-\rho V_d r^d) \left\{1 - \exp(\rho S_{d-1} r^{d-1} \delta r)\right\} \\ &= \exp(-\rho V_d r^d) \cdot \rho S_{d-1} r^{d-1} \delta r \\ &= \exp(-\rho V_d r^d) \cdot \rho V_d d r^{d-1} \delta r \end{aligned}$$

Letting $s = 1/\sqrt[d]{\rho V_d}$ be the scale of the average volume of d-dimensional hypercube per point, we obtain the Weibull distribution

$$p(r; s, d) = \frac{d}{s} \left(\frac{r}{s}\right)^{d-1} \exp\left\{-\left(\frac{r}{s}\right)^d\right\} \tag{4}$$

where s and d correspond to the so-called scale and shape parameters of the Weibull distribution, respectively. This distribution $p(r; s, d)$ has a mode at $r = s\sqrt[d]{(d-1)/d}$. For a fixed dimensionality, the mode depends only on the scale parameter s, which enables us to calculate the point density ρ from the mode.