Morphological Image Segmentation by Local Monotonicity

Joseph H. Bosworth and Scott T. Acton
The Oklahoma Imaging Laboratory
School of Electrical & Computer Engineering
Oklahoma State University
Stillwater, OK 74078
http://OIL.okstate.edu
Email: bosworth@ieee.org

Abstract

Image segmentation is a fundamental task in image processing and multimedia. We propose a general method for image segmentation based upon the relationship between mathematical morphology and the local monotonicity of signals. In this paper, we introduce a two-dimensional generalization of the concept of local monotonicity based on mathematical morphology. A multiscale representation of an image is morphologically generated, wherein the degree of local monotonicity determines scale. Then, a morphological edge detection operator exploits the monotonicity property in performing segmentation. The segmentation allows specification of object scale, edge detail, and contrast and is applicable to object-based video coding.

1. Introduction

Image segmentation is the partitioning of an image into semantically meaningful spatial regions. It is an essential task for many applications including object recognition, tracking, and recently coding and compression with the introduction of the MPEG-4 standard for video. With such an extensive and rapidly increasing variety of applications using a wide range of imagery, it is of fundamental interest to the signal processing community to develop robust image segmentation algorithms.

Here, we present a segmentation technique for two-dimensional graylevel imagery. By employing scale-space techniques and self-dual morphological operators, our algorithm is unbiased towards *a priori* knowledge of specific image content, foreground vs. background intensity, and spatial orientation. Results are shown for the case of two-dimensional graylevel images, but the method is applicable to three-dimensional imagery such as CAT and MRI medical data.

The segmentation is performed in three steps. First, a scale-space representation of the original image is generated by recursive morphological filtering. Next, a morphological edge detection algorithm is used to segment each scale. Finally, the segments at each scale are linked together in a coarse-to-fine manner.

This work was supported by the National Aeronautics and Space Administration under EPSCOR grant NCC5-171.

2. Local monotonicity

2.1 One-dimensional case

The scale-space includes a series of images with varying *degrees* of local monotonicity. In one-dimension (1-D), a signal is locally monotonic of degree n (lomo-n) if and only if it is either non-increasing or non-decreasing within each windowed interval of length n. This definition ensures a number of specific scale properties, which are discussed in further detail in [1]. For example, such a signal possesses constant-valued plateaus separating successive non-increasing and non-decreasing intervals. Also, level-set components, defined by a threshold decomposition of the image, have a minimum size or scale. An example of a 1-D discrete lomo-6 signal is shown in Figure 1, along with a corresponding level-set example.

There is presently no accepted generalization of this definition of local monotonicity to higher dimensions. We desire such a generalization for use in image segmentation. Though 1-D local monotonicity is related to the root signals of median filter, we avoid a 2-D generalization based on median filtering because of problems with oscillations and artifacts [2]. Instead, we propose a generalization based on mathematical morphology.

0-7803-5700-0/99/$10.00©1999 IEEE

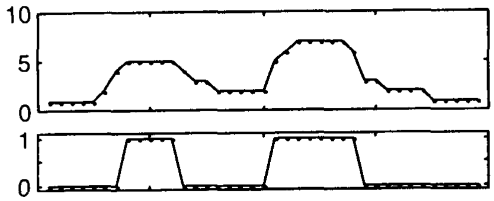


Figure 1. A discrete 1-D lomo-6 signal $f(x)$ and a level-set (binary threshold) for $f(x) \geq 5$.

It is known that a discrete locally monotonic 1-D signal (of degree n) is a root signal of the filters open and close, using symmetric zero-valued (flat) structuring elements of length n-1 [1]. Additionally, the filters open-close (open followed by close) and close-open each produce 1-D locally monotonic signals. These are *idempotent* filters, so further application of the filters effects no further change; a root signal is achieved in one pass.

2.1 Two-dimensional generalization

-In two-dimensions (2-D), a natural extension of the definition of local monotonicity would be to require that such a signal be a root signal of both the open and close filters of the appropriate structuring element. Such a structuring element would be a spatially symmetric (circular) flat disc of diameter n. This definition would properly correspond to the one-dimensional case, but no known filtering procedure would generate such two-dimensional signals. In 2-D, the open-close and close-open filters do not produce signals that are simultaneously roots of both open and close filters. In addition, these filters are dual filters of one another and are inherently biased in intensity direction due to the order of the application of open and close filters. We seek a self-dual filter that generates locally monotonic signals.

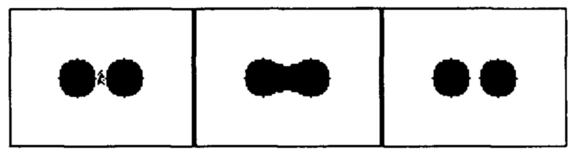
The filter given by the mean of open and close filters provides a solution to these problems. We define the *lomo filter* by the iterate:

$$f(x) \leftarrow \frac{f(x) \circ k(x) + f(x) \bullet k(x)}{2},$$

where the open and close filters of structuring element k(x) as described above are applied to the signal f(x) and their outputs averaged. Our first observation is that a 1-D locally monotonic signal is indeed a root signal of this filter, because it is a root of both open and close. In fact, being a root of this filter serves an equivalent definition of 1-D local monotonicity. This filter is self-dual but not idempotent. However, iterative application of the filter does converge to a root signal. Thus, the lomo filter serves both to define and to generate locally monotonic signals in 1-D.

It is natural to extend this viewpoint to 2-D, by defining a 2-D locally monotonic signal to be a root signal of this lomo filter. Again, the structuring element is spatially symmetric (circular) of diameter n-1. As in the 1-D case, the filter also provides a method for the generation of lomo signals of a given degree. Iterative application of the lomo filter again converges to a root signal.

In 2-D, however, this root signal is not everywhere a root of both the open and close filters simultaneously. While most points in the root signal are *regular* lomo points, *i.e.* unaffected by open or close filtering, there exist *saddlepoint* regions, where the signal possesses points that are local minima in one direction and local maxima in another direction. In these regions, the lomo filter converges to a root signal that is neither a root of the open nor of the close filters individually, but rather a compromise between the two. A simple example of such a situation is shown in Figures 2 and 3.



(a) (b) (c)

Figure 2. A simple 2-D example of a saddlepoint situation. From left to right: (a) original test image, (b) after the open (or close-open) filter of structuring element equal in size to the circular object, (c) after the close (or open-close) filter.

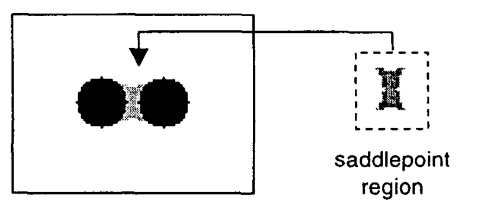


Figure 3. 2-D signal of Fig. 2(a) after lomo filtering (root signal reached after a single pass). The gray region is a saddlepoint region. All other points are regular lomo points.

While the possibility of saddlepoints complicates the 2-D generalization of local monotonicity, it should be noted that no morphological filtering could produce a root signal of both the open and close filters without altering regular lomo points in the region surrounding the saddlepoint. By generalizing local monotonicity to 2-D via the lomo filter rather than by the open and close filters individually, we avoid both the alteration of regular lomo points and the introduction of intensity bias. Also, as we shall see, the most significant 1-D scale properties of locally monotonic signals are generalized to 2-D.

Important scale properties for 2-D lomo signals are similar to those mentioned in Section 2.1 for the 1-D case. The 2-D root signal possesses plateau regions at each local extremum, and disallows level-set components smaller than the structuring element. These well-defined scale properties are advantageous for use in multiscale image processing applications such as segmentation.

A *lomo scale-space* can be generated by creating a series of lomo-n signals of increasing n. Starting with the original signal each scale is generated from the previous scale, rather than directly from the original image. This filtering order is similar to that of alternating sequential filters [4] and effectively removes noise and smaller features as scale is increased.

As shown in [1], faster convergence to a root signal may be obtained by the use of the alternative lomo filter:

$$f \leftarrow \frac{(f \circ k) \bullet k + (f \bullet k) \circ k}{2}.$$

For a given original image, this filter generally produces a slightly different root signal from that of the original lomo filter. However, root signals of this alternative lomo filter are also root signals of the original lomo filter, and differences are subtle. Therefore, due to the improved convergence rate, we generally employ this alternative filter in practical applications. An example of a scale-space generated in this manner is shown in Figure 4.

3. Morphological edge detection

Edge detection is performed on each scale-space layer independently. The detection scheme is similar to a second derivative or Laplacian edge detector. For the lomo-n layer of the scale-space, a self-dual filter defined by the mean of the dilate and erode filters is used to over-smooth sharp changes or edges in the signal:

$$s(x) = \frac{f(x) \oplus k(x) + f(x) \ominus k(x)}{2}.$$

This filter is called the midrange filter [3] and here uses the same structuring element as that used to create the lomo-n signal. Then, zero-crossings are detected on the difference between the lomo-n signal $f(x)$ and this over-smoothed signal $s(x)$.

This formulation, where zero-crossings are detected in the difference image $s(x)$-$f(x)$, may be interpreted as a morphological analogy to the Difference of Gaussians (DoG) or the more general difference-of-lowpass filters methods. Alternatively, it may be viewed as a morphological analogy to the relate Laplacian of Gaussian (LoG) edge detection. The relationship is explained in further detail as follows.

In 1-D, the morphological edge detection may be written as:

$$f''(x) \cong \frac{f(x) \oplus k(x) - 2f(x) + f(x) \ominus k(x)}{2},$$

which is referred to in the literature as the morphological Laplacian [6]. In fact, for lomo-n this reduces to the familiar discrete approximation to the second derivative given by:

$$f''(x) \cong \frac{f(x+\Delta) - 2f(x) + f(x-\Delta)}{2},$$

for any sampling interval $\Delta \leq n/2$. Note that this equivalence between the morphological Laplacian and simple second-derivative approximation relies explicitly on the property of local monotonicity.

In 2-D, the morphological formulation remains unchanged, sampling only three values of the signal. It is therefore more akin to a second directional derivative (in the direction of the gradient), than to a discrete approximation to the linear Laplacian [7]. For convenience and consistency with the literature, however, we will refer to the operation as the the morphological Laplacian.

As does the linear Laplacian, the morphological Laplacian requires a threshold on edge-stregth or gradient magnidude in order to alleviate the detection

Figure 4. Selected levels of a locally monotonic scale-space derived from the original cameraman image of Figure 5. From left, original 256-graylevel image (256 x 256 pixels), scale-space levels corresponding to structuring element diameters of 3, 7, and 11 pixels.

of spurious edges. Here, we apply a threshold to the difference image $s(x)$-$f(x)$ prior to zero-crossing. An example of this thresholding process is shown in Figure 5. Note that this threshold value is scale-invariant, whereas thresholds applied in the linear LoG edge-detection are highly scale-dependent. For example a large-scale step edge retains its magnitude throughout the lomo scale-space, while the same edge decreases in gradient magnitude within the Gaussian scale-space.

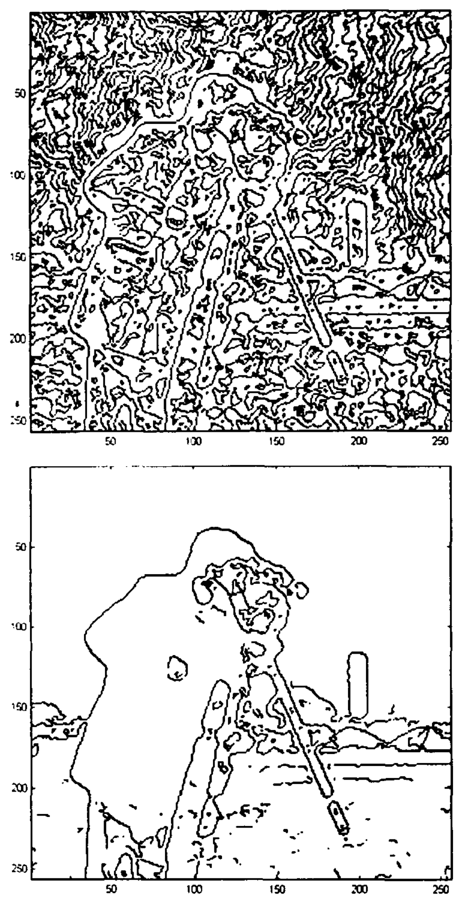


Figure 5. Zero-crossings of morphological Laplacian applied to the scale-space level corresponding to diameter 7 pixels. The top image is without an edge threshold, while the bottom image uses a threshold at graylevel 3.

4. Multiscale segmentation

The zero-crossing procedure with thresholding is used for edge detection. For segmentation, however, the edges detected in this manner should be modified in order to guarantee closed boundary contours between regions. Two additional steps are used to alleviate problems in zero-crossing detection and contour closing.

Even without thresholding, the detected zero-crossings can fail to form closed contours. At junctions between three or more regions, a gap the diameter of the structuring element may occur. This difficulty occurs at edges where the difference image $s(x)$-$f(x)$ is identically zero. To overcome this problem at junctions, and similar gaps caused by thresholding, a scale-determined edge dilation is employed. Edges are dilated by a structuring element of half the diameter used for the scale-space level, then re-eroded to single-pixel width using a small structuring element while avoiding the merger of separate regions. This process is illustrated in Figure 6. Additionally, the difference image $s(x)$-$f(x)$ is slightly blurred prior to zero-crossing detection in order to avoid identically zero intensities.

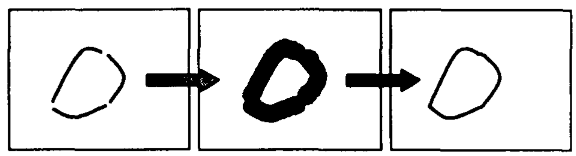


Figure 6. Illustration of single-scale contour closing. From left: zero-crossings of morphological Laplacian at a given scale-space level, scale-determined edge dilation, and final thinning.

Once each scale-space layer is independently segmented, a scale-space linking procedure is used. First, an initial scale-space level is selected where the segmentation represents the desired object scale. Then, a final scale-space level is selected where the desired fine-scale edge detail of segmented regions occurs. Between these scales, regions are linked based on the greatest spatial overlap of segments between consecutive scales, as illustrated in Figure 7. Fine-scale segments linked to the same coarse-scale segment are effectively merged. Thus, the object scale and edge detail can be independently adjusted in the final segmentation. Sample results of this multiscale segmentation procedure are shown in Figure 8.

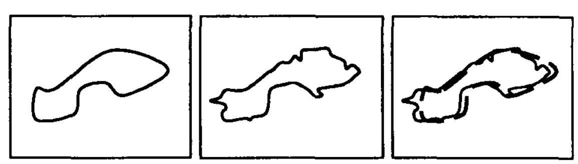


Figure 7. Illustration of multiscale segment linking. From left: a coarse-scale segment, a finer-scale segment, and spatial overlap through scale.

5. Conclusion

By employing scale-space techniques and self-dual filters, the entire segmentation method is very general and requires a minimum of parameters: object scale, edge detail, and contrast threshold. Also, by use of morphological filtering, noise resilience, edge localization, and efficient implementation are achieved. The segmentation method is especially appropriate for object-based coding applications, where control over scale and edge detail are critical. In many object-based coding algorithms, contour information is transmitted separately from intra-segment information. Because of the cost of coding contour detail, it is desirable to allow the selection of edge detail independently from object scale. Our segmentation process allows just this sort of control. Additionally, the method may be directly generalized to 3-D imagery, and we are presently pursuing research into its generalization into color and multi-spectral segmentation.

6. References

[1] J. Bosworth and S. T. Acton, "The morphological lomo filetr for mutliscale image processing," *IEEE International Conference on Image Processing*, ICIP-99, Kobe, Japan, 24-28 Oct., 1999.

[2] A. C. Bovik , "Streaking in median filtered images," *IEEE Trans. Acoust., Speech, Signal Process.*, vol. ASSP-35, pp. 493-503, 1987.

[3] P. Maragos and R.W. Schafer, "Morphological filters--part II: their relations to median, order statistic and stack filters," *IEEE Trans. Acoustics, Speech, Signal Processing*, vol. 35, pp. 1170-1184, 1987.

[4] J. A. Bangham, P. D. Ling, and R. Harvey, "Scale-space from nonlinear filters," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 18, no. 5, pp. 520-527, 1996.

[5] M. A. Schultze and J. A. Pearce, "Linear combinations of morphological operators: the midrange, pseudomedian, and LOCO filters," *IEEE Conference on Acoustics, Speech, and Signal Processing*, ICASSP-93, vol. 5, pp. 57-60, 1993.

[6] L. J. Van Vliet, I. T. Young, and G. L. Beckers, "A nonlinear Laplace operator as edge detector in noisy images," *Computer Vision, Graphics, and Image Processing*, vol. 45, pp. 167-195, 1989.

[7] R. M. Haralick, "Digital step edges from zero crossings of second directional derivatives," *IEEE Trans. Pattern Anal. Mach. Intell.*, PAMI-6, No. 1, pp. 56-68, 1984.

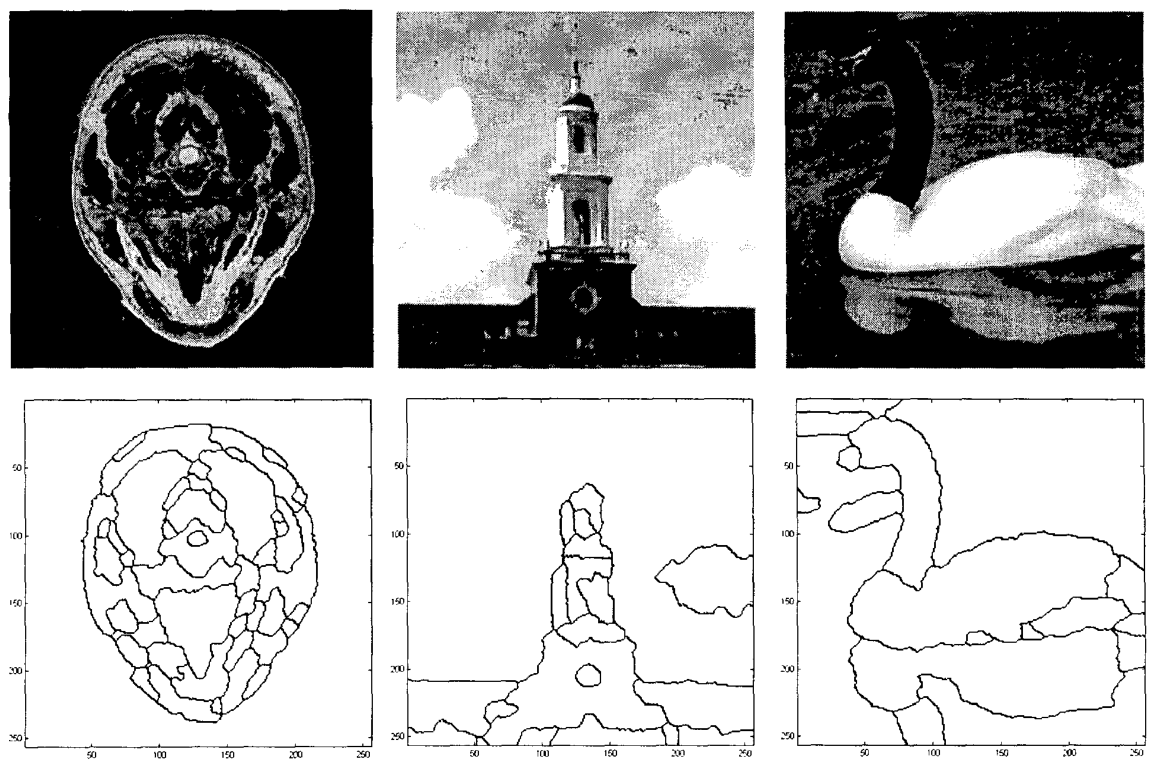


Figure 8. Examples of multiscale segmentation results. All original images are 256 graylevel, 256 x 256 pixel images. Initial scales (structuring element radii in pixels), final scales, and edge thresholds are as follows: 'brain' image: (2, 2, 11), 'library' image: (3, 2, 4), and 'swan' image: (4, 2, 3).