Will appear in *Proceedings of the IEEE International Joint Conference On Neural Networks,* 1998

Extracting Hydrographic Objects from Satellite Images Using a Two-layer Neural Network

Xiuwen Liu[1,2] DeLiang Wang[1,3] J. Raul Ramirez[2]
[1]*Department of Computer and Information Science*
[2]*Center for Mapping* [3]*Center for Cognitive Science*
The Ohio State University, Columbus, Ohio 43210, USA
{liux, dwang}@cis.ohio-state.edu {xiuliu raul}@cfm.ohio-state.edu

Abstract

This paper presents a two-layer network for extracting hydrographic objects, such as rivers, from satellite images. The first layer is a locally connected network, which performs nonlinear smoothing. A unique property of the network is that boundaries and junctions are preserved with high accuracy while noise within each region is greatly suppressed. A second layer is a locally excitatory globally inhibitory oscillator network (LEGION), which extracts the desired objects. The seeds of objects are selected separately. To find hydrographic objects, seed points are automatically identified from the original image, based on the assumption that water bodies are homogenous. Computationally, this approach is parallel and local and can be effectively implemented using hardware directly, the efficiency of which may provide a potential solution for real-time image processing. Experimental results using digital orthophoto images are provided.

1. Introduction

With the increasing availability of high-resolution digital orthophotography and satellite imagery, image processing techniques become critically important in providing accurate up-to-date spatial data. A fundamental issue is to group similar input elements together while segregating different groups. This problem is known as image segmentation, one of the central problems in machine vision. One reason for the difficulties is that images must be grouped into semantically meaningful entities. Specifically for the task here, pixels corresponding to hydrographic objects should be grouped together, while other undesired regions should be suppressed into the background. Because local pixel attributes are intrinsically ambiguous, i.e., pixels belonging to the same object can vary considerably while pixels belonging to different objects may have similar values, approaches based on local attributes such as various thresholding methods and classification techniques have only limited success. Also hydrographic objects are shapeless and a generic model cannot be defined, and thus template matching techniques including line following algorithms, active contour models [5] and deformable template matching techniques [4] cannot be applied effectively.

In this paper, we propose a two-layer network. The first layer is a locally connected network, which makes pixel values more reliable through nonlinear smoothing. Connection weights between two nodes are updated nonuniformly as a function of weights within their receptive fields. Information, which is encoded through connection weights, is propagated and thus salient boundaries are preserved while details within one region are smoothed out. The second layer is a LEGION network [16][20][21], which provides a computationally effective framework for image segmentation based on temporal correlation [8][18]. Unlike other oscillatory networks [6][15], LEGION has been shown analytically [16][21] that it converges rapidly with a local and parallel architecture and has been applied to segment real images [21].

The paper is organized as follows. Section 2 proposes a locally connected network for nonlinear smoothing after a brief review of nonlinear filters. Section 3 describes the LEGION network which is used for image segmentation. Section 4 presents experimental results for synthetic and real images. Finally, Section 5 discusses some advantages and potential limitations of the proposed network.

2. A network for nonlinear smoothing

2.1 A brief review on nonlinear filters

To make local attributes more reliable, noise that is introduced during image registration must be suppressed. Linear filters replace each pixel value by a weighted average from its local neighborhood. The main drawback of linear filters is that image structures will be degraded while noise is reduced [1]. To overcome this problem, nonlinear filters are proposed. Essentially, nonlinear filters try to adaptively change the weights in the smoothing neighborhood so that desired features such as edges, corners, and junctions are preserved by preventing smoothing across object boundaries. Obviously, the boundaries of the desired objects are not known beforehand (otherwise, the problem would have been solved already), and thus most nonlinear filters first estimate the evidence of the presence of desired features and incorporate the estimated evidence by changing the weights accordingly.

Therefore, the central issues in nonlinear filter design are how to estimate the presence of desired features and how to change the weights adaptively according to the estimation. To estimate the presence of edges, a popular choice is gradient. Therefore, many nonlinear filters are gradient-based. Lev et al. [7] proposed three similar methods of noise reduction by reflecting the local discontinuity through weight configurations to avoid blurring image structures. Tomita and Tsuji [17] proposed a method which replaces each pixel by the average of the most homogeneous neighborhood among the five overlapped rectangular ones. Nagao and Matsuyama [11] improved the result by using nine elongated bar masks instead of five rectangular masks. Gradient inverse-weighted smoothing filters [19] use the inverse of the gradient as the weighting factors. Recently, Perona et al. [12][13] posed image smoothing as anisotropic diffusion. The basic idea is to prefer diffusion along the direction where the gradient is small, making the diffusion nonuniform. While impressive results are obtained, there are several fundamental limitations. First the system converges to a uniform state and thus reasonable results exist only within a certain range of iterations, corresponding to scale selection in Gaussian linear scale space. Also in general, many iterations may be needed and the number of iterations for best results varies considerably from image to image. In addition, gradient is estimated in general from a very local window, making it difficult to distinguish salient boundaries from noise edge segments. Smith and Brady [14] proposed a filter that modulates the Gaussian smoothing with brightness smoothing. While the authors claimed that their filter combines the advantages of existing noise reducing filters, the filter is not effective when noise is large. For quantitative comparisons between selective nonlinear filters, see Chin and Yeh [3] and Smith and Brady [14].

2.2 A nonlinear smoothing filter family

Consider the Mumford-Shah energy functional [10], which Morel and Solimini [9] considered as a generic functional to unify many existing segmentation approaches:

$$E(u, K) = \int_{\Omega \backslash K} \left(|\nabla u|^2 + (u-g)^2 \right) dx + length(K) \qquad (1)$$

where g is the original image, u is the desired solution and K is the edge sets. For image segmentation, we want u to be piecewise constant. Because the last term is a global measure, it is computationally expensive to find the optimal or near optimal solutions. To overcome this difficulty, we instead use a local measurement at each pixel location. It is well known that grouping should be orientation sensitive and we assume that each pixel must be locally grouped with one of orientation sensitive subreceptive fields. Under this assumption, in order to actually minimize the corresponding energy functional, each pixel must be grouped with the subreceptive field that is most homogenous, which is consistent with [17][11].

Thus we obtained a nonlinear smoothing filter family. At each pixel location, we associate a set of orientation sensitive subreceptive fields. Smoothing is done only within the subreceptive field which is most homogenous.

2.3 A locally connected network

The proposed nonlinear filter family can be implemented using a locally connected network. Each node is associated with a receptive field and connected with other nodes in its receptive field. The receptive field is divided into orientation sensitive windows, which are called subreceptive fields. Initially we define the connection weight between node i and j as $w_{ij} = p_j - p_i$, where p's are the pixel values at the corresponding locations. The updating rule for the filter is given by:

$$\Delta w_{ij} = \frac{\sum_m e^{-2\rho V_j^{(m)}} A_j^{(m)}}{\sum_m e^{-2\rho V_j^{(m)}}} - \frac{\sum_m e^{-2\rho V_i^{(m)}} A_i^{(m)}}{\sum_m e^{-2\rho V_i^{(m)}}} \qquad (2)$$

where

$$A_i^{(m)} = \frac{\sum_{l \in R_i^{(m)}} w_{il}}{\left|R_i^{(m)}\right|}$$

and

$$V_i^{(m)} = \frac{\sum_{l \in R_i^{(m)}} \left|w_{il} - A_i^{(m)}\right|^2}{\left|R_i^{(m)}\right|}$$

Here $R_i^{(m)}$ is the mth subreceptive field and $|R|$ is the Cardinality of set R. Similar to soft-max algorithms, the parameter ρ is chosen to be large so that (2) actually chooses the most homogenous subreceptive field.

There are two conceptual steps involved in (2). The pixel value is first estimated statistically within a receptive field. Without prior knowledge, more reliable estimation can only be obtained through more samples. Here the assumption is used that the desired signal is piecewise constant. Then the signal is further estimated by nonuniformly integrating estimation from subreceptive fields. It can be shown easily that (2) will converge to a piecewise constant state, where an autonomous solution is feasible.

Equation (2) is a generic form for nonlinear filters. If we assume that the weight is a Gaussian function and the entire receptive field as one subreceptive field, we obtain a Gaussian filter. If we assume that the weight is simply the reciprocal of the connection weight, we obtain gradient inverse-weighted filter [19].

To demonstrate the effectiveness of the proposed nonlinear filter, we use a one-dimensional signal. In this case, there are two subreceptive fields, left and right windows. Figure 1(a) shows the ground truth signal, which is corrupted by Gaussian noise, as shown in Fig. 1(b). Figure 1(c) shows the result using a Gaussian smoothing filter. It is evident that the step edge was blurred and became a ramp. Figure 1(d) shows the result using the proposed method for 10 iterations. The step edge was preserved while noise was greatly reduced.

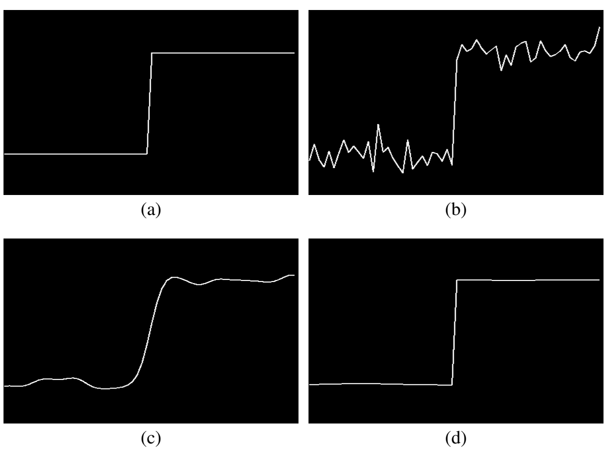


Fig. 1. Comparison of Gaussian filter and locally connected smoothing network for a one-dimensional signal. (a) Ground truth signal. (b) Signal corrupted by noise. (c) Result from a Gaussian smoothing filter. (d) Result from the proposed network after 10 iterations.

For two dimensional images, we need to specify subreceptive fields. To be easily implemented, each subreceptive is defined as an elongated rectangular window, given by W and L. Another parameter N defines the angular resolution. Subreceptive fields are obtained as rotated copies of some basis rectangular subreceptive fields.

3. Image segmentation

To extract hydrographic objects, pixels need to be grouped together to form semantically meaningful entities such as rivers. This is achieved using a LEGION network, which consists of locally coupled identical relaxation oscillators and a global inhibitor. A single oscillator i is modeled as a feedback loop between an excitatory unit x_i and an inhibitory unit y_i [21]:

$$\frac{dx_i}{dt} = 3x_i - x_i^3 + 2 - y_i + I_i H(p_i - \theta) + S_i + \rho \quad (3a)$$

$$\frac{dy_i}{dt} = \varepsilon(\gamma(1 + \tanh(x_i/\beta)) - y_i) \quad (3b)$$

where H stands for the Heaviside step function. I_i represents external stimulation for oscillator i, which is assumed to be applied at time 0. ρ denotes a Gaussian noise term to test the robustness of the system and play an active role in desynchronization. To suppress undesired noisy fragments and regions, p_i, a *potential* of the oscillator i, is introduced. Seed selection can be achieved by tuning θ, which is a threshold and $0 < \theta < 1$.

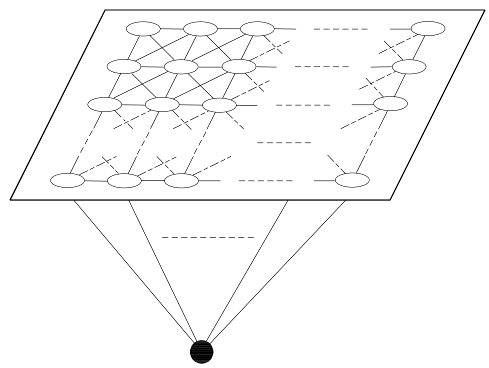


Fig. 2. Architecture of a two-dimensional LEGION network with eight-nearest neighbor coupling. An oscillator is indicated by an ellipse and the global inhibitor by a filled circle.

The term S_i denotes the coupling from other oscillators in the network, which is defined as:

$$S_i = \sum_{k \in N_c(i)} W_{ik} H(x_k - \theta_x) - W_z H(z - \theta_x) \quad (4)$$

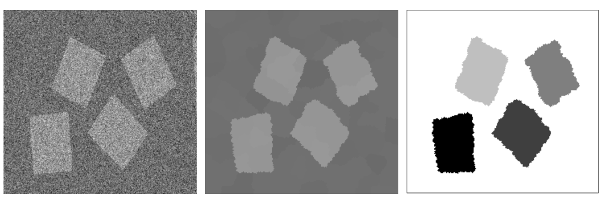
where θ_x is a threshold, and $N_c(i)$, the *coupling neighborhood* of i, is a set of neighboring oscillators of i. W_{ik} is the connection weight from oscillator k to i, defined as

$$W_{ik} = \frac{C}{|w_{ik}| + 1} \quad (5)$$

where C is a constant and w_{ik} is the connection weight from i to k in the first layer. In (4), W_z is the weight of inhibition from the global inhibitor z, defined by $dz/dt = \phi\ (\sigma_\infty - \theta_z)$. Here, ϕ is a parameter, $\sigma_\infty = 1$ if $x_i \geq \theta_z$ for at least one oscillator i, and $\sigma_\infty = 0$ otherwise, where θ_z is a threshold. The network architecture used for image segmentation is a locally coupled network with inhibitory connections to the global inhibitor, as shown in Fig. 2.

4. Experimental results

Heavy computation is involved when the two-layer network is simulated on serial computers. To avoid that, we derived an algorithmic version which is suitable for implementation. An algorithm for the first layer can be derived easily. For the segmentation results shown in this paper, the smoothing images were obtained by running 10 iterations. For the second layer, a LEGION algorithm derived by Wang and Terman [21] is used with a separate seed selection procedure for digital orthophoto images.



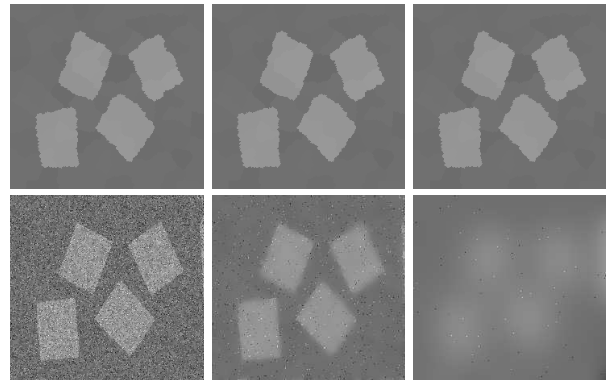
(a) (b) (c)

Fig. 3. Segmentation result for a synthetic image. (a) Original image. (b) Smoothed image after 10 iterations, where *W* = 10, *L* = 3 and *N* = 8. (c) Segmentation result represented as a gray map. Images are 240 x 230 pixels.

4.1 A synthetic image

We used a noisy synthetic image to demonstrate the segmentation capability and do a comparison with anisotropic diffusion. Figure 3(a) shows a noisy image, which consists of four corrupted rectangles. Figure 3(b) shows the smoothed image using the proposed method. Salient boundaries are preserved while detail segments due to noise are smoothed out. Figure 3(c) shows the segmentation result using a gray map [21], where objects with different phases are encoded through distinct gray levels.

As a comparison, we also applied anisotropic diffusion to the image shown in Fig. 3(a). The diffusion was implemented using a program available in [13]. Figure 4 shows the results along with the results from the proposed method after the same number of iterations. It shows clearly that the proposed filter converges very quickly to a piecewise constant state while anisotropic diffusion approaches to a uniform state slowly. While 100 iterations of anisotropic diffusion produces a good result, a separate procedure or an operator is needed to detect that.



(a) 10 iterations (b) 100 iterations (c) 1000 iterations

Fig. 4. Comparison between the proposed filter and anisotropic diffusion. The top row shows the results from our proposed filter and the bottom shows the results from anisotropic diffusion with the same number of iterations.

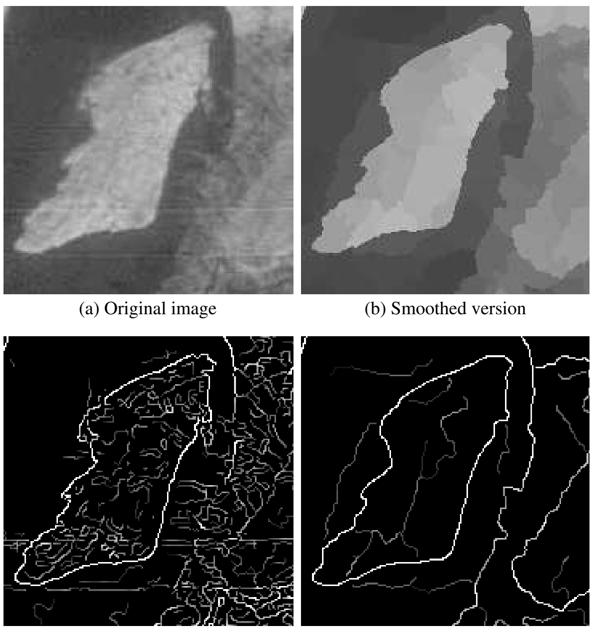
4.2 Digital orthophoto images

Digital orthophoto images are obtained from high resolution satellite images by compensating variations in actual pixel sizes on ground due to perspective projection using a nonlinear transformation. The real images used in this experiment were taken from digital orthophoto quadrangles provided by United States Geological Survey. The images were first smoothed by the proposed filter family for 10 iterations with the following parameters: $W = 10$, $L = 3$, and $N = 8$. The smoothed images were then segmented. For hydrographic features, seed pixels are identified, based on the assumption that water bodies in the image are homogenous. Because rivers and streams tend to be narrow and long, elongated masks are used to calculate the potential. Seed selection was implemented by applying the masks along different orientations, and using the most homogenous mask to calculate the potential. This is almost identical to the proposed nonlinear smoothing filter.

Figure 5(a) shows an image with a river which has several branches. Figure 5(b) shows the smoothed image. We can see clearly that the regions became more homogenous and noise was reduced. Figure 5(c) shows the segmentation result superimposed on the original image. The river was segmented and boundaries were marked with high accuracy. The small island in the upper part of the river was smoothed out because it is smaller than the size of the filter kernel.

(a) (b) (c)

Fig. 5. A digital orthophoto image with a river with several branches. (a) The original image. (b) The smoothed image. (c) The segmented river marked as white and imposed on (a). Images are 512 x 512 pixels.



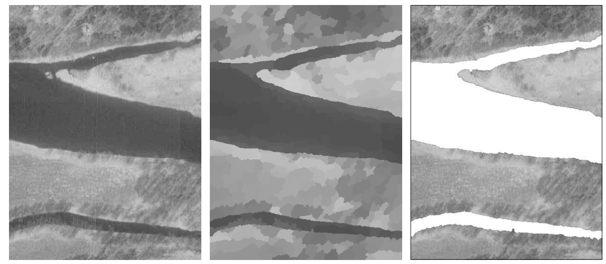
(a) Original image (b) Smoothed version

(c) Edge map for (a) (d) Edge map for (b)

Fig. 6. Smoothing effects on a real image. (a) A small portion from Fig. 5(a). (b) The corresponding portion from Fig. 5(b). (c) The edge map of (a) using a Canny edge detector. (d) The corresponding edge map for (b). Images are 181 x 181 pixels.

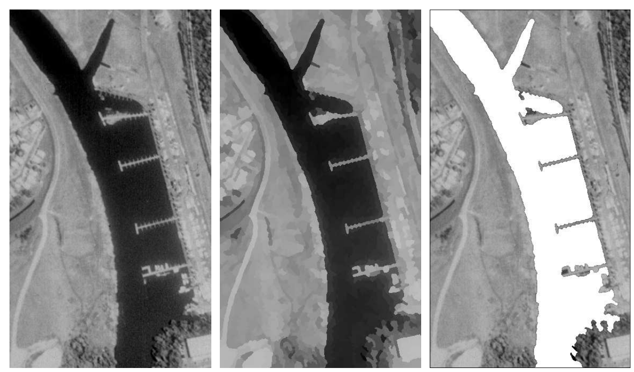
To show the effectiveness of our proposed filter, Fig. 6(a) shows a small portion from Fig. 5(a) and the corresponding portion of the smoothed image is shown in Fig. 6(b). Figures 6(c) and 6(d) show the corresponding edge maps by a Canny edge detector [2] with $\sigma = 1.0$ and the same hysteresis thresholds. Most salient edges in Fig. 6(c) are preserved with high accuracy in Fig. 6(d) while noisy segments are suppressed greatly. The river boundaries became continuous salient edges in Fig. 6(d) while they were corrupted into noisy segments in Fig. 6(c). This also shows clearly that grouping based on edge elements from an edge detector would be difficult to form salient boundaries.

Figure 7 shows an image with narrow rivers. Figure 7(a) shows the original image, consisting of a river and a narrow stream. Figure 7(b) shows the smoothed image, where the river boundaries are enhanced without location displacement. Figure 7(c) shows the result. Here two river objects with different phases in LEGION are marked as white to show the accuracy of the boundaries. Because the elongated masks were used to calculate potentials, the narrow stream was correctly segmented.



(a) (b) (c)

Fig. 7. A river with a narrow branch and a stream. See caption for Fig. 5 for arrangement. Images are 384 x 512 pixels.



(a) (b) (c)

Fig. 8. A river with complicated boundaries. See caption for Fig. 5 for arrangement. Images are 378 x 670 pixels.

Figure 8(a) shows a more complicated image with small objects which make the river boundaries irregular. Figure 8(b) shows the smoothing result. We can see that the linear features are preserved well, except that the tiny details were smoothed out, demonstrating the feature preserving and noise reduction property of the proposed filter. Figure 8(c) shows the segmentation result. The system segmented the river correctly even along noisy banks in the bottom part. Also pixel values within the river vary considerably and a classification technique would have difficulty to handle the variations.

5. Conclusions

In this paper we proposed a two-layer neural network and applied it to extract hydrographic objects in digital orthophoto images. To avoid heavy computation on serial computer simulations, algorithms that were extracted from the network were used for segmenting real images. The proposed filter integrates a local grouping for pixel value estimation and nonlinear smoothing together. Because the system converges to a piecewise constant state, it overcomes some of the limitations of the anisotropic diffusion method [12][13]. The experimental results also showed that salient edges are preserved without position displacement while noisy edge segments are smoothed out. Segmentation was achieved using a LEGION network, which offers a neural and parallel solution for image segmentation.

Methodologically, this approach has several major advantages. First, the flexible representation of temporal labeling in the oscillatory network promises to provide a generic solution to image segmentation. Our network could be applied to identify other objects, such as road networks when appropriate seed selection is incorporated. Secondly, the system is guaranteed to be very efficient through local and parallel architectures. Finally, it is a dynamic system with parallel and distributed computation. The unique property of neural plausibility makes it particularly feasible for analog VLSI implementation, the efficiency of which may pave the way to achieve real-time image segmentation.

There are several limitations of this approach when applied to arbitrary images. Because the method is based on the piecewise constant assumption, the system may produce unsatisfactory results if the assumption does not hold. Another problem is that the proposed filter tends to smooth out small island features, even though they may be salient, as shown in Fig. 5. One way to overcome that is to represent the results in a hierarchical manner. When major boundaries are identified, we then reduce the filter size to preserve island features.

Acknowledgments This work is partially supported by an ONR grant (N00014-93-1-0335), an NSF grant (IRI-9423312) and an ONR Young Investigator Award to DLW (N00014-96-1-0676). The authors would also like to thank National Imagery and Mapping Agency (NIMA) and The Ohio State Univeristy Center for Mapping for supporting this research.

References

[1] S. T. Acton, "Image restoration using generalized deterministic annealing," *Digital Signal Processing*, vol. 7(2), pp. 94-104, 1997.

[2] J. Canny, "A computational approach to edge detection," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 8(6), pp. 679-698, 1986.

[3] R. T. Chin and C. L. Yeh, "Quantitative evaluation of some edge-preserving noise-smoothing techniques," *Computer Vision, Graphics, and Image Processing*, vol. 23(1), pp. 67-91, 1983.

[4] A. K. Jain, Y. Zhong, and S. Lakshmanan, "Object matching using deformable templates," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 18(3), pp. 267-277, 1996.

[5] M. Kass, A. Witkin, and D. Terzopoulos, "Snakes: Active contour models," *International Journal of Computer Vision*, vol. 1, pp. 321-331, 1988.

[6] P. Konig and T. B. Schillen, "Stimulus-dependent assembly formation of oscillatory responses: I. Synchronization," *Neural Computation*, vol. 3, pp. 155-166, 1991.

[7] A. Lev, S. W. Zucker, and A. Rosenfeld, "Iterative enhancement of noisy images," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7(6), pp. 435-442, 1977.

[8] P. M. Milner, "A model for visual shape recognition," *Psychological Review*, vol. 81(6), pp. 521-535, 1974.

[9] J. M. Morel and S. Solimini, *Variational Methods for Image Segmentation: With Seven Image Processing Experiments*. Birkhauser, Boston, 1995.

[10] D. Mumford and J. Shah, "Optimal approximations of piecewise smooth functions and associated variational problems," *Communications on Pure and Applied Mathematics*, vol. 42(4), pp. 577-685, 1989.

[11] M. Nagao and T. Matsuyama, "Edge preserving smoothing," *Computer Graphics and Image Processing*, vol. 9(4), pp. 394-407, 1979.

[12] P. Perona and J. Malik, "Scale space and edge detection using anisotropic diffusion," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 12(7), pp. 16-27, 1990.

[13] P. Perona, T. Shiota, and J. Malik, "Anisotropic diffusion," In *Geometry-Driven Diffusion in Computer Vision*, edited by B. M. ter Haar Romeny, Kluwer Academic Publishers, Dordrecht, pp. 73-92, 1994.

[14] S. M. Smith and J. M. Brady, "SUSAN - a new approach to low level image processing," *International Journal of Computer Vision,* in press, 1997 (also available at *www.robots.ox.ac.uk/~steve*).

[15] M. Stoecker, H. J. Reitboeck, and R. Eckhorn, "A neural network for scene segmentation by temporal coding," *Neurocomputing*, vol. 11(2-4), pp. 123-134, 1996.

[16] D. Terman and D. L. Wang, "Global competition and local cooperation in a network of neural oscillators," *Physica D,* vol. 81(1-2), pp. 148-176, 1995.

[17] F. Tomita and S. Tsuji, "Extraction of multiple regions by smoothing in selected neighborhoods," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 7(6), pp. 107-109, 1977.

[18] C. von der Malsburg, *The Correlation Theory of Brain Function*, Internal Report 81-2, Max-Planck-Institute for Biophysical Chemistry, 1981.

[19] D. C. C. Wang, A. H. Vagnucci, and C. C. Li, "Gradient inverse weighted smoothing scheme and the evaluation of its performance," *Computer Graphics and Image Processing*, vol. 15, pp. 167-181, 1981.

[20] D. L. Wang and D. Terman, "Locally excitatory globally inhibitory oscillator networks," *IEEE Transactions on Neural Networks,* vol. 6(1), pp. 283-286, 1995.

[21] D. L. Wang and D. Terman, "Image segmentation based on oscillatory correlation," *Neural Computation*, vol. 9(4), pp. 805-836, 1997.