In *Proceedings of GIS/LIS*, 1997, pp. 365-373

AUTOMATIC EXTRACTION OF HYDROGRAPHIC OBJECTS IN DIGITAL ORTHOPHOTO IMAGES

Xiuwen Liu[†‡] and J. Raul Ramirez[‡]

[†]*Department of Computer and Information Science*

[‡]*The Center for Mapping*

The Ohio State University

1216 Kinnear Road, Columbus, OH 43212, USA

{xiuliu, raul}@ra.cfm.ohio-state.edu

ABSTRACT

This paper presents a new approach to extracting hydrographic objects such as rivers, lakes, and other water bodies from digital orthophoto images. A family of directional-sensitive operators is derived and applied to suppress the noise while preserving the desired features. The unique property of this operator family is that boundaries and junctions are preserved with high accuracy while noise within each object is greatly suppressed. A region-based grouping technique, which was derived from locally excitatory globally inhibitory network (LEGION) dynamics, is then applied to extract desired objects, the seeds of which 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. This approach is general and can be used to find other objects when appropriate seed selections are incorporated. Computationally, this approach is parallel and local in nature and can be efficiently implemented using a neural network. Experimental results using real 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 GIS data. A fundamental issue is to group similar input elements together while segregating different groups. This problem is known as image segmentation in machine vision. One reason for the difficulties is that

The authors would like to thank Prof. DeLiang Wang for many valuable ideas and comments and Ms. Beverly Henderson for making the paper more readable.

images must be segmented into semantically meaningful entities. Specifically for the task here, hydrographic objects in digital orthophoto images 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 the same values, approaches based only on local attributes such as various thresholding methods (Castleman, 1996) have only limited success. In addition, noise introduced during image registration makes local attributes unreliable, and thus, worsens the problem.

Therefore, first we need to recover local attributes that are corrupted by noise. In this paper, after we justify a general assumption about the underlying structures, we propose a directional feature-preserving filter family. When noise is reduced, we need to actually identify the hydrographic objects through grouping. Many techniques have been proposed to solve the segmentation problem (see Haralick and Shapiro (1985), Lowe (1985), Sarkar and Boyer (1993), and Pal and Pal (1993) for reviews on the subject). Very recently, Terman and Wang (Terman and Wang, 1995; Wang and Terman, 1995, 1997) proposed a LEGION network as a representation framework for image segmentation, based on temporal correlation theory (Milner, 1974; von der Malsburg, 1981). Unlike other methods, it has been shown that it converges very rapidly (Terman and Wang, 1995) even though the network architecture is very local and parallel. Here we use the algorithm derived based on the dynamics of the system for real images.

The paper is organized as follows. Section 2 proposes a directional feature-preserving filter family after a review of nonlinear smoothing filters. Section 3 describes the grouping technique. Section 4 presents experimental results for synthetic and real images. Finally, Section 5 discusses advantages and some potential limitations of the proposed method and issues for future research.

2. DIRECTIONAL FEATURE-PRESERVING SMOOTHING OPERATOR FAMILY

2.1 A Brief Review on Nonlinear Smoothing Filters

To make local attributes more reliable, digital orthophoto images must be enhanced. Linear filters, including the most widely used Gaussian filter, replace each pixel value by a weighted average from its local neighborhood. The main drawback of linear filters is that image structures, mainly edges, will be degraded when noise is reduced (Canny, 1986). To overcome this problem, nonlinear filters are proposed. Essentially, most nonlinear filters try to adaptively change the weights in the smoothing neighborhood so that desired features such as edges, corners, or 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). To overcome this difficulty, the evidence of the presence of desired features is first estimated and then the estimated evidence is incorporated into the filter 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. Some nonlinear filters, such as order-statistic neighbor operators, do not perform the estimation explicitly; their assumptions about the underlying structures reflect that. For example, the median filter (Castleman, 1996) assumes that the median of the smoothing window is closest to the true value of the central pixel. To estimate the presence of edges, a good and popular choice is gradient. Therefore, many nonlinear filters are gradient-based. Lev et al. (1977) proposed three similar methods of noise reduction by reflecting the local discontinuity through weight configurations to avoid blurring image structures. Gradient inverse-weighted smoothing filter (Wang et al., 1981) uses the inverse of the gradient as the weighting factors. Recently, Perona (Perona and Malik, 1987; Perona et al., 1994) posed the image smoothing problem as anisotropic diffusion. The basic idea is to prefer diffusion along the direction where the gradient is small. In general, many iterations may be needed and the final results tend to be piecewise constant (Perona et al., 1994).

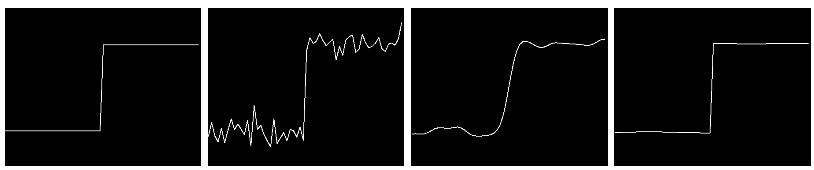
If we explicitly assume that the underlying structures in images are piecewise constant, which corresponds to the step edge model in designing many edge detectors, including the Canny edge detector (Canny, 1986), estimating the presence of edges can be achieved through homogeneity testing. Tomita and Tsuji (1977) proposed a method which replaces each pixel by the average of the most homogeneous neighborhood among the five overlapped rectangular ones. Nagao and Matsuyama (1979) improved the result by using nine elongated bar masks instead of five rectangular masks. These methods are computationally efficient and effective in noise reduction. However, two-dimensional features, such as corners and junctions could be smoothed. The assumption of piecewise constancy is also used by Mumford and Shah (1989) to propose a generic energy functional for image segmentation, which was shown by Morel and Solimini (1995) to be a unified model for many edge detection and image segmentation techniques. For a quantitative comparison of some of the nonlinear filters, see Chin and Yeh (1983) and Smith and Brady (1997).

2.2 Directional Feature-Preserving Filter Family

Based on the observation that the hydrographic features are locally linear and the underlying structures are piecewise constant, we propose a family of feature-preserving smoothing operators suitable to digital orthophoto image processing. To illustrate the main points, we first consider the one dimensional case. Suppose we have an ideal step edge as shown in Fig. 1(a) which is corrupted by adding 10% Gaussian noise. The resulting signal is shown in Fig. 1(b). Figure 1(c) shows the smoothing result using a Gaussian filter with σ=2 and *W*= 5, given by:

$$G(k) = \frac{e^{-k^2/2\sigma^2}}{\sum_{i \in [-W, W]} e^{-i^2/2\sigma^2}} \tag{1}$$

Here, G is defined only on [-*W*, *W*], and *i* and *k* are integers. It is evident that the step edge becomes a ramp, even though the noise was not greatly reduced. Because the filter is structurally symmetrical, the filter inevitably smooths across different regions, which leads to the blurring effects of edges. Also, the filter is not able to take advantage of the explicit assumption that the true signal is piecewise constant.



(a) (b) (c) (d)

Fig. 1. Comparison of Gaussian filter and feature-preserving filter 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 filter.

These observations lead to a structurally unsymmetrical filter family. If we can estimate the boundaries between regions, the problem becomes to recover a constant signal within each region. Then by viewing all of the points within a region as random samples of a constant signal, we can statistically estimate the true value using an equally weighted average filter. Obviously, the boundaries are not known in advance. But with the additional assumption that the underlying signal is piecewise constant, a good estimation can be obtained by homogeneity testing. Specifically, if all points are random samples of a constant, they should have similar local statistical quantities, such as average and variance. This leads to a filter family which consists of three equally weighted filters: left window [-*W*, 0], right window [0, *W*], and full window [-*W*, *W*]. The following algorithm can be used to recover the true signal.

Procedure Selective-smoothing
For each point in the signal, do the following:
    calculate the statistical quantities of left and right windows
    if those quantities are similar
        locally use the full window filter
    else

use the filter of the window which is more homogenous
endif
endfor

It is easy to see that this procedure can be repeated to improve the result incrementally. Figure 1(d) shows the smoothing result after 10 iterations. In that implementation, average and variance from each window are used. Variance is used to indicate homogeneity. It is clear that the step edge was preserved while noise was reduced greatly. Furthermore, this algorithm is not sensitive to the filter parameter because smoothing is restricted to one region with same underlying signal and is propagated only within that region. Another important property of the algorithm is quick convergence. Because of the equally weighted averaging, no spurious local maximum and minimum can occur. This assures that the system will approach a piecewise constant state.

To extend the algorithm to two dimensional cases, we need two additional parameters. Now each window needs two parameters to specify its size: *W* and *H*. Alternatively, we can specify the size with *W* and the ration between *H* and *W*. As shown by Canny (1986), the elongated kernel performs better for linear features. Another parameter is to specify the orientations along which the filters will be applied. An obvious choice is to apply the filters at several orientations and use the filter with the most homogenous window. This can be specified by *N*, the number of quantized orientations. Mathematically, we obtained the following smoothing operator family.

$$P_n^+(k_x, k_y) = \frac{F(k_x, k_y)}{\sum\limits_{\substack{R_x(i_x, i_y, \theta_n) \in [0, W] \\ R_y(i_x, i_y, \theta_n) \in [0, H]}} F(i_x, i_y)} \tag{2a}$$

$$P_n^-(k_x, k_y) = \frac{F(k_x, k_y)}{\sum\limits_{\substack{R_x(i_x, i_y, \theta_n) \in [0, W] \\ R_y(i_x, i_y, \theta_n) \in [-H, 0]}} F(i_x, i_y)} \tag{2b}$$

Here, $\theta_n = 2n\pi/N$, i_x and i_y are integers within [-*W*, *W*]. *F* can generally be any low- pass filter function. P_n^+ and P_n^- are only defined in the region where the corresponding conditions are satisfied. R_x and R_y are rotation functions, given by

$$R_x(x, y, \theta) = x\cos\theta + y\sin\theta \tag{3a}$$

$$R_y(x, y, \theta) = -x\sin\theta + y\cos\theta \tag{3b}$$

Figure 2 shows a directional feature-preserving smoothing family where F is an elongated Gaussian function. The two basis filters are structurally unsymmetrical elongated kernels. Then rotated copies of the basis filters are generated and normalized so that the sum is always unit.

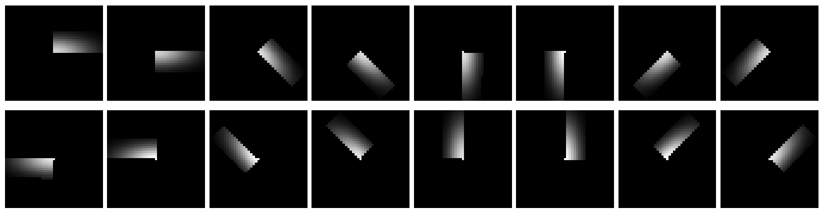
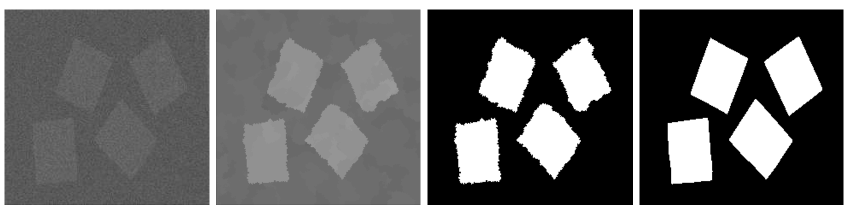


Fig. 2. A directional feature-preserving filter family with W=10, H=3, and N=8.

Figure 3 shows the smoothing result for a synthetic image. Figure 3(a) is the noisy input image, the ground truth of which is shown in (d). Figure 3(b) shows the results after six iterations. Boundaries are preserved with high accuracy. Corners of the rectangles are also preserved well.



(a) (b) (c) (d)

Fig. 3. Segmentation result for a noisy synthetic image. (a) Original noisy image. (b) Smoothed image. (c) Segmentation result. (d) Ground truth image.

3. IMAGE SEGMENTATION

After smoothing, pixels need to be grouped together to form semantically meaningful entities such as rivers. This was achieved using an algorithm derived from LEGION dynamics proposed by Terman and Wang (Terman and Wang, 1995; Wang and Terman, 1995, 1997) for image segmentation. In the original LEGION network, a single oscillator i is modeled as a feedback loop between an excitatory unit x_i and an inhibitory unit y_i (Wang and Terman, 1997):

$$\frac{dx_i}{dt} = 3x_i - x_i^3 + 2 - y_i + I_i\, H(p_i - \theta) + S_i + \rho \qquad (4a)$$

$$\frac{dy_i}{dt} = \varepsilon(\gamma(1 + \tanh(x_i/\beta)) - y_i) \tag{4b}$$

Here, H stands for the Heaviside step function. I_i represents external stimulation for the 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, the *potential* of the oscillator i, is introduced. Seed selection can be achieved by tuning θ, which is a threshold and $0 < \theta < 1$. S_i denotes the coupling from other oscillators in the network, which is defined as:

$$S_i = \sum_{k \in N_c(i)} W_{ik} H(x_k - \theta_x) - W_z H(z - \theta_{xz}) \tag{5}$$

Here, θ_x and θ_{xz} are thresholds. W_{ik} is the connection weight from oscillator k to i, and $N_c(i)$, the *coupling neighborhood* of i, is the set of neighboring oscillators of i. W_z is the weight of inhibition from the global inhibitor z, defined by $dz/dt = \phi(\sigma_\infty - \theta_{zx})$. Here, ϕ is a parameter, $\sigma_\infty = 1$ if $x_i \geq \theta_{zx}$ for at least one oscillator i, and $\sigma_\infty = 0$ otherwise, where θ_{zx} 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. 4.

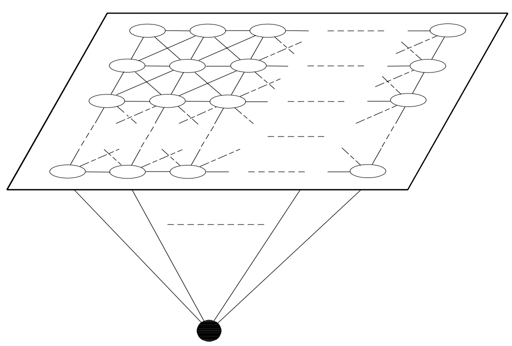


Fig. 4. Architecture of a two-dimensional LEGION network with eight-nearest neighbor coupling. The empty ellipses indicate oscillators, and the filled circle indicates the global inhibitor.

There are well-established results regarding to the LEGION network's performance. Mainly, it has been proved that oscillators belonging to the same region will be synchronized and different oscillator groups will be desynchronized within a finite number of cycles (Terman and Wang, 1995). This guarantees the efficiency of the system. On the other hand, many local and parallel systems, such as cooperative algorithms, including the widely used relaxation labeling technique (Hummel and Zucker, 1983) tend to be slow (Marr, 1982; Gigus and Malik, 1991), and convergence has not generally been established. In addition, the information in the oscillatory network is flexibly encoded by the phase of the oscillators; the system does not need to be provided with the number of labels (critical for the relaxation labeling technique (Hummel and Zucker, 1983)). Further, this system will effectively develop the labels through dynamics. Beyond theoretical properties, the network has been applied successfully for segmenting natural scene and medical images (Wang and Terman, 1997), texture images (Cesmeli and Wang, 1997), and range images (Liu and Wang, 1997).

While there are many parameters involved in the system, most of them are intrinsic. Only two parameters, namely W_{ik} and W_z need to be provided or determined by other methods. For digital orthophoto images, W_{ik} can be defined by:

$$W_{ik} = \frac{C}{|p_i - p_k| + 1} \tag{6}$$

Here C is a constant and p's are the corresponding pixel values. This is only one example of reasonable choices. Theoretically, any monotonically decreasing function of the absolute difference might be used.

As indicated in Fig. 4, the network can be readily implemented by VLSI chips and real-time segmentation could be achieved. On serial computers, however, intensive numerical computation is required to solve the differential equations if the network is applied directly for real images. To reduce the numerical computation on a serial computer, an algorithm is extracted from the dynamics (Wang and Terman, 1997). The algorithm follows the major steps of the original network and exhibits the essential properties. Figure 3(c) shows the segmentation result for the image in Fig. 3(a). The system correctly segmented the four rectangles out.

4. EXPERIMENTAL RESULTS

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 feature-preserving filter family for six iterations with the following parameters: W= 10, H=3, and N=8, and F is a constant function. The smoothed images were then segmented. For hydrographic features, the seed points were identified, based on the assumption that the water bodies in the image are homogenous. Because rivers and streams can be narrow and linear, elongated masks were used to calculate the potential. The system was implemented by applying the masks along different orientations, and using the most homogenous mask to calculate the potential. It is easy to see that this is almost identical to the selective smoothing algorithm discussed earlier. Seed points were actually generated during smoothing.

Figure 5(a) shows an image with a river that has several branches. Figure 5(b) shows the smoothed image. We can see clearly that the regions became more homogenous and the 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 within the river was smoothed out because it is smaller than the size of the filter kernel.

(a) (b) (c)

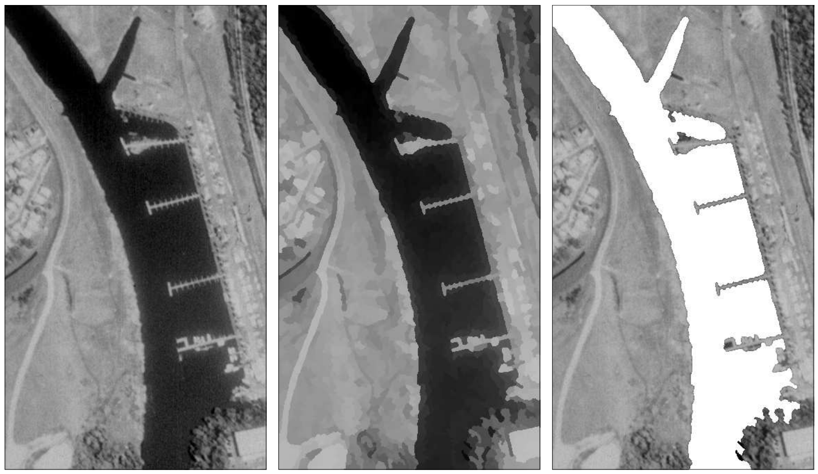
Fig. 5. Segmentation result for a digital orthophoto image. (a) Original image. (b) Smoothed image after six iterations. (c) Segmentation result superimposed on the original image. The river is marked as white.

Figure 6(a) shows another image with a large river with a stream branch and a separate stream. Figure 6(b) shows the smoothing result after six iterations. Figure 6(c) shows the segmentation result superimposed on the original image. We can see that even details along the stream boundaries were preserved and marked accurately. The junctions and corners were accurately located. This clearly demonstrates the performance of the feature-preserving noise-reducing capability of this approach.

(a) (b) (c)

Fig. 6. Segmentation result for a digital orthophoto image. For arrangement, see figure caption for Fig. 5.

Figure 7(a) shows a more complicated image with small objects. Figure 7(b) shows the smoothing result. We can see that the small linear features are preserved, except that the tiny details were smoothed out. The segmentation result shown in Figure 7(c) was very accurate. It segmented the river correctly even along noisy banks in the lower bottom part.



(a) (b) (c)

Fig. 7. Segmentation result for a digital orthophoto image. The arrangement is same as Fig. 5.

5. CONCLUSIONS

In this paper we proposed a feature-preserving smoothing operator family and applied it to identify hydrographic objects in digital orthophoto images. Segmentation was achieved using an oscillatory network. The filter family was derived based on the explicit assumption that the underlying structures are piecewise constant. This nonlinear filter family preserves junctions and corners, as well as edges, with high accuracy as demonstrated in the experimental results.

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 method can be applied to identify other objects, such as road networks. 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 digital orthophoto image processing.

There are several limitations of this approach when applied to arbitrary images. First of all, because the method is derived based the assumption of piecewise constancy, it may not produce satisfactory results if the assumption does not hold. Secondly, the filter tends to smooth out small island features. One way to overcome that is to represent the results in a hierarchy. When major boundaries are identified, we then reduce the filter size to preserve island features.

REFERENCES

Canny, J. (1986). A computational approach to edge detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **8**(6):679-698.

Castleman, K. R. (1996). *Digital Image Processing*. Prentice Hall, New Jersey.

Cesmeli, E. and Wang, D. L. (1997). Texture segmentation using Gaussian Markov random fields and LEGION. *Proceedings of IEEE International Conference on Neural Networks* (*ICNN'97*), Vol. III, pp. 1529-1534.

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

Gigus, Z. and Malik, J. (1991). Detecting Curvilinear Structure in Images. Technical Report UCB/CSD 91/619, Computer Science Division, University of California at Berkeley.

Haralick, R. M. and Shapiro, L. G. (1985). Image segmentation techniques. *Computer Vision, Graphics, and Image Processing*, **29**(1):100-132.

R. H. Hummel and Zucker, S. W. (1983). On the foundations of relaxation labeling processes. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **5**(3):267-286.

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

Liu, X. and Wang, D. L. (1997). Range image segmentation using an oscillatory network. *Proceedings of IEEE International Conference on Neural Networks* (*ICNN'97*), Vol. III, pp. 1656-1660.

Lowe, D. G. (1985). *Perceptual Organization and Visual Recognition*. Kluwer Academic, Boston.

Marr, D. (1982). *Vision*. W. H. Freeman and Company, New York.

Milner, P. M. (1974). A model for visual shape recognition. *Psychological Review*, **81**(6):521-535.

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

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

Nagao, M. and Matsuyama, T. (1979). Edge preserving smoothing. *Computer Graphics and Image Processing*, **9**(4):394-407.

Pal, N. R. and Pal, S. K. (1993). Review on image segmentation techniques. *Pattern Recognition*, **26**(9):1277-1294.

Perona, P. and Malik, J. (1987). Scale space and edge detection using anisotropic diffusion. *Proceedings of IEEE Computer Society Workshop on Computer Vision*, pp. 16-27.

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

Sarkar, S. and Boyer, K. L. (1993). Perceptual organization in computer vision: A review and a proposal for a classificatory structure. *IEEE Transactions on Systems, Man, and Cybernetics*, **23**(2):382-399.

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

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

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

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

Wang, D. L. and Terman, D. (1995). Locally excitatory globally inhibitory oscillator networks. *IEEE Transactions on Neural Networks,* **6**(1):283-286.

Wang, D. L. and Terman, D. (1997). Image segmentation based on oscillatory correlation. *Neural Computation*, **9**(4):805-836.