Range Image Segmentation Using an Oscillatory Network

Xiuwen Liu† and DeLiang Wang†‡
†*Department of Computer and Information Science*
‡*Center for Cognitive Science*
The Ohio State University
Columbus, Ohio 43210, USA
{liux, dwang}@cis.ohio-state.edu

Abstract

We use a locally excitatory globally inhibitory oscillator network (LEGION) as a framework for range image segmentation. Each oscillator in the LEGION network has excitatory lateral connections to the oscillators in its neighborhood as well as a connection with a global inhibitor. The lateral connection between two oscillators is established based on the similarity between their feature vectors which consist of the surface normal and curvature at the corresponding pixel locations. The emergent behavior of the LEGION network gives rise to the segmentation result. Unlike other methods, our scheme needs no assumption about the underlying structures in image data and no prior knowledge regarding the number of regions. Experimental results for real range images are presented.

I. Introduction

Image segmentation has long been considered in computer vision as an important and difficult task. Due its practical importance, many techniques have been proposed for range image segmentation, mainly region-based [4] [2] [1] [7]. The region-based algorithms first estimate surface normal and curvature at each pixel, and then try to cluster range data into a set of homogeneous regions. They can be implemented by region growing or feature vector clustering techniques. In region-growing algorithms [1] [7], pixels are first coarsely classified based on the sign of mean and Gaussian surface curvature, and the final results are obtained by iterative region growing using surface fitting. One of the shortcomings of this approach is that a good surface model for the range image data must be provided. In feature vector clustering techniques [4] [2], range image segmentation is posed as a vector quantization problem by associating each pixel with an appropriate feature vector. One of the limitations is that the number of regions needs to be specified *a priori*, which, in general, is not available.

Recently, Koh et al. [5] proposed a hierachical self-organizing network for range image segmentation. At each level, a self-organizing feature map (SOFM) [6] is used to segment range images into a given number of regions. An *abstract tree* [5] is constructed to represent the output of the hierachical SOFM network. The final segmentation is obtained by searching through this abstract tree, which is sequential and similar to region-based methods. Thus it suffers the problems associated with region-based algorithms.

In this paper, we use a novel neural network approach for segmenting range images, which overcomes some of the above limitations. Locally excitatory globally inhibitory oscillator network (LEGION) [8] provides a biologically plausible framework for image segmentation in general. Each oscillator in the LEGION network connects excitatorily with the oscillators in its neighborhood as well as inhibitorily with a global inhibitor. For range image segmentation, the feature detector associated with each oscillator estimates the surface normal and curvature at its corresponding pixel location. The lateral connection between two oscillators is set based on a similarity measure of their feature vectors. The segmentation process is the emergent behavior of the oscillator network. Experimental results using real range images are provided.

II. Overview of LEGION dynamics

As the building block of LEGION, a single oscillator *i*, is defined as a feedback loop between an excitatory unit x_i and an inhibitory unit y_i [9]:

$$\frac{dx_i}{dt} = 3x_i - x_i^3 + 2 - y_i + I_i H(p_i - \theta) + S_i + \rho \quad (1a)$$

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

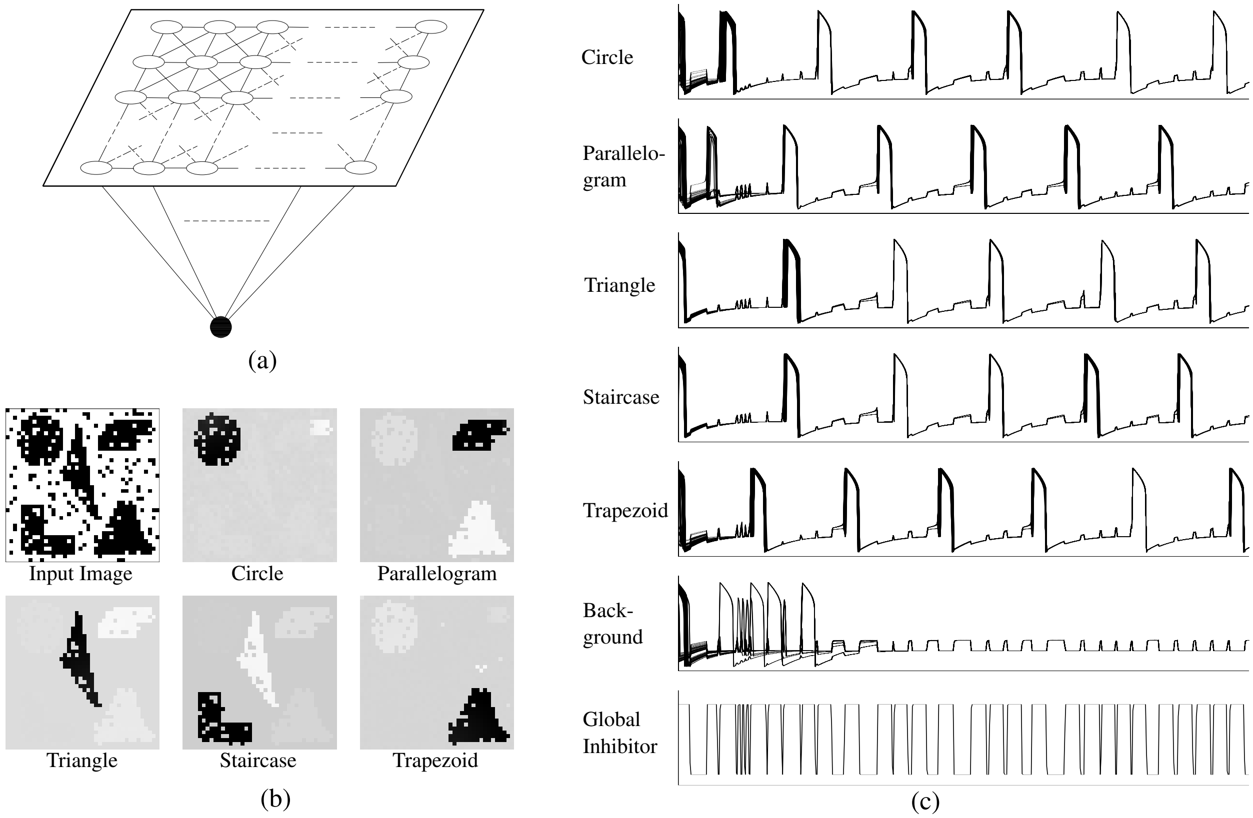


Fig. 1. (a). Architecture of a two-dimensional LEGION network. (b) The noisy input image to the network and the snapshots when one object is in active phase. The input image is binary with 40 x 40 pixels. (c). The temporal evolution of the LEGION network. The upper five plots show the activities of the oscillator blocks representing individual objects. The sixth plot shows the activities of the noisy stimulated oscillators, called the background. They became silent after two cycles. The bottom one shows the activity of the global inhibitor. The parameter values for the simulation: ε = 0.02, β = 0.1, γ = 12.0, θ_x = -0.5, ϕ = 3.0, $\theta_{zx} = \theta_{xz}$ = 0.1, θ = 0.8, W_z = 2.0, W_{ik} = 0.75, and I_i = 0.2 for stimulated oscillators, and I_i = -0.02 otherwise.

Here H stands for the Heaviside step function. I_i represents external stimulation for oscillator i, and ρ denotes a Gaussian noise term. A *potential* for the oscillator i, p_i, is introduced to allow the network to distinguish between major regions and noisy fragments. See [9] for a detailed discussion.

The parameter ε is a small positive number, and in this case (1), without any coupling or noise but with constant stimulation, corresponds to a standard relaxation oscillator. If $I > 0$ and $H = 1$, (1) produces a stable periodic orbit for all sufficiently small values of ε, alternating between *silent* and *active* phases of near steady-state behavior. An oscillator is in the silent phase if its excitatory unit has low activity, and in the active phase if its excitatory unit has high activity. The transition between the two phases occurs rapidly.

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

$$S_i = \sum_{k \in N_c(i)} W_{ik}\, H(x_k - \theta_x) - W_z\, H(z - \theta_{xz}) \qquad (2)$$

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 a set of neighboring oscillators of i. W_z is the weight of inhibition from the global inhibitor z, which is 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 we use, shown in Fig. 1(a) is two dimensional, where each oscillator is connected with its eight immediate neighbors, as well as to a global inhibitor. We illustrate the capability of the network for image segmentation by a binary synthetic image. Fig. 1(b) shows the noisy input image with 40 x 40 pixels, and snapshots of the network activity when one object is in the active phase. The oscillators that belong to same objects are synchronized. Fig. 1(c) shows the temporal evolution of all the stimulated oscillators. All the five oscillator groups representing objects are popped up alternately after three cycles. Formally, it has been proved [8] that the LEGION network can rapidly achieve synchronization among the oscillators corresponding to the same object and desynchronization between groups corresponding to different objects.

III. Similarity measure for range images

For real images with a large number of pixels, it involves prohibitively intensive numerical computation to solve the differential equations if the LEGION network is applied directly. To reduce the numerical computation on a serial computer, an algorithm extracted from these equations [9] is applied to range image segmentation. Given LEGION dynamics, the main task is to establish lateral connections based on some similarity measure. Intuitively, pixels belonging to the same surface should be similar and thus should have strong connections with each other in the LEGION network. Thus the similarity measure should depend on surface normal and curvature in addition to raw depth data. Similar to [5], we chose (z, n_x, n_y, n_z, C) as the feature vector for each oscillator. Here z is the depth value, (n_x, n_y, n_z) is the surface normal, and C is the mean curvature, at the corresponding pixel location.

Based on the work in [1], (n_x, n_y, n_z) is calculated from the first-order partial derivatives:

$$\left(n_x, n_y, n_z\right) = \frac{(\frac{\partial z}{\partial x}, 0, 1) \times (0, \frac{\partial z}{\partial y}, 1)}{\left\|(\frac{\partial z}{\partial x}, 0, 1) \times (0, \frac{\partial z}{\partial y}, 1)\right\|} \quad (3)$$

and the partial derivatives are estimated using the following formula:

$$\frac{\partial z}{\partial x} \approx D_x = \boldsymbol{d}_0 \cdot \boldsymbol{d}_1^T \quad (4a)$$

$$\frac{\partial z}{\partial y} \approx D_y = \boldsymbol{d}_1 \cdot \boldsymbol{d}_0^T \quad (4b)$$

Here T indicates transpose, $\boldsymbol{d}_0$ is the equally weighted average operator, and $\boldsymbol{d}_1$ is the least-square derivative estimator [1]. For a 5 x 5 window, $\boldsymbol{d}_0$ is 1/5 $[1, 1, 1, 1, 1]^T$, and $\boldsymbol{d}_1$ is 1/10 $[-2, -1, 0, 1, 2]^T$. This normal estimation method works well if the estimation window is within one surface. When the window crosses different surfaces, especially those with different depth values, the estimation results tend to be inaccurate. In order to improve the results near surface boundaries, $\boldsymbol{d}_0$ and $\boldsymbol{d}_1$ are applied only to pixels that have similar pixel values with respect to the central pixel. These operators are called *context preserving* operators.

Following [4], the surface curvature at point $\boldsymbol{p}$ in the direction of point $\boldsymbol{q}$ is estimated by:

$$\kappa_{p,q} = \begin{cases} \dfrac{\|\boldsymbol{n}_p - \boldsymbol{n}_q\|}{\|\boldsymbol{p} - \boldsymbol{q}\|}, & \text{if } \|\boldsymbol{p}-\boldsymbol{q}\| \le \left\|(\boldsymbol{n}_p + \boldsymbol{p}) - (\boldsymbol{n}_q + \boldsymbol{q})\right\| \\ -\dfrac{\|\boldsymbol{n}_p - \boldsymbol{n}_q\|}{\|\boldsymbol{p} - \boldsymbol{q}\|}, & \text{otherwise} \end{cases} \quad (5)$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ refer to the 3-D coordinate vectors of the corresponding pixels, and $\boldsymbol{n}_p$ and $\boldsymbol{n}_q$ are the unit normal vectors at points $\boldsymbol{p}$ and $\boldsymbol{q}$ respectively, which are available from (3). Here the 3-D coordinate vector of a pixel is composed of its 2-D location in the image and its depth value. The condition in (5) is to assign a positive curvature value for pixels on a convex surface and a negative one for pixels on a concave surface. The mean curvature C_i of oscillator i is defined as the average of all possible values with respect to a neighborhood $N_k(i)$ of oscillator i.

In summary, the similarity measure between two oscillators is given by: $W_{ik} = 255/\psi(i, k)$, and the lateral connections are set up accordingly. Here 255 is the possible maximum value, and $\psi(i, k)$ is a disparity measure between two oscillators, and is given by:

$$\psi(i, k) = \begin{cases} K_z |z_i - z_k| + K_n \|\boldsymbol{n}_i - \boldsymbol{n}_k\| + B, & \text{if } |C_i| \le T_c \\ K_c |C_i - C_k| + B, & \text{otherwise} \end{cases} \quad (6)$$

Here K_z, K_n, and K_c are weights for different disparity measures. B is a constant biased term, and T_c is a threshold for planar surface testing.

IV. Experimental results

We applied our network to synthetic and real images. We selected some of them and showed here to demonstrate the performance of the network. For all the results, we use a single system configuration. That is, the surface normal is estimated from a 5 x 5 window, the mean curvature is estimated from a 3 x 3 window, and each oscillator has lateral connections with its eight immediate neighbors as depicted in Fig. 1(a). The parameters for the disparity measure in (6) are set as the following: $K_z = 1$, $K_n = 50$, $K_c = 200$, $T_c = 0.005$, and $B = 0.1$.

Fig. 2 shows a complete example from the LEGION network for a real range image. From its 3-D model, the input image consists of a cylinder and a rectangular parallelepiped. All of the four visible surfaces in the image, including the background, are popped up and thus correctly segmented. The segmentation result shown here is obtained by putting the popped up regions (excluding the background) together into a *gray map* [9], where each region is shown using a single gray level.

Fig. 3 shows the segmentation results for several other

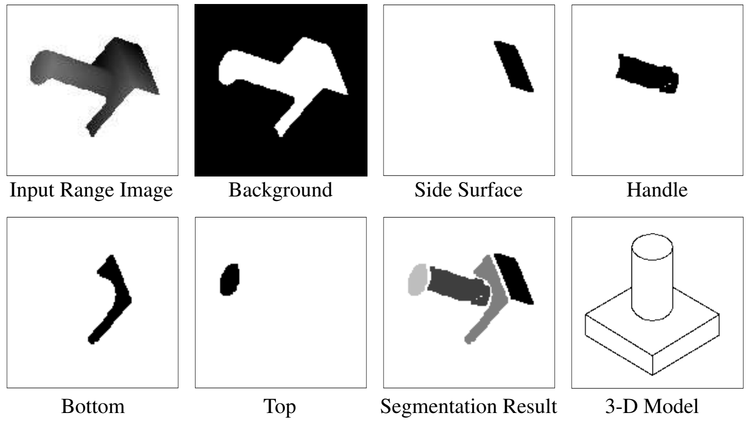


Fig. 2. Segmentation result of the LEGION network for a range image.

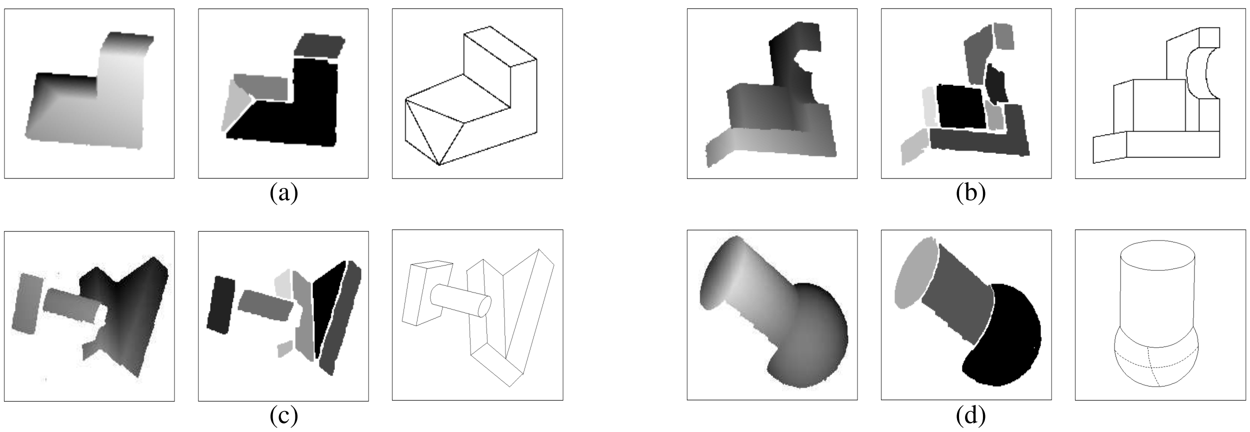


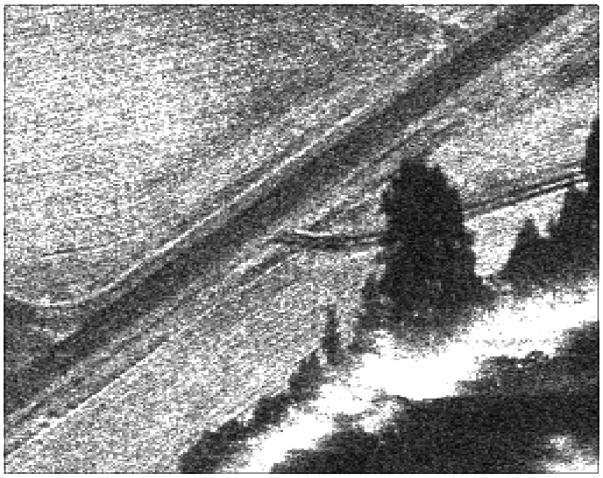
Fig. 3. Segmentation results of the network for real range images. In each subfigure, left frame shows the input range image, the middle is the segmentation result represented by a gray map, and the right is the 3-D construction model for evaluation purposes.

real range images. In Fig. 3(a), an object with only planar surfaces is segmented into four surface regions. Boundaries between surfaces are precisely marked. This is because that surface normal can be estimated relatively well for large planar surfaces. Fig. 3(b) shows an object with several planar surfaces and a cylindrical one. As in Fig. 3(a), all the planar surfaces are segmented with boundaries marked precisely. The cylindrical surface is segmented correctly even though it is not complete because of the shadowing effect. Fig. 3(c) and 3(d) consist of two combined objects. All surfaces are segmented correctly with boundaries marked clearly. These results demonstrate that our network can handle different types of surfaces. We have also applied our network to natural scene images. Fig. 4(a) shows a SAR image, and the gray values in it are interpreted as range data because of the lack of range images of natural scenes. As shown in Fig. 4(b), the main features, such as the main road, the cloud, and the bushes are segmented out correctly.

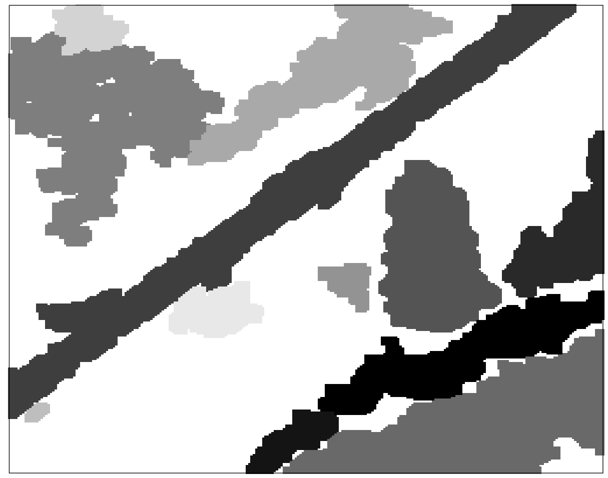
V. Discussion

A LEGION network has been constructed and applied for range image segmentation. The network successfully segments real range images, suggesting a generic range image segmentation method. For the range images with well defined objects, our results are comparable with results from other methods [4] [2] [1] [5]. Based on the comparison framework proposed in [3], our network segments all the images in Fig. 2 and Fig. 3 correctly, without any *over-segmentation*, *under-segmentation*, *missed* classification or *noise* classification [3].

Methodologically, our approach has several major advantages over other methods. Firstly, the segmentation is achieved by local computation mainly. Unlike region-growing method [1] [7], we do not need to define a surface model to fit the image data. Secondly, the segmentation result is the emergent behavior of the LEGION network. We do not need to do any searching. In contrast, for the



(a). Original SAR image



(b). Segmentation result

Fig. 4. The segmentation result for a natural scene image.

neural network approach proposed in [5], the final segmentation result was obtained by searching through the abstract tree, which is sequential and not neurally implementable. Furthermore, our approach is a dynamical system with parallel and neural computation. These unique properties make it feasible for analog VLSI implementation, the efficiency of which could lead to real-time range image segmentation.

Acknowledgments

This work is 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 like to thank the Pattern Recognition and Image Processing Laboratory at Michigan State University for providing a range image database, and the MIT Lincoln Laboratory for providing the SAR image used in this study.

References

[1] P. J. Besl and R. C. Jain, "Segmentation through variable-order surface fitting", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 10(2), pp. 167-192, 1988.

[2] R. Hoffman and A. K. Jain, "Segmentation and classification of range images", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 9(5), pp. 608-620, 1987.

[3] A. Hoover, G. Jean-Baptiste, and et al, "An experimental comparison of range image segmentation algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 18(7), pp. 673-689, 1996.

[4] D. J. Ittner and A. K. Jain, "3-D surface discrimination from local curvature measures", in *Proceedings of the Computer Vision and Pattern Recognition Conference,* pp. 119-123, 1985.

[5] J. Koh, M. Suk, and S. M. Bhandarkar, "A multilayer self-organizing feature map for range image segmentation", *Neural Networks,* vol. 8(1), pp. 67-86, 1995.

[6] T. Kohonen, *Self-Organizing Maps,* Berlin Heidelberg: Springer, 1995.

[7] A. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation as the search for the best description of the image in terms of primitives", in *Proceedings of the International Conference on Computer Vision,* pp. 121-125, 1990.

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

[9] D. L. Wang and D. Terman, "Image segmentation based on oscillatory correlation", *Neural Computation* (in press); see also Technical Report 19, Center for Cognitive Science, The Ohio State University, 1996.

Range Image Segmentation Using an Oscillatory Network

Xiuwen Liu[†] and DeLiang Wang[†‡]
†*Department of Computer and Information Science*
‡*Center for Cognitive Science*
The Ohio State University
Columbus, Ohio 43210, USA
{liux, dwang}@cis.ohio-state.edu

Abstract

We use a locally excitatory globally inhibitory oscillator network (LEGION) as a framework for range image segmentation. Each oscillator in the LEGION network has excitatory lateral connections to the oscillators in its neighborhood as well as a connection with a global inhibitor. The lateral connection between two oscillators is established based on the similarity between their feature vectors which consist of the surface normal and curvature at the corresponding pixel locations. The emergent behavior of the LEGION network gives rise to the segmentation result. Unlike other methods, our scheme needs no assumption about the underlying structures in image data and no prior knowledge regarding the number of regions. Experimental results for real range images are presented.

I. Introduction

Image segmentation has long been considered in computer vision as an important and difficult task. Due its practical importance, many techniques have been proposed for range image segmentation, mainly region-based [4] [2] [1] [7]. The region-based algorithms first estimate surface normal and curvature at each pixel, and then try to cluster range data into a set of homogeneous regions. They can be implemented by region growing or feature vector clustering techniques. In region-growing algorithms [1] [7], pixels are first coarsely classified based on the sign of mean and Gaussian surface curvature, and the final results are obtained by iterative region growing using surface fitting. One of the shortcomings of this approach is that a good surface model for the range image data must be provided. In feature vector clustering techniques [4] [2], range image segmentation is posed as a vector quantization problem by associating each pixel with an appropriate feature vector. One of the limitations is that the number of regions needs to be specified *a priori*, which, in general, is not available.

Recently, Koh et al. [5] proposed a hierachical self-organizing network for range image segmentation. At each level, a self-organizing feature map (SOFM) [6] is used to segment range images into a given number of regions. An *abstract tree* [5] is constructed to represent the output of the hierachical SOFM network. The final segmentation is obtained by searching through this abstract tree, which is sequential and similar to region-based methods. Thus it suffers the problems associated with region-based algorithms.

In this paper, we use a novel neural network approach for segmenting range images, which overcomes some of the above limitations. Locally excitatory globally inhibitory oscillator network (LEGION) [8] provides a biologically plausible framework for image segmentation in general. Each oscillator in the LEGION network connects excitatorily with the oscillators in its neighborhood as well as inhibitorily with a global inhibitor. For range image segmentation, the feature detector associated with each oscillator estimates the surface normal and curvature at its corresponding pixel location. The lateral connection between two oscillators is set based on a similarity measure of their feature vectors. The segmentation process is the emergent behavior of the oscillator network. Experimental results using real range images are provided.

II. Overview of LEGION dynamics

As the building block of LEGION, a single oscillator *i,* is defined as a feedback loop between an excitatory unit x_i and an inhibitory unit y_i [9]:

$$\frac{dx_i}{dt} = 3x_i - x_i^3 + 2 - y_i + I_i H(p_i - \theta) + S_i + \rho \quad \text{(1a)}$$

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

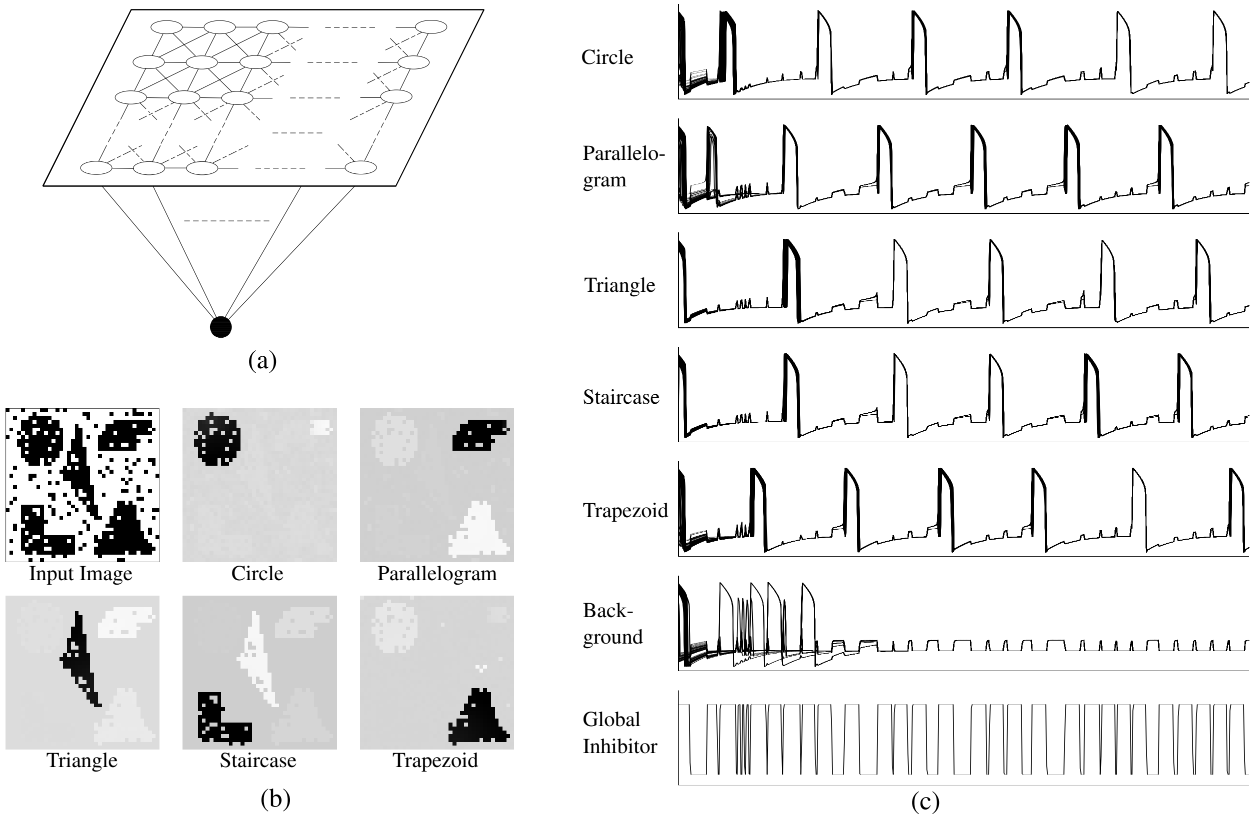


Fig. 1. (a). Architecture of a two-dimensional LEGION network. (b) The noisy input image to the network and the snapshots when one object is in active phase. The input image is binary with 40 x 40 pixels. (c). The temporal evolution of the LEGION network. The upper five plots show the activities of the oscillator blocks representing individual objects. The sixth plot shows the activities of the noisy stimulated oscillators, called the background. They became silent after two cycles. The bottom one shows the activity of the global inhibitor. The parameter values for the simulation: ε = 0.02, β = 0.1, γ = 12.0, θ_x = -0.5, ϕ = 3.0, $\theta_{zx} = \theta_{xz}$ = 0.1, θ = 0.8, W_z = 2.0, W_{ik} = 0.75, and I_i = 0.2 for stimulated oscillators, and I_i = -0.02 otherwise.

Here H stands for the Heaviside step function. I_i represents external stimulation for oscillator i, and ρ denotes a Gaussian noise term. A *potential* for the oscillator i, p_i, is introduced to allow the network to distinguish between major regions and noisy fragments. See [9] for a detailed discussion.

The parameter ε is a small positive number, and in this case (1), without any coupling or noise but with constant stimulation, corresponds to a standard relaxation oscillator. If $I > 0$ and $H = 1$, (1) produces a stable periodic orbit for all sufficiently small values of ε, alternating between *silent* and *active* phases of near steady-state behavior. An oscillator is in the silent phase if its excitatory unit has low activity, and in the active phase if its excitatory unit has high activity. The transition between the two phases occurs rapidly.

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

$$S_i = \sum_{k \in N_c(i)} W_{ik} H(x_k - \theta_x) - W_z H(z - \theta_{xz}) \qquad (2)$$

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 a set of neighboring oscillators of i. W_z is the weight of inhibition from the global inhibitor z, which is 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 we use, shown in Fig. 1(a) is two dimensional, where each oscillator is connected with its eight immediate neighbors, as well as to a global inhibitor. We illustrate the capability of the network for image segmentation by a binary synthetic image. Fig. 1(b) shows the noisy input image with 40 x 40 pixels, and snapshots of the network activity when one object is in the active phase. The oscillators that belong to same objects are synchronized. Fig. 1(c) shows the temporal evolution of all the stimulated oscillators. All the five oscillator groups representing objects are popped up alternately after three cycles. Formally, it has been proved [8] that the LEGION network can rapidly achieve synchronization among the oscillators corresponding to the same object and desynchronization between groups corresponding to different objects.

III. Similarity measure for range images

For real images with a large number of pixels, it involves prohibitively intensive numerical computation to solve the differential equations if the LEGION network is applied directly. To reduce the numerical computation on a serial computer, an algorithm extracted from these equations [9] is applied to range image segmentation. Given LEGION dynamics, the main task is to establish lateral connections based on some similarity measure. Intuitively, pixels belonging to the same surface should be similar and thus should have strong connections with each other in the LEGION network. Thus the similarity measure should depend on surface normal and curvature in addition to raw depth data. Similar to [5], we chose (z, n_x, n_y, n_z, C) as the feature vector for each oscillator. Here z is the depth value, (n_x, n_y, n_z) is the surface normal, and C is the mean curvature, at the corresponding pixel location.

Based on the work in [1], (n_x, n_y, n_z) is calculated from the first-order partial derivatives:

$$\left(n_x, n_y, n_z\right) = \frac{\left(\frac{\partial z}{\partial x}, 0, 1\right) \times \left(0, \frac{\partial z}{\partial y}, 1\right)}{\left\|\left(\frac{\partial z}{\partial x}, 0, 1\right) \times \left(0, \frac{\partial z}{\partial y}, 1\right)\right\|} \tag{3}$$

and the partial derivatives are estimated using the following formula:

$$\frac{\partial z}{\partial x} \approx D_x = \boldsymbol{d}_0 \cdot \boldsymbol{d}_1^T \tag{4a}$$

$$\frac{\partial z}{\partial y} \approx D_y = \boldsymbol{d}_1 \cdot \boldsymbol{d}_0^T \tag{4b}$$

Here T indicates transpose, $\boldsymbol{d}_0$ is the equally weighted average operator, and $\boldsymbol{d}_1$ is the least-square derivative estimator [1]. For a 5 x 5 window, $\boldsymbol{d}_0$ is 1/5 $[1, 1, 1, 1, 1]^T$, and $\boldsymbol{d}_1$ is 1/10 $[-2, -1, 0, 1, 2]^T$. This normal estimation method works well if the estimation window is within one surface. When the window crosses different surfaces, especially those with different depth values, the estimation results tend to be inaccurate. In order to improve the results near surface boundaries, $\boldsymbol{d}_0$ and $\boldsymbol{d}_1$ are applied only to pixels that have similar pixel values with respect to the central pixel. These operators are called *context preserving* operators.

Following [4], the surface curvature at point $\boldsymbol{p}$ in the direction of point $\boldsymbol{q}$ is estimated by:

$$\kappa_{p,q} = \begin{cases} \dfrac{\|\boldsymbol{n}_p - \boldsymbol{n}_q\|}{\|\boldsymbol{p} - \boldsymbol{q}\|}, & \text{if } \|\boldsymbol{p}-\boldsymbol{q}\| \le \|(\boldsymbol{n}_p+\boldsymbol{p})-(\boldsymbol{n}_q+\boldsymbol{q})\| \\ -\dfrac{\|\boldsymbol{n}_p - \boldsymbol{n}_q\|}{\|\boldsymbol{p} - \boldsymbol{q}\|}, & \text{otherwise} \end{cases} \tag{5}$$

where $\boldsymbol{p}$ and $\boldsymbol{q}$ refer to the 3-D coordinate vectors of the corresponding pixels, and $\boldsymbol{n}_p$ and $\boldsymbol{n}_q$ are the unit normal vectors at points $\boldsymbol{p}$ and $\boldsymbol{q}$ respectively, which are available from (3). Here the 3-D coordinate vector of a pixel is composed of its 2-D location in the image and its depth value. The condition in (5) is to assign a positive curvature value for pixels on a convex surface and a negative one for pixels on a concave surface. The mean curvature C_i of oscillator i is defined as the average of all possible values with respect to a neighborhood $N_k(i)$ of oscillator i.

In summary, the similarity measure between two oscillators is given by: $W_{ik} = 255/\psi(i, k)$, and the lateral connections are set up accordingly. Here 255 is the possible maximum value, and $\psi(i, k)$ is a disparity measure between two oscillators, and is given by:

$$\psi(i, k) = \begin{cases} K_z |z_i - z_k| + K_n \|\boldsymbol{n}_i - \boldsymbol{n}_k\| + B, & \text{if } |C_i| \le T_c \\ K_c |C_i - C_k| + B, & \text{otherwise} \end{cases} \tag{6}$$

Here K_z, K_n, and K_c are weights for different disparity measures. B is a constant biased term, and T_c is a threshold for planar surface testing.

IV. Experimental results

We applied our network to synthetic and real images. We selected some of them and showed here to demonstrate the performance of the network. For all the results, we use a single system configuration. That is, the surface normal is estimated from a 5 x 5 window, the mean curvature is estimated from a 3 x 3 window, and each oscillator has lateral connections with its eight immediate neighbors as depicted in Fig. 1(a). The parameters for the disparity measure in (6) are set as the following: $K_z = 1$, $K_n = 50$, $K_c = 200$, $T_c = 0.005$, and $B = 0.1$.

Fig. 2 shows a complete example from the LEGION network for a real range image. From its 3-D model, the input image consists of a cylinder and a rectangular parallelepiped. All of the four visible surfaces in the image, including the background, are popped up and thus correctly segmented. The segmentation result shown here is obtained by putting the popped up regions (excluding the background) together into a *gray map* [9], where each region is shown using a single gray level.

Fig. 3 shows the segmentation results for several other

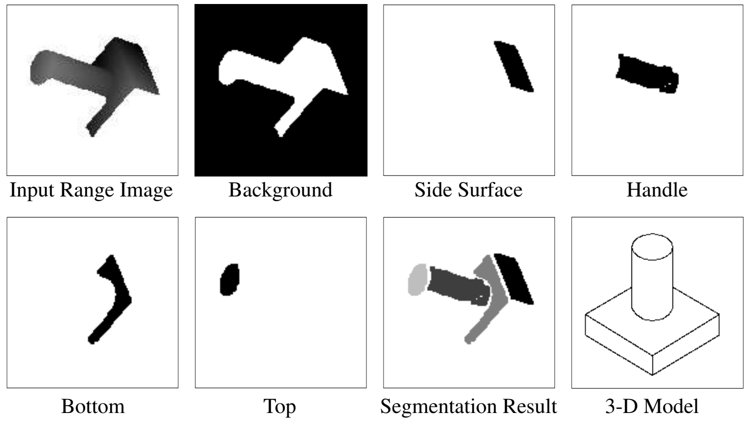


Fig. 2. Segmentation result of the LEGION network for a range image.

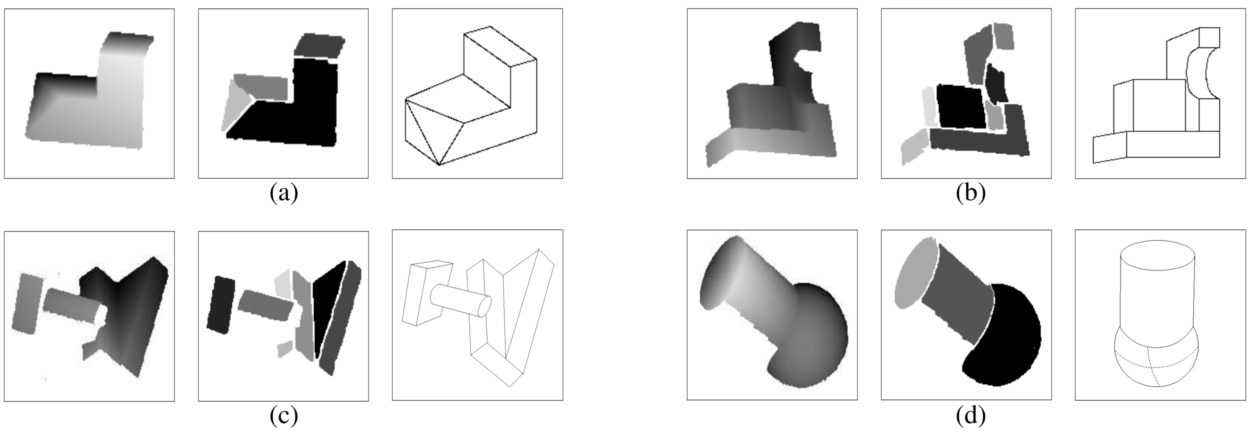


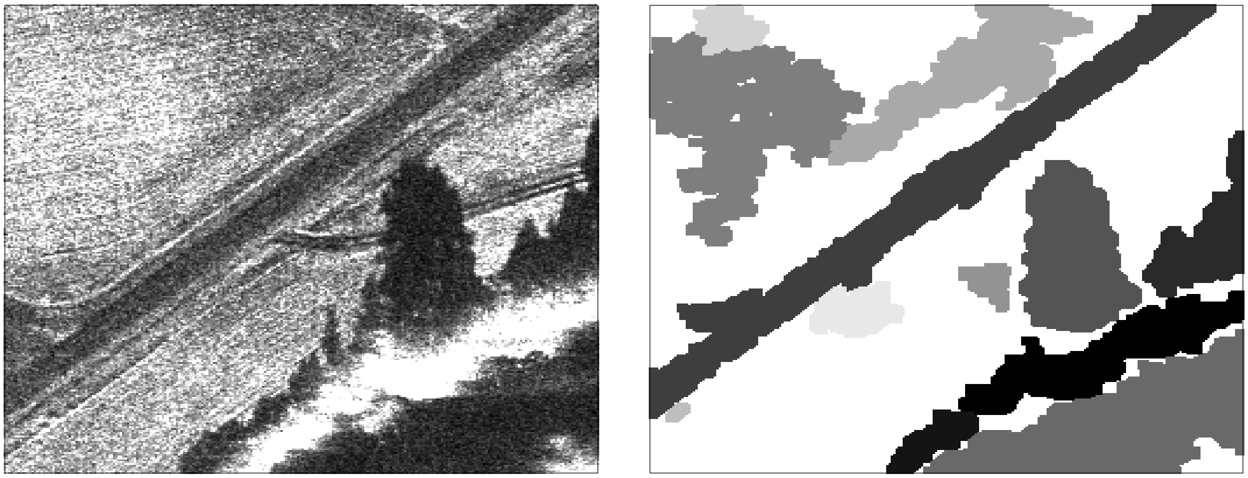
Fig. 3. Segmentation results of the network for real range images. In each subfigure, left frame shows the input range image, the middle is the segmentation result represented by a gray map, and the right is the 3-D construction model for evaluation purposes.

real range images. In Fig. 3(a), an object with only planar surfaces is segmented into four surface regions. Boundaries between surfaces are precisely marked. This is because that surface normal can be estimated relatively well for large planar surfaces. Fig. 3(b) shows an object with several planar surfaces and a cylindrical one. As in Fig. 3(a), all the planar surfaces are segmented with boundaries marked precisely. The cylindrical surface is segmented correctly even though it is not complete because of the shadowing effect. Fig. 3(c) and 3(d) consist of two combined objects. All surfaces are segmented correctly with boundaries marked clearly. These results demonstrate that our network can handle different types of surfaces. We have also applied our network to natural scene images. Fig. 4(a) shows a SAR image, and the gray values in it are interpreted as range data because of the lack of range images of natural scenes. As shown in Fig. 4(b), the main features, such as the main road, the cloud, and the bushes are segmented out correctly.

V. Discussion

A LEGION network has been constructed and applied for range image segmentation. The network successfully segments real range images, suggesting a generic range image segmentation method. For the range images with well defined objects, our results are comparable with results from other methods [4] [2] [1] [5]. Based on the comparison framework proposed in [3], our network segments all the images in Fig. 2 and Fig. 3 correctly, without any *over-segmentation*, *under-segmentation*, *missed* classification or *noise* classification [3].

Methodologically, our approach has several major advantages over other methods. Firstly, the segmentation is achieved by local computation mainly. Unlike region-growing method [1] [7], we do not need to define a surface model to fit the image data. Secondly, the segmentation result is the emergent behavior of the LEGION network. We do not need to do any searching. In contrast, for the



(a). Original SAR image (b). Segmentation result

Fig. 4. The segmentation result for a natural scene image.

neural network approach proposed in [5], the final segmentation result was obtained by searching through the abstract tree, which is sequential and not neurally implementable. Furthermore, our approach is a dynamical system with parallel and neural computation. These unique properties make it feasible for analog VLSI implementation, the efficiency of which could lead to real-time range image segmentation.

Acknowledgments

This work is 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 like to thank the Pattern Recognition and Image Processing Laboratory at Michigan State University for providing a range image database, and the MIT Lincoln Laboratory for providing the SAR image used in this study.

References

[1] P. J. Besl and R. C. Jain, "Segmentation through variable-order surface fitting", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 10(2), pp. 167-192, 1988.

[2] R. Hoffman and A. K. Jain, "Segmentation and classification of range images", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 9(5), pp. 608-620, 1987.

[3] A. Hoover, G. Jean-Baptiste, and et al, "An experimental comparison of range image segmentation algorithms", *IEEE Transactions on Pattern Analysis and Machine Intelligence,* vol. 18(7), pp. 673-689, 1996.

[4] D. J. Ittner and A. K. Jain, "3-D surface discrimination from local curvature measures", in *Proceedings of the Computer Vision and Pattern Recognition Conference,* pp. 119-123, 1985.

[5] J. Koh, M. Suk, and S. M. Bhandarkar, "A multilayer self-organizing feature map for range image segmentation", *Neural Networks,* vol. 8(1), pp. 67-86, 1995.

[6] T. Kohonen, *Self-Organizing Maps,* Berlin Heidelberg: Springer, 1995.

[7] A. Leonardis, A. Gupta, and R. Bajcsy, "Segmentation as the search for the best description of the image in terms of primitives", in *Proceedings of the International Conference on Computer Vision,* pp. 121-125, 1990.

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

[9] D. L. Wang and D. Terman, "Image segmentation based on oscillatory correlation", *Neural Computation* (in press); see also Technical Report 19, Center for Cognitive Science, The Ohio State University, 1996.