Motion Segmentation Based on Motion/Brightness Integration and Oscillatory Correlation

Erdogan Çesmeli, *Member, IEEE,* and DeLiang Wang, *Associate Member, IEEE*

Abstract—A segmentation method based on the integration of motion and brightness is proposed for image sequences. The method is composed of two parallel pathways that process motion and brightness, respectively. Inspired by the visual system, the motion pathway has two stages. The first stage estimates local motion at locations with reliable information. The second stage performs segmentation based on local motion estimates. In the brightness pathway, the input scene is segmented into regions based on brightness distribution. Subsequently, segmentation results from the two pathways are integrated to refine motion estimates. The final segmentation is performed in the motion network based on refined estimates. For segmentation, locally excitatory globally inhibitory oscillator network (LEGION) architecture is employed whereby the oscillators corresponding to a region of similar motion/brightness oscillate in synchrony and different regions attain different phases. Results on synthetic and real image sequences are provided, and comparisons with other methods are made.

Index Terms—LEGION, motion and brightness integration, motion estimation, neural networks, oscillatory correlation.

I. INTRODUCTION

COMPUTATIONAL investigation of motion perception presents three major challenges. The first one is the well-known aperture problem [1], which states the insufficiency of local motion estimation to uniquely determine the true underlying motion. This suggests that the integration of a collection of local estimates is essential. The second one, known as the blank wall problem [26], arises from homogeneous regions and is closely related to the boundary assignment problem [17]. When a region has relatively uniform brightness, local motion estimates at its inner locations are not unique. This problem is illustrated in Fig. 1, where a homogeneous circular region is moving on a homogeneous background. Consistent with the stimulus, subjects perceive the circular region moving. Although the circular boundary and its motion can be extracted, whether the moving boundary belongs to the circular (inner) region or to the (outer) background is computationally ambiguous. A solution to this problem seems to require a combination of the estimates at the inner locations of a surface and those along its boundary. Hence, both of the
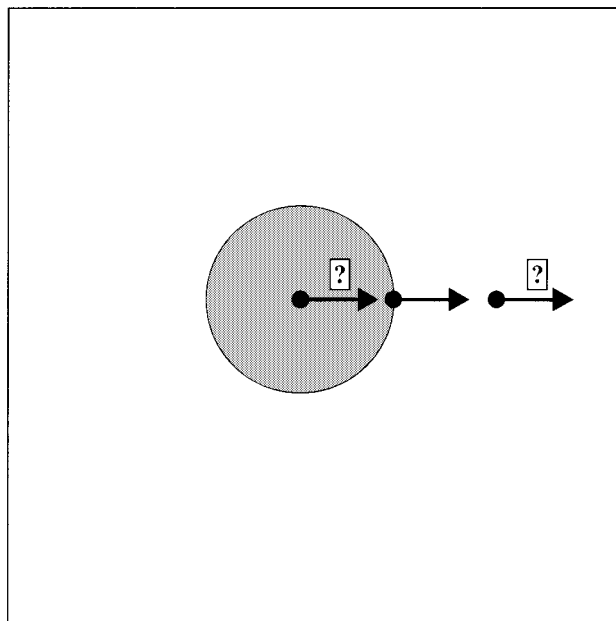
Fig. 1.   Example illustrating the blank wall problem.

challenges call for a selective integration of local estimates to uniquely identify the underlying motion. The third challenge is how to represent multiple motions over the same area, which is necessary in order to model motion transparency [13].

Neural-network methods consider only a subset of these challenges. Most of them address the aperture problem by providing a scheme to selectively integrate local estimates [10], [11], [18], [20], [21], [24], [25], [28], [30], [35]. Few consider the blank wall problem, e.g., [11]. Similarly, motion transparency has not received much attention in neural networks [28], [21]. Most of the neural-network methods are designed to model neuronal and perceptual behavior [11], [18], [25], [28], usually using psychophysical stimuli, e.g., plaids and random dot images, and have not been tested using real motion scenes.

Unlike neural-network methods, image processing algorithms can deal with real scenes [3], [5], [7], [8], [16], [26], [29], [33], [34]. However, they also fail to consider all of the three challenges simultaneously. One class of algorithms employs motion energy filters for motion detection [2], [21], [26], [33]. These algorithms can represent motion transparency. However, they are fundamentally limited to translational motion and provide limited spatial localization. Starting from the pioneering study by Horn and Schunck [15], another class utilizes the brightness constraint equation wherein local motion components are related to temporal and spatial derivatives [5], [7], [8], [15], [16]. These algorithms can represent translational,

affine, and planar motions. When translational motion is assumed, cross-correlation methods, such as the one proposed by Anandan [3], can be utilized for motion estimation. Black and colleagues have employed robust statistics [5], [7], [8] to overcome the difficulty of using fixed neighborhoods in the assessment of derivatives, especially in the presence of noise. Finally, only a few image processing algorithms are biologically relevant, and even in these algorithms the biological link is limited to local motion estimation [2], [33], [21], [26].

Except for the studies by Black and Jepson [8], Bober *et al.* [9], and Chey *et al.* [11], image processing algorithms and neural-network methods perform motion analysis by using only motion cues. Numerous psychophysical and neurobiological studies reveal the importance of the integration of different cues, such as stereo, form, and brightness [8], [22].

The aim of our study is to find a solution to all of the three challenges. We address these challenges by using an oscillatory neural network to perform motion-based segmentation by integrating motion and brightness cues. Our model consists of two parallel pathways for motion and brightness, respectively. The motion pathway has two stages. Assuming translational motion, the first stage estimates motion by performing adaptive temporal block matching that is a variant of the Reichardt detector [23]. Unlike others, the size of our matching block is adapted at each location to obtain a reliable estimate, or, in the absence of strong evidence, motion is not estimated. In the second stage, locations are grouped based on motion similarity in a multilayer locally excitatory globally inhibitory oscillator network (LEGION) [31], [27]. LEGION is based on the idea of oscillatory correlation [19], [27], [31], whereby phases of neural oscillators encode region labeling. Oscillators corresponding to one region have the same phase, which is different from those of other regions. Due to the absence or the inaccuracy of motion estimates at some locations, a stationary scene analysis is performed in parallel with the brightness pathway, also using a LEGION network. The partial solutions from the two pathways are combined in the integration stage to refine motion estimates, based on which the final result is obtained in the motion network.

The following section describes the two building blocks of our model. Next is its detailed presentation. Subsequently, the performance of our model is demonstrated on both synthetic and real sequences. Finally, comparisons with other methods are made and conclusions are drawn.

II. METHOD COMPONENTS

A. Temporal Block Matcher

The Reichardt bilocal correlator [23] is composed of two symmetric parts: rightward and leftward motion detectors. The rightward motion of a moving dot is detected by correlating brightness at location and time $(i - \Delta i, t - \Delta t)$ with that of (i, t). Here, the model implicitly assumes the speed of $\Delta i / \Delta t$. Similarly, the leftward motion detector correlates the brightness at $(i + \Delta i, t - \Delta t)$ and (i, t). Depending on the sign of the difference between the right and the left correlation, the motion of the dot is determined to be rightward or leftward.

A temporal block matcher [3] compares brightness within a block, N_B, instead of a single location. Selectivity to dif-
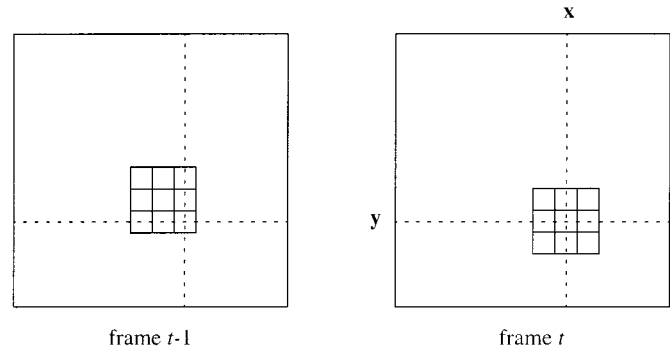

Fig. 2. Brightness within matching blocks ($N_B = 3 \times 3$) in two consecutive frames are cross-correlated by a temporal block matcher to detect a lower right motion of one pixel per frame.

ferent velocities is achieved by sliding correlation blocks in different quantities and different directions on the previous snapshot (frame) of the scene, as shown in Fig. 2. For a temporal block matcher employing two consecutive image frames, the correlation corresponding to displacement $\boldsymbol{r} = (r_i, r_j)$, at location (i, j) and time t can be expressed as

$$\tilde{v}_{\boldsymbol{r}}(i, j, t) = \sum_{(k, l) \in N_B(i, j)} \cdot \frac{I(k, l, t)}{|I(k, l, t) - I(k - r_i, l - r_j, t - 1)|} \quad (1)$$

where $I(i, j, t)$ is the brightness at location (i, j) in frame t and $N_B(i, j)$ is centered at location (i, j). The denominator of the above expression, as in the following equations, should include a small number to avoid division by zero. A large correlation, $\tilde{v}_{\boldsymbol{r}}(i, j, t)$, implies a high probability of the displacement, $\boldsymbol{r}$, at location (i, j) and time t.

One can view $\tilde{v}_{\boldsymbol{r}}$ as the response of a velocity detector. A spatial collection of detectors tuned to the same displacement can be visualized as a velocity layer. Different displacements correspond to different layers. Detectors from different layers corresponding to the same location form a velocity column. When a location has a single motion, only the corresponding detector in its velocity column should have a substantial response. For a location in a region producing motion transparency, multiple detectors should attain high responses in the column at that location.

B. LEGION

LEGION is based on the idea of oscillatory correlation, where the phases of the neural oscillators encode the grouping of corresponding locations. The building block of a LEGION network is a single relaxation oscillator, (i, j), defined as a feedback loop between an excitatory unit x_{ij} and an inhibitory unit y_{ij} [31], [27]

$$\dot{x}_{ij} = 3x_{ij} - x_{ij}^3 + 2 - y_{ij} + I_{ij} + S_{ij} + \rho \quad (2a)$$

$$\dot{y}_{ij} = \varepsilon \left(\alpha \left(1 + \tanh \left(\frac{x_{ij}}{\beta} \right) \right) - y_{ij} \right). \quad (2b)$$

Here, I_{ij} represents external stimulation to oscillator (i, j), S_{ij} denotes overall coupling, and ρ is the variance of Gaussian
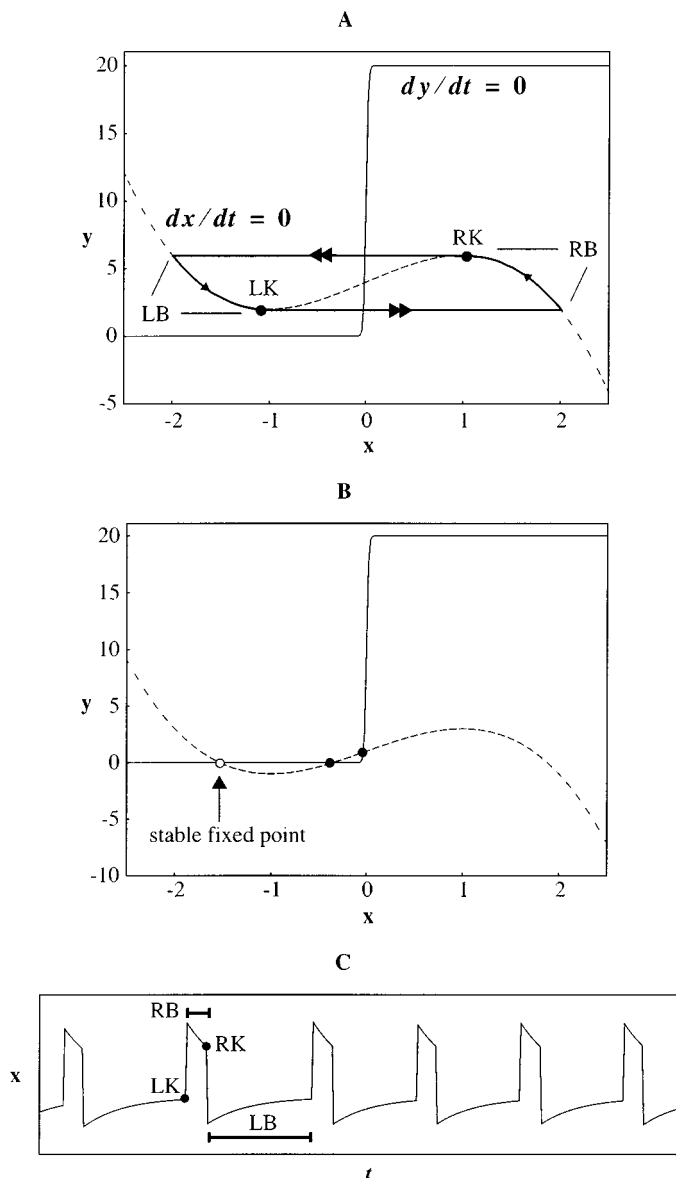
A

B

C

Fig. 3. Phase plane diagram of a single oscillator with $S_{ij} = 0$. The x-nullcline is the dotted line and the y-nullcline is the solid line. (a) When the nullclines intersect only along the middle branch of the x-nullcline, the oscillator produces a limit cycle, indicated as a thick solid line. The parameters are $\varepsilon = 0.005$, $\alpha = 10.0$, $\beta = 0.02$, $I_{ij} = 2.0$, and $\rho = 0.02$. (b) When the nullclines in addition intersect along the left branch of the x-nullcline, the oscillator approaches a stable fixed point and does not oscillate. All parameters are the same as in (a) except for $I_{ij} = -1.0$. (c) Temporal activity of an oscillator that oscillates. Salient portions of a limit cycle are labeled for one period.

noise. The parameter ε is chosen to be a small positive number so that (2) defines a relaxation oscillator with two time scales. The x-nullcline ($\dot{x} = 0$) of (2) is a cubic curve while the y-nullcline ($\dot{y} = 0$) is a sigmoid function, as shown in Fig. 3. When $I_{ij} > 0$ and $S_{ij} = 0$, the two nullclines intersect along the middle branch of the cubic and the system is oscillatory [Fig. 3(a)]. Otherwise ($I_{ij} < 0$ and $S_{ij} = 0$), the nullclines intersect at a stable fixed point along the left branch of the cubic and the system does not oscillate [Fig. 3(b)]. In the former case, oscillator (i, j) travels along the left branch (LB) until reaching the left knee (LK) and then jumps to the right branch

(RB). When it is on RB, the oscillator is called active. After traveling along RB, the oscillator reaches the right knee (RK), where it jumps back to LB. This completes a limit cycle. The temporal trace of several limit cycles is depicted in Fig. 3(c). α is used to control the ratio of the times that an oscillator spends on RB or on LB. The larger is α, the shorter is the time the oscillator spends on RB. Finally, β defines the steepness of the sigmoid function and is chosen to be small.

S_{ij} includes a variable called lateral potential and coupling from neighboring oscillators and a global inhibitor

$$S_{ij} = \sum_{kl \in N(i,j)} W_{ij,kl} H(x_{kl}) + W_p H(p_{ij} - 0.5) - W_z H(z - 0.5) \tag{3}$$

where

$W_{ij,kl}$	connection weight from oscillator (k, l) to oscillator (i, j);
H	heaviside step function;
W_p	weight for the lateral potential;
W_z	weight for the the global inhibition;
N	local coupling neighborhood (e.g. four nearest neighbors).

The lateral potential is introduced for each oscillator to distinguish a homogeneous region from a noisy one [32]

$$\dot{p}_{ij} = (1 - p_{ij}) H \left[\sum_{kl \in N_p(i,j)} H(x_{kl}) - \theta_p \right] - \varepsilon p_{ij} \tag{4}$$

where N_p is the potential neighborhood, which is generally larger than N. The potential of an oscillator is initially set high but continuously decays. When an oscillator is active and has a number of active neighbors greater than θ_p, its potential rises to one. Oscillators that maintain high potentials are referred to as leaders while others are called followers. Due to the potential term, only a sizable group that contains a leader can form a segment. Small and noisy groups cannot produce leaders and stop oscillating after a short beginning period. Such oscillators form the network background, which should be distinguished from the background part of an image (referred to as image background later).

Whether an oscillator is active at a particular time also depends on the activity of the global inhibitor, z, which is defined by

$$\dot{z} = H \left[\sum_{kl} H(x_{kl}) - 1 \right] - z. \tag{5}$$

When no oscillator is active, z decays to zero; otherwise it rises to one. A leader can jump to RB when the inhibitor is not triggered. A follower, however, can be active only when it has strong excitatory couplings with active oscillators in its neighborhood through the first term in the right-hand-side of (3). A common form of the LEGION network used for segmentation is a two-dimensional (2-D) array of oscillators and a global inhibitor, where each oscillator corresponds to one location in an input scene [see Fig. 6(b)].
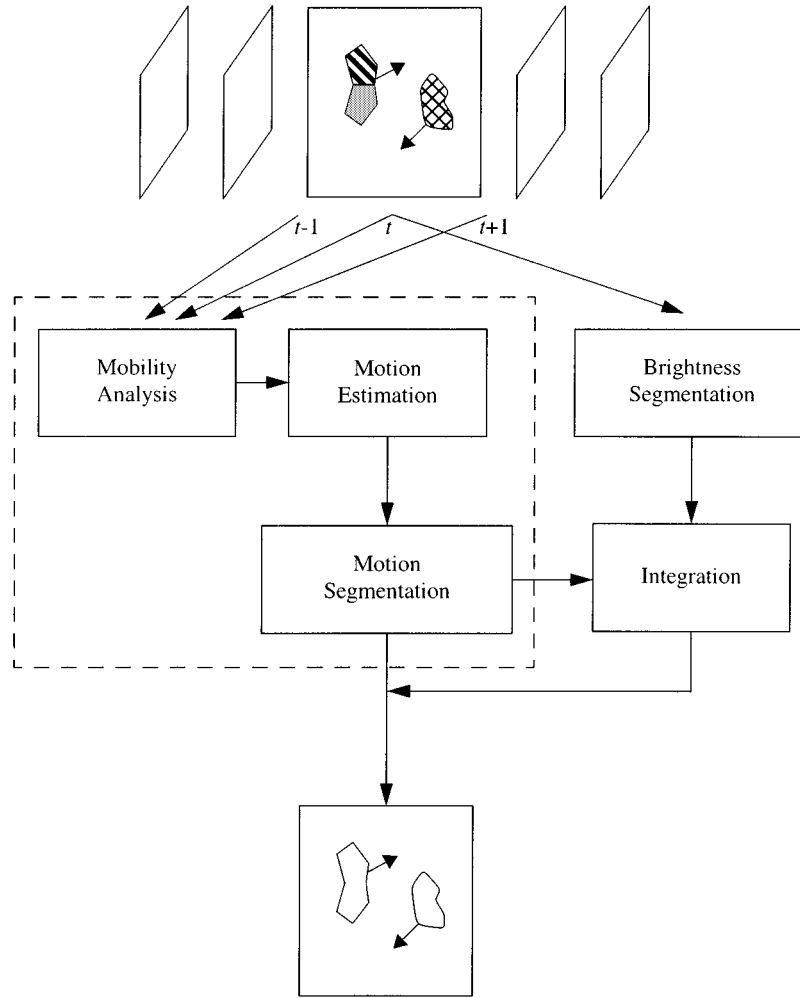
Fig. 4.   Flow diagram of the proposed model. Processing starts from the top and proceeds downward. Following the analyzes in the motion and the brightness pathways, results are combined in the integration stage to refine motion estimates. Finally, segmentation is performed in the motion network based on refined motion estimates.

As discussed previously [27], [32], [36], there is a large and robust range to choose the values of intrinsic LEGION parameters: α, β, ε, and ρ. As specified in Section IV-A, an algorithmic version of LEGION is used in our study to save computing time; the above four parameters are absorbed in the algorithmic version.

III. MODEL DESCRIPTION

The two parallel pathways in our model are shown in Fig. 4. One of the pathways, contained in a dash box, performs motion analysis. The other pathway segments the scene based on brightness distribution. Combining the results from the pathways, the integration stage refines initial motion estimates. The final segmentation is performed in the motion network based on refined motion estimates.

A. Motion Pathway

There is strong evidence from both psychology and neurophysiology supporting at least two separate stages of motion processing in the visual system (see [25] for a review). Moti-

vated by this, our method has two stages for motion estimation and grouping, respectively.

Spatial Pooling: Before we employ our adaptive temporal block matcher to estimate local motion, we choose a matching block, namely, a pooling neighborhood, at each location using a reliability criterion. This criterion serves two purposes. First, locations with strong motion evidence are identified. Second, pooling neighborhoods, N_B, at these locations are adaptively selected to produce more accurate estimates. Similar to the measure used by Irani *et al.* [16], we calculate a mobility value at each location and time, (i, j, t)

$$M(i, j, t) = \frac{\displaystyle\sum_{(k, l) \in N_M(i, j)} |I(k, l, t) - I(k, l, t - 1)|}{\displaystyle\sum_{(k, l) \in N_M(i, j)} I(k, l, t)} \quad (6)$$

where N_M is a mobility neighborhood. Our mobility analysis captures temporal brightness changes in a local neighborhood. Note that changes are normalized by summated local brightness. A large value of M at a location indicates a high probability of
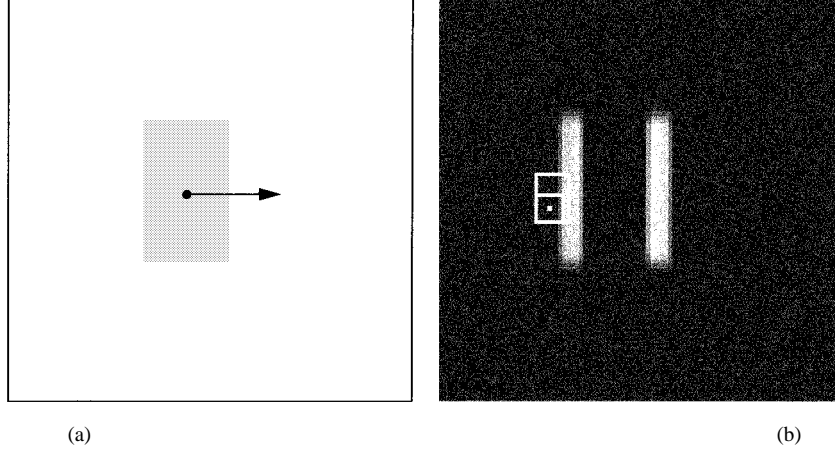
Fig. 5. Adaptive spatial pooling. (a) Middle frame of an input sequence. (b) Mobility image for the frame in (a), where intensities indicate the corresponding mobility values. The initial and the final pooling blocks are depicted as white rectangles for a location near the boundary (white dot). Note that the neighborhood extends along the boundary to encompass locations with large mobility values.

motion presence at that location. Otherwise, the presence of motion is uncertain since noise could also cause a temporal brightness change. Thus, we do not obtain motion estimates at locations with small mobility values.

Initially, a small neighborhood N_B is assumed to be sufficient for spatial pooling. When the sum of the mobility values M_s and the average mobility M_a over N_B are large enough, matching is performed. When M_s or M_a is small, N_B is expanded in the direction that maximally increases M_s. The expansion continues until either both M_s and M_a are sufficiently large or the size of N_B reaches an upper limit. In the latter case, estimation is not performed. In the former case, N_B is used for estimation. Fig. 5 illustrates the adaptive changes in N_B for an image composed of a rectangular region and a background, both homogeneous [Fig. 5(a)]. Among the three consecutive input frames, only the middle one is depicted. Fig. 5(b) shows the mobility image, where the intensity at a location indicates its mobility value. In this image, initial and final N_B's are depicted at a location near one boundary. Initially, $N_B = 5 \times 5$ and the stopping condition is not met. Since boundaries and surface markings of moving regions have large mobility values, N_B usually extends toward them. For the location considered in the image, N_B expands to cover a considerable portion of the boundary. When a location does not have large mobility values in its N_M, e.g. a location away from moving boundaries in Fig. 5(b), M_s or M_a is small. In this case, N_B will reach the upper limit and motion will not be estimated at that location. Such a location is inferred to be unreliable.

Motion Estimation: Having selected N_B at reliable locations, we apply an adaptive temporal block matcher to three consecutive frames, instead of two, to improve accuracy. We assume only translational motion in our method. The correlation corresponding to displacement $\boldsymbol{r}$ is

$$v_{\boldsymbol{r}}(i, j, t) = \tilde{v}_{\boldsymbol{r}}(i, j, t) + \tilde{v}_{\boldsymbol{r}}(i, j, t+1) \tag{7}$$

where $\tilde{v}_{\boldsymbol{r}}(i, j, t)$ and $\tilde{v}_{\boldsymbol{r}}(i, j, t+1)$ are the correlations at (i, j) between the image frames at t and $t-1$ and those at $t+1$ and t, as given in (1). We have a total number of $L = (2R+1)^2$ different displacements, hence velocities, corresponding to a set of displacements varying from $-R$ to R in the i- and j-direction. Our additional modification is to match not only N_B's but also local spatial correlation surfaces to make estimation more robust against temporal illumination variations. Consider a location, P, and its neighbors, Q's in a neighborhood centered at P, $N_S(P)$. Correspondingly, P and Q's have pooling neighborhoods $N_B(P)$ and $N_B(Q)$'s. A spatial correlation surface (SCS), which has the size of N_S, is obtained at P by matching $N_B(P)$ and $N_B(Q)$'s. Note that this operation is exactly the same as in (1). Instead of a neighborhood in the next frame, a neighborhood in the same frame but centered at a neighboring location is matched with $N_B(P)$. SCS's at all reliable locations including those in the neighboring frames are obtained similarly. Subsequently, both pooling neighborhoods, N_B, and SCS's of size N_S are matched across time for different displacements to obtain the temporal correlation surface at a location. Using (7), matching two SCS's results in $c_{\boldsymbol{r}}(i, j, t)$ for displacement $\boldsymbol{r}$ at (i, j, t). Consequently, the temporal correlation at (i, j, t) for displacement $\boldsymbol{r}$ is

$$V_{\boldsymbol{r}}(i, j, t) = v_{\boldsymbol{r}}(i, j, t) c_{\boldsymbol{r}}(i, j, t). \tag{8}$$

The displacement resulting in the maximum temporal correlation at a location yields the local motion estimate at that location.

Network: Under certain circumstances, a region can be perceived to have multiple (usually two) different motions simultaneously, a phenomenon known as motion transparency [13]. In order to allow for multiple overlapping motions, we adopt a multilayer structure in the velocity representation where a 2-D LEGION network corresponds to each layer as depicted in Fig. 6(a). Oscillators at the same location in all layers form a velocity column and correspond to a single location in a scene. The coupling weight between two oscillators, (i, j) and (k, l), on a velocity layer $\boldsymbol{r}$ is determined based on their temporal correlation for displacement $\boldsymbol{r}$. Dropping time t from the expressions for convenience, the coupling weight is given by

$$W_{\boldsymbol{r}, ij, kl} = \frac{V_{\boldsymbol{r}}(i, j) + V_{\boldsymbol{r}}(k, l)}{|V_{\boldsymbol{r}}(i, j) - V_{\boldsymbol{r}}(k, l)|}. \tag{9}$$
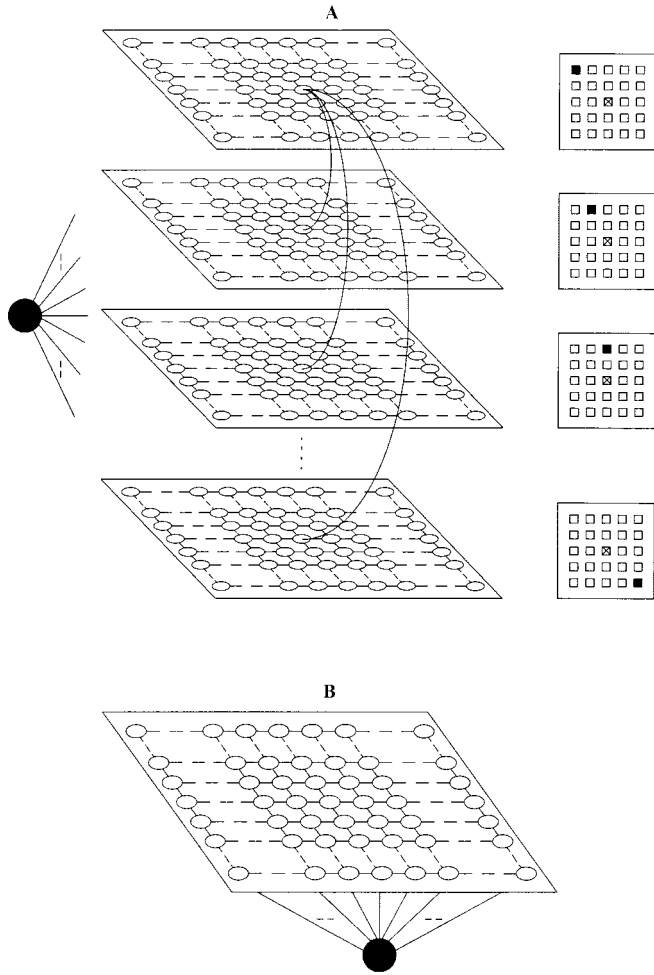
Fig. 6. Neural-network architecture of our model. Small circles represent oscillators and the big black circle the global inhibitor. The global inhibitor is connected to all oscillators in each pathway. (a) Multilayer motion network where each velocity layer corresponds to a LEGION network. Each layer is tuned to a particular velocity indicated by a filled square in a small rectangular grid representing the set of velocities varying from $-R$ to R in the i- and j-direction. Units interact vertically within a velocity column and horizontally within each velocity layer. Vertical interactions are shown only for one oscillator. (b) Brightness network. Each oscillator is locally coupled with its four nearest neighbors.

When oscillators have similar correlations for a particular displacement, they are strongly coupled in the corresponding velocity layer. At locations without estimates, couplings are set to zero. We replace S_{ij} in (3) by $S_{\boldsymbol{r},ij} = H(\tilde{S}_{\boldsymbol{r},ij} - \theta_M)$ where θ_M is a threshold and

$$\tilde{S}_{\boldsymbol{r},ij} = \sum_{kl \in N(i,j)} W_{\boldsymbol{r},ij,kl} H(x_{kl}) - W_z H(z - 0.5)$$

$$+ W_p H(p_{\boldsymbol{r},ij} - 0.5) H(M_{ij} - 0.25)$$

$$H\left[\sum_{\boldsymbol{q}=1}^{L} H(V_{\boldsymbol{r},ij} - V_{\boldsymbol{q},ij}) - L\right]. \quad (10)$$

Note that the major change is the two multiplicative terms in the potential part, which extends the standard definition of a leader. The first one ensures that only locations with large mobility values can become leaders. The second one allows for just a single leader within each velocity column. Due to the

inhibitory interaction within each column, only the oscillator with the largest correlation becomes a winner. If such a winner also has sufficiently high mobility, $M_{ij} > 0.25$, and potential, $p_{\boldsymbol{r},ij} > 0.5$, it becomes a leader and forms a segment. A leader recruits oscillators on its layer through local couplings. As in the original LEGION network [31], [27], [32], these recruited oscillators synchronize. Because of the global inhibition across all layers, different segments desynchronize.

The complete activity of all layers can be captured in an output network where each unit summates the activity of the entire corresponding velocity column. Since at most one oscillator is active in a column at any time, the output network displays segments without ambiguity.

Note that the selection of a single leader within each column does not prevent the network from representing multiple motions at a location. An oscillator can become active if it is a leader or a follower that is recruited by a leader on its layer. Since there can be several followers in a velocity column in addition to a leader, more than one oscillator at a single location can become active, thus representing different motions. This is key to our representation of motion transparency.

B. Brightness Pathway

In the brightness pathway, the middle frame of the sequence analyzed in the motion pathway is processed in terms of brightness. In this pathway, a 2-D LEGION network is employed as depicted in Fig. 6(b), where each oscillator corresponds to a single location in the input scene. Assuming that each region has similar brightness, the coupling weight between two oscillators at two neighboring locations, (i, j) and (k, l), is defined as

$$W_{ij,kl} = \frac{I(i, j) + I(k, l)}{|I(i, j) - I(k, l)|}. \quad (11)$$

Thus, locations that have relatively similar brightness establish a strong coupling weight. We employ the network model given in (3) where S_{ij} is replaced by $H(S_{ij} - \theta_B)$ and θ_B is a threshold. Similar to the motion network, a strongly coupled group of oscillators with at least one leader becomes synchronized. Oscillators corresponding to different regions become active at different times (desynchronization). Textured regions, on the other hand, tend not to produce leaders or segments, and thus they are distinguished from homogeneous ones. This property has a significant role in the subsequent integration stage.

C. Integration Stage

So far, the motion and the brightness pathways perform segmentation independently. Relatively homogeneous regions have low mobility values in their interior areas and thus, only motion estimates along their moving boundaries are obtained. Conversely, textured regions, due to their high brightness variations, generally cannot form their own segments and thus stay in the network background. The integration stage combines the partial results from the two pathways to yield more accurate results. Our integration stage has two steps, namely, unreliable estimate elimination and motion filling-in.

Elimination of Unreliable Estimates: Unreliable estimates are eliminated based on the brightness analysis. The mobility
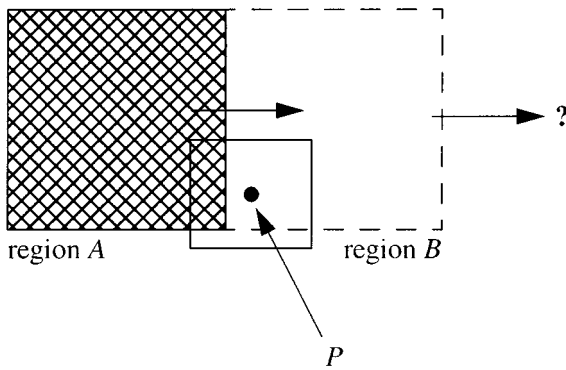
Fig. 7. Schematic illustration of induced motion. Regardless of the motion of region B, location P possesses an erroneously induced motion estimate due to the texture in moving region A and the homogeneity in region B.

analysis in the motion pathway is only effective in identifying inner locations of homogeneous regions. Locations along motion boundaries are assumed to be reliable due to their large mobility values. A typical case where this assumption does not hold is illustrated in Fig. 7. Assume that region A is textured and moves in the right direction. Without loss of generality, a neighboring region B is assumed to be homogeneous and stationary. Due to the texture, the locations across A have reliable motion estimates. The inner locations of B, on the other hand, do not have estimates since they have low mobility values. However, location P in B obtains an estimate because its pooling neighborhood, indicated by the small square in the figure, partially covers an area of A, which has large mobility values. Because of the motion in A and the homogeneity in B, an induced motion similar to that of A is estimated at P. Other locations similar to P are also assigned an induced motion. Note that these estimates are erroneous. The same analysis applies when B moves but with a different velocity than that of A.

Based on the above analysis, we propose an elimination rule to remove these erroneous estimates. First, the number of oscillators belonging to the network background of the brightness network is obtained for each motion segment. When they occupy more than half of a motion segment, we consider that the segment belongs to a moving textured surface. The locations that are in this motion segment but also belong to a brightness segment are inferred to have induced motion estimates and thus, their estimates are eliminated. Otherwise, these estimates stay intact. In other words, when a homogeneous region is occluded by a moving textured surface, our method eliminates the estimates in the homogeneous region along the boundary of the textured surface.

Motion Filling-In: Recall that there is a winner within each velocity column. Following the removal of unreliable estimates, the distribution of winner velocities within each brightness segment is obtained in a 2-D form. The dimensions correspond to translational velocity components in the i- and j-direction. The dominant motion of each brightness segment is identified by finding the maximum of the motion distribution. Subsequently, locations within each brightness segment are filled in with this dominant motion. As a result of filling in, estimates become available for all locations in homogeneous moving regions. This represents our solution to the blank wall problem. Since texture

regions cannot produce brightness segments, filling-in does not take place in these regions. On the other hand, locations within these regions should have reliable motion estimates.

To be consistent with the perception of homogeneous image backgrounds, an additional constraint is incorporated into our model. Depending on moving regions, there can be locations on a homogeneous background with induced motion estimates. As shown in Fig. 9, which is discussed later in detail, locations in the image background along the boundary of a texture region are assigned incorrect induced estimates. With the elimination process, these estimates are removed. When induced motions are caused by the motion of a homogeneous region occluding the image background (cf. Fig. 1), our method does not eliminate them. Consequently, the motion of the occluding region may incorrectly become the dominant motion for the image background. In other words, the image background appears to move with the occluding region. However, a homogeneous background extending beyond the image border should be perceived as stationary. In order to simulate this perception, we introduce, for homogeneous regions along the image border, zero-velocity estimates at these locations that do not have estimates. As a result, when an image background is homogeneous, a large portion of its area takes on the zero velocity and hence, the motion distribution of the image background is biased toward the zero velocity.

Final Output: Following the integration stage, couplings in the motion network are updated based on the refined estimates and the final segmentation result is obtained. Note that the time scale in the evolution of networks is faster than that of an input scene. The analysis described so far takes place in the network time, while time (frame) t with respect to the input sequence stays the same.

IV. RESULTS

The performance of our method is demonstrated on two sets of image sequences. The first set includes synthetic scenes where we illustrate our design principles. The second set is a collection of real scenes acquired under various viewing conditions.

A. Parameter Values

We choose $N_M = N_S = 5 \times 5$ which is also the initial size of N_B. The thresholds for M_s, M_a, and the maximum one-dimensional (1-D) length for N_B are $\theta_s = 10$, $\theta_a = 0.1$, $e_m = 15$, respectively. These thresholds are selected to maximize the size of N_B while allowing it to have locations only from a single moving surface.

The selection of the number of velocity layers, $L = (2R + 1)^2$, is not critical as long as the maximum speed in motion scenes can be accommodated. We employ $R = 4$ for synthetic, $R = 10$ for real scenes, corresponding to a set of displacements varying from -4 to $+4$ and from -10 to $+10$ in the i- and j-direction, respectively. Images are of size 100×100 for the first set and 160×120 for the second set, except for the first scene of the second set, which is of size 120×170. Due to image border effects, output images are smaller. In order to minimize these effects, we simplify our estimation method for the locations near

the image border. We obtain estimates by using only a fixed spatial neighborhood, $N_B = 5 \times 5$, when these locations have an average mobility larger than $\theta_a = 0.1$. Thus, the size of output images is limited only by the maximum speed considered. The larger the maximum speed is considered, the wider is the border band without estimates [see Fig. 10(d)]. As a result, we obtain output images of size 92×92, 140×100, and 100×150 for the above three different image sizes, respectively. Because we are not aware of psychophysical conditions under which a homogeneous background should be perceived as stationary, we empirically assign the zero velocity to locations in all homogeneous regions if these locations lack estimates and are within a distance of $R + 10$ pixels from the image border.

In the simulation of LEGION networks, an algorithmic version is employed for computational efficiency where the parameters ε, ρ, α, and β are incorporated into the algorithm (see [32] for details). The weights except for $W_{ij,kl}$'s in (3) are fixed. We choose $\sum_{(k,l) \in N(i,j)} W_{ij,kl} H(x_{kl}) + I_{ij} > W_z$ and $W_p < W_z$ to achieve proper synchronization and desynchronization. The first condition, which also applies to $W_{r,ij,kl}$'s, implies that an oscillator receiving strong local excitation can overcome the global inhibition. The second condition ensures that no leader of a different segment can jump to RB when there are oscillators on RB. Specifically, we use $I_{ij} = -0.25$, $W_p = 0.74$, and $W_z = 1.0$ in all simulations. Regarding various neighborhoods in our method, the smaller N is, the more accurate is the localization of region boundaries but the more sensitive is the network to image noise. The smaller N_p is, the smaller is the size of regions that can be segmented. In our simulations, $N = 3 \times 3$, $N_p = 7 \times 7$, and $\theta_P = 0.75 \times \|N_p\| = 0.75 \times 49 = 36.75$ (0.75 indicates 75%). Finally, for different scenes, we adjust $\theta_{M,1}$ and $\theta_{M,2}$, the thresholds for the initial and the final segmentation in the motion network, respectively, and θ_B, that of the brightness network. It is desirable that $\theta_{M,1}$ and θ_B initially do not group locations with significantly different motion and brightness. Since refined estimates are more accurate, the selection of $\theta_{M,2}$ is relatively easy. We consider three frames for each scene since our temporal cross-correlation requires only three snapshots.

B. Synthetic Image Sequences

We illustrate the segmentation process on three synthetic image sequences. The first sequence is shown in Fig. 8, where two homogeneous rectangular regions are moving on a homogeneous background. As depicted in Fig. 8(a), the larger rectangle is moving two pixels per frame rightward and the smaller one is moving three pixels per frame downward, occluding the former. First, mobility values at all locations are determined in the motion pathway and the result is depicted in Fig. 8(b). Following the selection of N_B's, estimates are obtained for reliable locations and are depicted as a needle diagram in the left image of Fig. 8(c). For better visualization, estimates in needle diagrams are displayed after a spatial subsampling by a factor of three in all results. Based on the initial estimates, the motion network segments the scene into regions as shown in the right image of Fig. 8(c). Due to the missing and induced estimates, the segmentation result is bad.
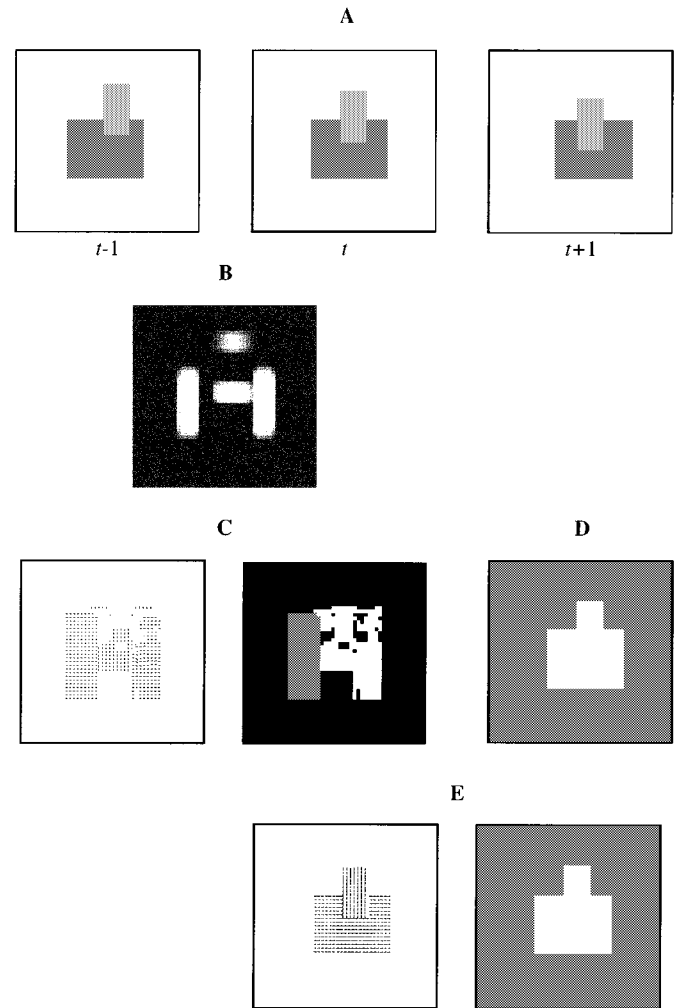

Fig. 8. Application of the proposed method to a synthetic image sequence, where the blank wall problem is present. (a) Three consecutive frames of the input sequence. (b) Mobility image. (c) Initial motion estimates and the segmentation result in the motion pathway. (d) Segmentation result in the brightness pathway. (e) Refined motion estimates and the final segmentation result. Here, $\theta_{M,1} = \theta_{M,2} = \theta_B = 10$.

In parallel, the brightness pathway segments the middle frame of the input sequence as shown in Fig. 8(d). The regions are accurately segmented in the brightness network. Because of the empty network background in the brightness pathway, the first step of the integration stage does not modify any estimate. After including zero velocity estimates in the image background along the image border, motion distributions within brightness segments are obtained in the second step of the integration. Due to this bias, the dominant motion in the segment corresponding to the image background is identified to be of zero velocity, which, in turn, is assigned to all locations in the image background. Similarly, the dominant motions of rectangular regions are determined and spread over all locations in these regions. The refined estimates over the entire scene are provided in the left image of Fig. 8(e). Subsequently, the motion network correctly segments the scene based on the refined estimates as depicted in the right image of Fig. 8(e). Note that, without the zero-velocity bias, the image background would not have the zero velocity. Also, incorrectly induced motions due to
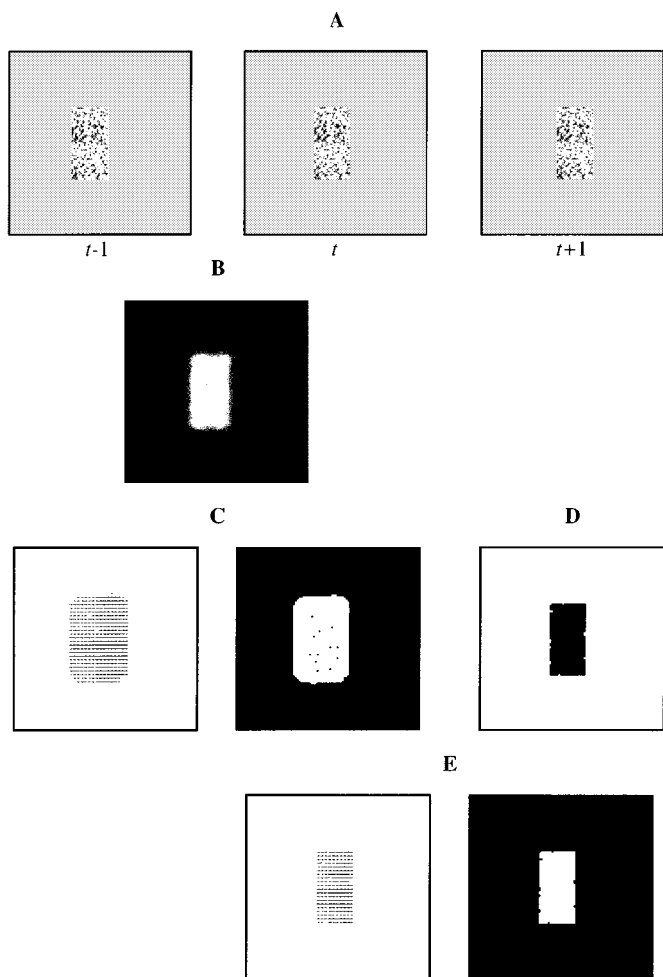
Fig. 9. Application of the proposed method to a synthetic image sequence, illustrating the elimination of erroneously induced motion estimates. (a) Three consecutive frames of the input sequence. (b) Mobility image. (c) Initial motion estimates and the segmentation result in the motion pathway. (d) Segmentation result in the brightness pathway. (e) Refined estimates and the final segmentation result where $\theta_{M,1} = \theta_{M,2} = \theta_B = 10$.
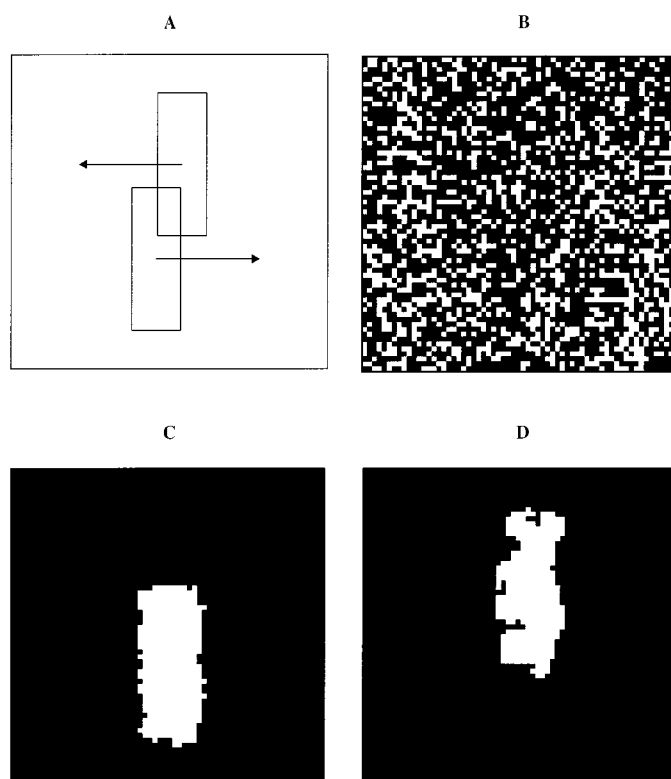

Fig. 10. Application of the proposed method to a synthetic scene, illustrating motion transparency. (a) Schematic diagram of the input, where two oppositely moving rectangles partially overlap. (b) Middle frame of the input sequence. (c)–(d) Final segmentation result in the motion network. The overlapping part is grouped with both rectangles. Here, $\theta_{M,1} = \theta_{M,2} = 20$ and $\theta_B = 50$.

rectangular regions would spread over the image background, causing the large rectangle to merge with the background. This sequence illustrates how the brightness pathway is utilized to address the blank wall problem.

In the second sequence, a textured rectangular region is moving two pixels per frame rightward on a homogeneous background as depicted in Fig. 9(a). Locations in the rectangular region are determined to have large mobility values as shown in Fig. 9(b). Motion is estimated at all locations with large mobility values. Since the locations in the image background along the boundary of the rectangular region also have large mobility values, they are assigned incorrect motions (cf. Fig. 7). Due to motion similarity, the motion segment corresponding to the rectangular region includes these background locations as shown in the right image of Fig. 9(c). Unlike the scene in Fig. 8, the brightness pathway is only able to group the locations in the image background. Since the oscillators corresponding to the texture region do not have strong couplings, they form the background in the brightness network indicated as the dark region in Fig. 9(d). Next, the number of the oscillators in the

network background of the brightness network is obtained for the motion segment corresponding to the rectangle. Since they occupy the majority of the segment, the estimates for the locations in this segment that are part of the image background are eliminated, and the motion distribution within the brightness segment corresponding to the image background is obtained. Since the induced motions are eliminated, the injected zero velocity becomes the dominant motion and fills in the entire image background. The final result is depicted in the left image of Fig. 9(e). Based on the refined estimates, the rectangular region and the background are accurately segregated as shown by the right image of Fig. 9(e). This scene illustrates how two partial segmentation results from the two pathways are integrated to obtain accurate segmentation.

In the last synthetic scene shown in Fig. 10, two random-dot rectangular regions are horizontally moving two pixels per frame in the opposite directions on a background of random-dot noise. During this motion sequence, there is an overlapping area where both motions are simultaneously perceived [13], which indicates motion transparency. A schematic diagram of the input scene and the middle frame of the sequence are provided in Fig. 10(a) and (b), respectively. Note that the rectangles cannot be segregated using a stationary brightness analysis. Thus, the initial estimates are not modified in the integration stage. Our method obtains two rectangular segments in the corresponding layers. As explained in the dynamics of the motion network, leaders in nonoverlapping areas of the
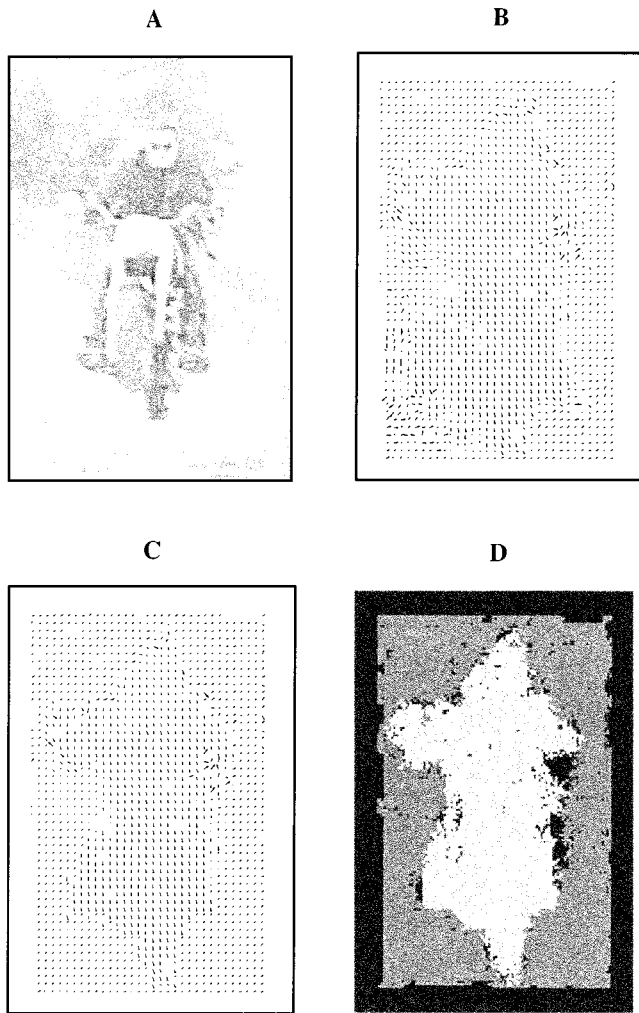
Fig. 11. A real scene sequence. (a) Middle frame of the input sequence. (b) Initial motion estimates. (c) Refined motion estimates after integrating the results from the two pathways. (d) Final segmentation. Here, $\theta_{M,1} = 20$, $\theta_{M,2} = 10$, and $\theta_B = 50$.

two rectangles recruit the locations in each area as well as the overlapping area. Thus, the overlapping area is correctly grouped with both rectangles [Fig. 10(c) and (d)].

C. Real Image Sequences

The second set of results illustrates the performance of our method on real scenes. The first sequence is from a scene, where a motorcycle rider jumps to a (dry) canal with his motorcycle while the camera is tracking him. This is shown in Fig. 11(a). Here, only the middle frame of the input sequence is depicted. Due to the camera motion, the rider and his motorcycle have a downward motion with a small rightward component and the image background has an upright diagonal motion. Besides the mobility analysis and estimation in the motion pathway [Fig. 11(b)], the scene is segmented into regions in the brightness pathway. Note that locations on the canal wall (lighter homogeneous region) occluded by the lower part of the motorcycle have induced motion estimates. In this particular area, since the motorcycle is part of the network background in the brightness network and the canal wall forms

a brightness segment, the estimates at the occluded locations of the canal wall are identified to be incorrect and eliminated in the integration stage. These locations, in turn, are assigned the dominant motion of their brightness segment [Fig. 11(c)]. Note that some locations near the image border do not have estimates since their average mobility is small. Despite the zero-velocity bias, none of the homogeneous regions in the image background, such as the sky in the upper right corner, assume the zero velocity; their motions are correctly estimated. Finally, based on the refined estimates, the rider with his motorcycle is accurately segmented from the image background as depicted in Fig. 11(d). As in the segment of the rider and his motorcycle, regions with different texture and brightness are grouped into a single segment due to common motion. Although some regions have similar brightness, e.g., the jacket of the rider and the trees, they are separated due to different motion. The boundary between the rider segment and the background segment in the right part of the scene is not as well defined, due to motion blurring and the inhomogeneous image background.

Fig. 12 shows the results for the four remaining real scenes. Only the middle frame of each input sequence, the refined estimates, and the final segmentation result are depicted. In the first scene, an airplane is flying diagonally in the lower left direction on the relatively homogeneous sky and there is a mountain in the lower right corner. The mountain appears to be moving rightward due to the camera motion. Following the initial segmentation in the two pathways, the integration stage refines the estimates based on which the three segments, namely, the airplane, the mountain, and the sky are segmented. Without the integration stage, the airplane segment would have a round shape due to erroneously induced estimates. Because of the zero-velocity bias, the sky segment is stationary, consistent with its perceived motion. As in the case of the sky and the middle part of the mountain, when regions have different motions, brightness similarity does not lead to incorrect grouping.

In the second scene depicted in the second row of Fig. 12, as a result of the camera motion, a woman and a big and a small satellite dishes all have different motions due to their depth differences. Along the boundary between the woman's hat and the big dish, there are induced motions, which are not eliminated in the first integration step. However, the brightness pathway accurately segments the two regions. Following that, the dominant motion in each segment fills in not only the inner locations but also the locations near the common boundary. Due to the zero-velocity bias, the homogenous image background is correctly analyzed. The woman, the dishes, and the sky are the resulting segments.

In the third scene, the camera tracks a person falling off a dam. Due to the camera motion, the dam has an upward motion while the person appears to be stationary. Since the person occupies a relatively small area in the scene, spatial pooling within a fixed neighborhood would fail to detect the person. However, our adaptive pooling scheme is well suited to the task. The integration stage eliminates the induced motions at the locations along the inner boundary of the person and segregates the person from the image background. Consistent with perception, the sky in the upper left corner is grouped with the dam despite the zero velocity bias.
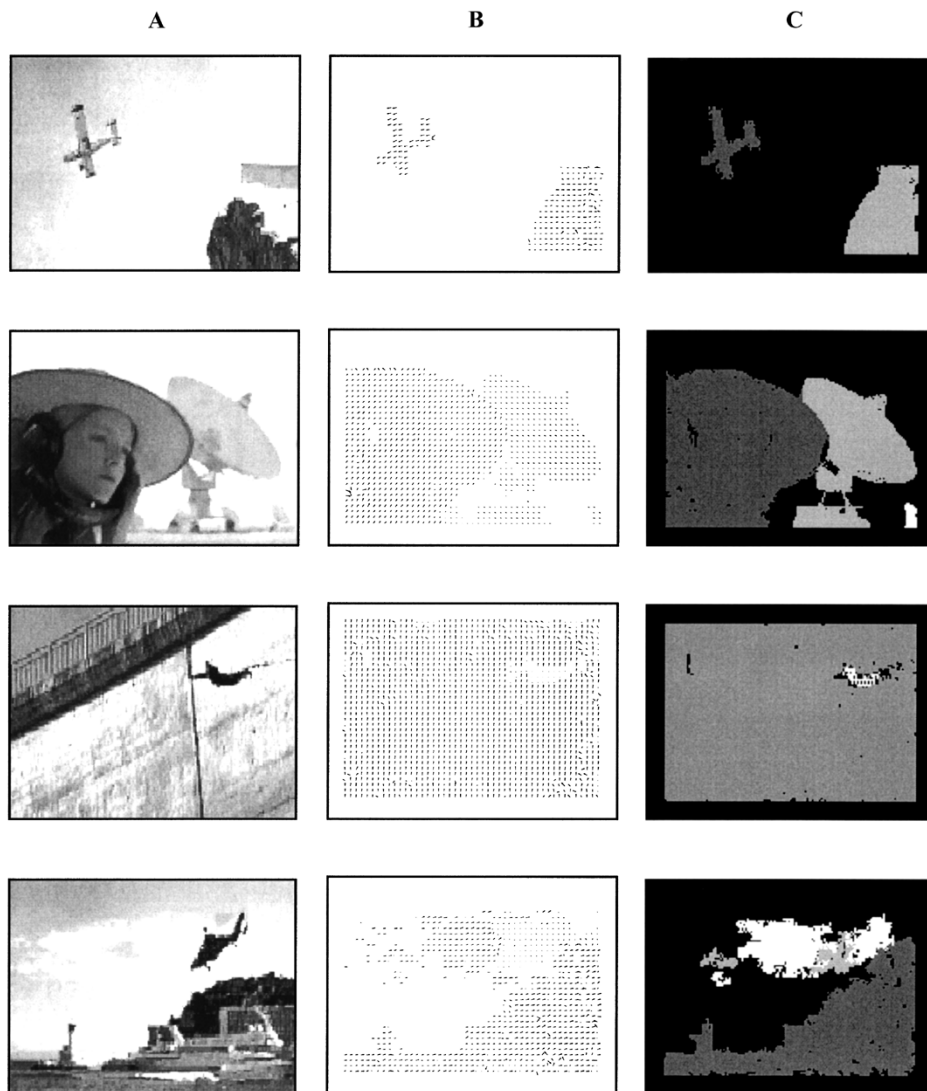
Fig. 12. Additional real scene sequences. (a) Middle frame of the input sequences. (b) Final motion estimates after integrating the results from the two pathways. (c) Final segmentations. Here, the corresponding thresholds $(\theta_{M,1}, \theta_{M,2}, \theta_B)$ for each sequence are (20, 2, 20), (20, 10, 10), (10, 2, 10), and (10, 4, 30), respectively.

In the last scene of Fig. 12, the camera tracks a helicopter flying in the cloudy sky over a port. Due to the camera motion, the clouds, the mountain, the port, and the lighthouse appear to move rightward while the helicopter appears relatively stationary. As shown in the middle image, the helicopter is detected to be approximately stationary while the other regions are estimated to have a rightward motion. The sky is assigned the zero velocity. The helicopter segments from the remaining regions which, despite their similar motion, result in five different segments. Note that our method does not group locations with similar motion when they belong to spatially discontiguous regions.

V. COMPARISON

Our method differs from others in several aspects. First, local motion is estimated in a pooling neighborhood that expands adaptively, conditioned on a mobility analysis. Second, motion analysis and brightness analysis are integrated in a unique way. Third, the building architecture of our model, LEGION, performs flexible grouping and segmentation in time, and is sup-

ported by neurobiological findings of coherent oscillations in the visual cortex. Taken together, our model is the only one, to our knowledge, that addresses simultaneously the three challenges described in Section I.

Since only image processing algorithms are tested on real motion scenes, we compare in the following our method against these algorithms. To focus our comparison, we choose three representative algorithms proposed by Horn and Schunck [15], Anandan [3], and Black [5], [7], [8], respectively, and the scene given in Fig. 11(a). We use the implementation of the first two algorithms by Barron [4] available online and the last one by Black [6] also available online. The algorithm by Horn and Schunck is a pioneering study on motion analysis. The study by Anandan, like ours, employs a temporal block matcher. Finally, Black's algorithm represents recent approaches that use robust statistics. Using the parameter sets suggested in their implementations, we employ eleven, two, and three consecutive frames for the three algorithms, respectively.

The results of the algorithms of Horn and Schunck, Anandan, and Black for the scene in Fig. 11(a) are given in Fig. 13(a)–(c),

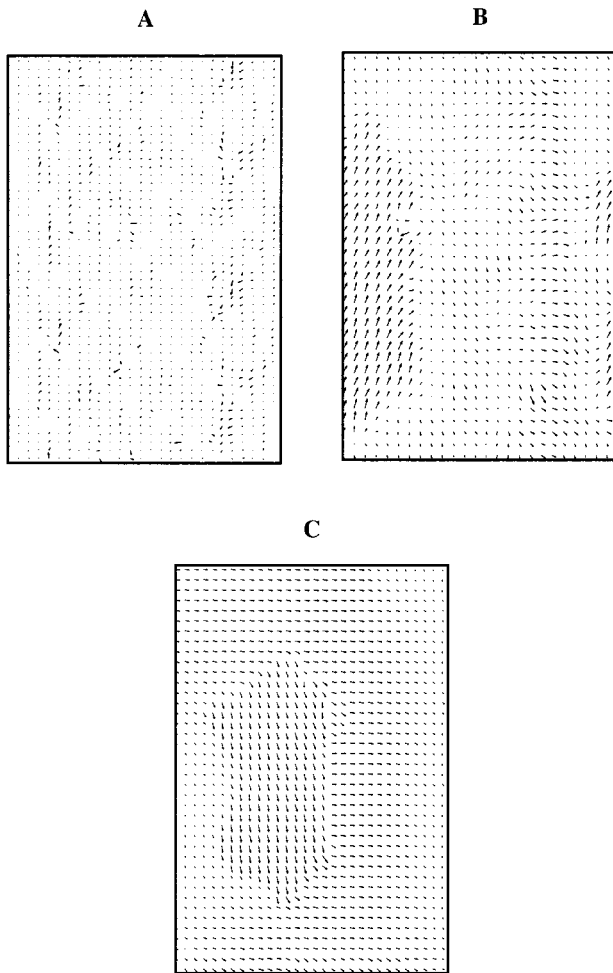A   B

C

Fig. 13.   Performance comparison using the sequence in Fig. 11(a). (a) Result of the Horn and Schunck algorithm. (b) Result of the Anandan algorithm. (c) Result of the Black algorithm.

respectively. Due to the use of a fixed pooling neighborhood and the relatively poor assessment of derivatives, the Horn and Schunck algorithm cannot capture the motion of the regions [Fig. 13(a)]. Similarly, the Anandan algorithm cannot accurately localize the motion boundaries [Fig. 13(b)]. This poor performance of the Anandan algorithm seems to result from a fixed pooling neighborhood, the employment of only two image frames, and the use of only motion cues. On the other hand, this algorithm estimates motion at the inner locations of surfaces, e.g., the rider, accurately. The Black algorithm offers the best result among the three. Unlike the other algorithms or ours, the Black algorithm requires the number of regions as an input parameter. When two regions are assumed, it can group the locations into two segments that best account for the motion distribution in the scene. Thus, despite the incorrect motion direction assigned to the background, this algorithm successfully segregates the rider and his motorcycle from the background. Since the algorithm pools locations that are not necessarily proximate, the resulting estimates vary smoothly across the scene. As a result, local motions are not estimated accurately and the motion boundary between the rider/motorcycle segment and the image background is not well localized.

By comparing Figs. 11 and 13, it is clear that the performance of our model is highly competitive. Except for the limitation that our model does not produce motion estimates near the image border while others do, our neural-network model yields the most accurate motion boundaries.

VI. CONCLUDING REMARKS

A. Contributions

Our proposed method is the first study that simultaneously addresses the aperture, blank wall, and motion transparency problems. The method is composed of two biologically plausible building blocks, i.e., an adaptive temporal block matcher and LEGION. The former is based on Reichardt's classic study of the fly visual system [23]. We have extended common versions of the temporal block matcher by using adaptive pooling and spatial correlation surfaces.

We have employed the LEGION architecture for motion-based segmentation, and extended the one-layer version to a multilayer LEGION network that allows for multiple motions at the same location, and thus naturally represents motion transparency. LEGION perform segmentation in phase space in a flexible way, and is supported by synchronous oscillations recorded in the visual cortex [12], [14].

Our segmentation method integrates partial solutions from the motion pathway and the brightness pathway. This integration contributes to motion analysis in two ways. First, motion ambiguity in homogeneous regions, i.e., the blank wall problem, is addressed. Second, the number of erroneously induced motion estimates is minimized, significantly improving boundary localization. We have also introduced the zero-velocity bias for a homogeneous image background to obtain perceptually valid results.

B. Limitations

Our current model has a number of limitations, which need to be addressed in future research. Translational motion is used in our model for simplicity, and must be replaced by a more realistic affine model in order to capture a larger set of motion types. Our method considers only three consecutive frames for each motion scene, while it is clear that longer sequences of frames are needed to process more complex motion scenes. Our method may need to be extended in order to model motion transparency induced by overlapping homogeneous regions, such as moving plaids. In the integration stage, we have employed a simple technique to resolve the depth relationship among surfaces, where the network background is exploited to infer the occlusion relationship in neighboring textured and homogeneous regions. More sophisticated techniques, such as detailed analysis of motion-induced changes of surface occlusion [17], may be more effective generally.

REFERENCES

[1] E. H. Adelson and J. Movshon, "Phenomenal coherence of moving visual patterns," *Nature*, vol. 300, pp. 523–525, 1982.

[2] E. H. Adelson and J. R. Bergen, "Spatiotemporal energy models for the perception of motion," *J. Opt. Soc. Amer. A.*, vol. 2, no. 2, pp. 284–299, 1985.

[3] P. Anandan, "Measuring visual motion from image sequences," Ph.D. dissertation, COINS Dept., Univ. Massachusetts, Amherst, MA, 1987.

[4] J. L. Barron.. [Online]ftp://ftp.csd.uwo.ca/pub/vision

[5] M. J. Black, "Robust incremental optical flow," Ph.D. dissertation, Dept. Comput. Sci., Yale Univ., New Heaven, CT, 1992.

[6] M. J. Black.. [Online]http://www.parc.xerox.com/spl/members/black/code.html

[7] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vis. Image Understand.*, vol. 63, no. 1, pp. 75–104, 1996.

[8] M. J. Black and A. D. Jepson, "Estimating optical flow in segmented images using variable-order parametric models with local deformations," *IEEE Trans. Pattern. Anal. Machine Intell.*, vol. 19, pp. 972–986, Oct. 1996.

[9] M. Bober, M. Petrou, and J. Kittler, "Non-linear motion estimation using the supercoupling transform," *IEEE Trans. Pattern. Anal. Machine Intell.*, vol. 20, pp. 550–555, 1997.

[10] T. Chen, W. Lin, and C. Chen, "Artificial neural networks for 3-D motion analysis—Part I: Rigid motion," *IEEE Trans. Neural Networks*, vol. 6, pp. 1386–1393, Nov. 1995.

[11] J. Chey, S. Grossberg, and E. Mingolla, "Neural dynamics of motion grouping: From aperture ambiguity to object speed and direction," *J. Opt. Soc. Amer. A.*, vol. 14, no. 10, pp. 2570–2594, 1997.

[12] E. Eckhorn *et al.*, "Coherent oscillations: A mechanism of feature linking in the visual cortex?," *Biol. Cybern.*, vol. 60, pp. 121–130, 1988.

[13] E. J. Gibson, J. J. Gibson, O. W. Smith, and H. Flock, "Motion parallax as a determinant of perceived depth," *J. Exp. Psychol.*, vol. 58, no. 1, pp. 40–51, 1959.

[14] C. M. Gray, P. König, A. K. Engel, and W. Singer, "Oscillatory responses in cat visual cortex exhibit inter-columnar synchronization which reflects global stimulus properties," *Nature*, vol. 338, pp. 334–337, 1989.

[15] B. K. P. Horn and B. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, pp. 185–203, 1981.

[16] M. Irani, B. Rousso, and S. Peleg, "Computing occluding and transparent motions," *Int. J. Comput. Vis.*, vol. 12, no. 1, pp. 5–16, 1994.

[17] G. A. Kaplan, "Kinetic disruption of optical texture: The perception of depth at an edge," *Percept. Psychophys.*, vol. 6, no. 4, pp. 193–198, 1969.

[18] K. Kita, S. Osamu, and Y. Nishikawa, "Neural-network model of long-range apparent motion in the human vision," in *Proc. IEEE Int. Conf. Neural Net.*, 1994, pp. 2254–2257.

[19] C. von der Malsburg and W. Schneider, "A neural cocktail-party processor," *Biol. Cybern.*, vol. 54, pp. 29–40, 1986.

[20] J. A. Marshall, "Self-organizing neural networks for perception of visual motion," *Neural Networks*, vol. 3, pp. 45–74, 1991.

[21] S. J. Nowlan and T. J. Sejnowski, "Filter selection for motion segmentation and velocity integration," *J. Opt. Soc. Amer. A.*, vol. 11, no. 12, pp. 3177–3200, 1994.

[22] N. Qian, R. A. andersen, and E. H. Adelson, "Transparent motion perception as detection of unbalanced motion signals. III. Modeling," *J. Neurosci.*, vol. 14, no. 12, pp. 7381–7392, 1994.

[23] W. Reichardt, "Autokorrelationsauswertung als Funktionsprinzip des Zentralnervensystems," *Z. Naturforsch.*, vol. 12b, pp. 447–457, 1957.

[24] R. Samy, "Infrared target motion estimation using neural network," *Proc. SPIE*, vol. 1294, pp. 238–245, 1990.

[25] M. E. Sereno, *Neural Computation of Pattern Motion: Modeling Stages of Motion Analysis in the Primate Visual Cortex.* Cambridge, MA: MIT Press, 1993.

[26] E. P. Simoncelli, "Distributed representation and analysis of visual motion," Ph.D. dissertation, Dept. Elect. Eng. Comp. Sci., Mass. Inst. Technol., Cambridge, 1993.

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

[28] R. Wang, "A network model of motion processing in area MT of primates," *J. Comput. Neurosci.*, vol. 4, pp. 287–308, 1997.

[29] J. Y. A. Wang and E. H. Adelson, "Representing moving images with layers," *IEEE Trans. Image Processing*, vol. 3, pp. 635–638, 1994.

[30] H. T. Wang, M. Mathur, and C. Koch, "Computing optical flow in the primate visual system," *Neural Comput.*, vol. 1, pp. 92–103, 1989.

[31] D. L. Wang and D. Terman, "Locally excitatory globally inhibitory oscillator networks," *IEEE Trans. Neural Networks*, vol. 6, pp. 283–286, Jan. 1995.

[32] ——, "Image segmentation based on oscillatory correlation," *Neural Comput.*, vol. 9, pp. 805–836, 1997.

[33] A. B. Watson and A. J. Ahumada, "Model of human visual motion sensing," *J. Opt. Soc. Amer. A.*, vol. 2, no. 2, pp. 322–341, 1985.

[34] Y. Weiss, "Bayesian motion estimation and segmentation," Ph.D. dissertation, Dept. Brain Cognitive Sci., Mass. Inst. Technol., Cambridge, 1998.

[35] A. Yamane, N. Ohnishi, and N. Sugie, "Robust extraction of moving objects through grouping edges along with optical flow," in *Proc. IEEE Int. Joint Conf. Neural Networks*, 1993, pp. 1267–1272.

[36] D. L. Wang and D. Terman, "Errata to 'Image segmentation based on oscillatory correlation'," *Neural Comput.*, vol. 9, pp. 1623–1626, 1997.

Edrogan Çesmeli (S'96–M'99) received the B.S. degree in electrical electronics engineering from Bilkient University, Turkey, in 1992. He received the M.S. and Ph.D. degrees in biomedical engineering from The Ohio State University, Columbus, in 1995 and 1999, respectively.

He has been with the Corporate Research and Development Department, General Electric Company since 1999. His research interests include biologically plausible neural networks, statistical image processing, and medical imaging.

DeLiang Wang (M'94–A'95) received the B.S. degree in 1983 and the M.S. degree in 1986 from Peking (Beijing) University, Beijing, China, and the Ph.D. degree in 1991 from the University of Southern California, Los Angeles, all in computer science.

From July 1986 to December 1987, he was with the Institute of Computing Technology, Academia Sinica, Beijing. Since 1991, he has been with the Department of Computer and Information Science and the Center for Cognitive Science, The Ohio State University, Columbus, where he is currently an Associate Professor. From October 1998 to September 1999, he was a Visiting Scholar in the Vision Sciences Laboratory, Harvard University, Cambridge, MA. His research interests include neural networks for perception, neurodynamics, neuroengineering, and computational neuroscience.

Dr. Wang is a member of IEEE Computer Society and the International Neural Network Society. He is a recipient of the 1996 U.S. Office of Naval Research Young Investigator Award.