Pulse-coupled neural nets: translation, rotation, scale, distortion, and intensity signal invariance for images

John L. Johnson

The linking-field neural network model of Eckhorn *et al.* [Neural Comput. **2**, 293–307 (1990)] was introduced to explain the experimentally observed synchronous activity among neural assemblies in the cat cortex induced by feature-dependent visual activity. The model produces synchronous bursts of pulses from neurons with similar activity, effectively grouping them by phase and pulse frequency. It gives a basic new function: grouping by similarity. The synchronous bursts are obtained in the limit of strong linking strengths. The linking-field model in the limit of moderate-to-weak linking characterized by few if any multiple bursts is investigated. In this limit dynamic, locally periodic traveling waves exist whose time signal encodes the geometrical structure of a two-dimensional input image. The signal can be made insensitive to translation, scale, rotation, distortion, and intensity. The waves transmit information beyond the physical interconnect distance. The model is implemented in an optical hybrid demonstration system. Results of the simulations and the optical system are presented.

1. Introduction

Nonpulsed neural network models use as their output signal a variable that represents the average firing rate of pulsed neurons. They have adaptive weighted interconnects and a nonlinear sigmoidal relationship to map the internal activity to the output activity. Excitatory and inhibitory receptive fields are employed for a variety of processing functions. The applications emphasize adaptive classification through supervised and unsupervised learning and recall. A pulsed model must retain the versatility and the functionality of these models, and in the limit of pulse averaging it must reduce to the nonpulsed models.

A pulse-based neural network must consider the pulse-forming mechanisms, the methods of transmission and reception, and the interactions based on the phase properties of the pulse trains. The linking-field model of Eckhorn *et al.*[1] has strong neurophysiological support. It was proposed as a minimal model to explain the experimentally observed synchronous feature-dependent activity of neural assemblies over large cortical distances in the cat cortex.[2] It is a cortical model. It emphasizes synchronizations of oscillatory spindles that occur in the limit of strong linking fields and distinguishes two major types of activity: forced, or stimulus-locked, synchronous activity, which is produced by abrupt temporal changes such as movement, and induced synchronous activity, which occurs when the pulse-train structures of the outputs of groups of cells are similar.[3]

This paper is concerned with the behavior of the linking-field model in the limit of weak-to-moderate linking strengths.[4,5] Strong linking is characterized by synchronous bursts of pulses. When the linking strength is reduced, the neurons no longer fire in bursts but still have a high degree of phase and frequency locking. This is the regime of moderate linking strength. Further reduction continuously lowers the degree of linking to where locking can occur only for small phase and frequency differences. This is the weak linking regime in which it is possible to encode spatial input distributions into corresponding temporal patterns with enough structure to have an object-specific time series for each input pattern. The pulse phase patterns in the time series are often found to be periodic and coherent in nature. Both in simulations and in an optical hybrid laboratory demonstration system, coherence and periodicity are ob-

The author is with the U. S. Army Missile Command, Research, Development, and Engineering Center, Weapons Sciences Directorate, AMSMI-RD-WS-PO, Redstone Arsenal, Alabama 35898-5248.

Received 12 July 1993; revised manuscript received 23 March 1994.

0003-6935/94/266239-15$06.00/0.

served to be the rule rather than the exception. The time series can be made invariant against translation, rotation, and scale changes of the input-image distribution by an appropriate choice of the structure of the receptive-field weight patterns. Invariance against scene illumination and image distortion was observed in simulations. Section 2 reviews the basic model. Time series and invariances are covered in Sections 3 and 4. The optical hybrid system's performance is shown in Section 5.

2. Basic Model

This section reviews the basic model of Eckhorn *et al.* as discussed in Ref. 1–3 and 6–8. The model neuron is a neuromime,[9] modified with two receptive fields per neuron and a linking mechanism added. It is shown in Fig. 1. There are three parts to the model neuron: the dendritic tree, the linking modulation, and the pulse generator. Each part is described separately, and then the operation of the complete model is discussed.

The dendritic tree is divided into two principal branches in order to make two distinct inputs to the linking part of the jth neuron. They are the primary input, termed the feeding input F_j, and the secondary input, which is the linking input L_j. These are given in Eqs. (1) and (2), respectively, for the case of continuous time. For discrete time steps the digital filter model is used, as given in the appendix of Ref. 1. Each input is a weighted sum from the synaptic connections on its dendritic branch. The synapses themselves are modeled as leaky integrators. An electrical version of a leaky integrator is a capacitor and a resistor in parallel, charged by a brief voltage pulse and decaying exponentially. Likewise, when a synapse receives a pulse, it is charged and its output amplitude rises steeply. The amount of rise depends on the amplitude-gain factor assigned to the synapse. It then decays exponentially according to its time constant. These postsynaptic signals are summed to form the total signal out of that branch of the dendritic tree, as indicated in Fig. 1. The amplitude-gain factors and the decay time constants of the synapses characterize the signals. The synapses in the feeding branch are assumed[1] to have smaller time constants than those of the linking branch. This assumption lets the feeding signal have a long decay tail on which the spikelike linking input can operate through the linking modulation process. The linking and the feeding inputs are given by

$$L_j = \sum_k L_{kj} = \sum_k [W_{kj} \exp(-\alpha_{kj}^L t)] * Y_k(t), \quad (1)$$

$$F_j = \sum_k F_{kj} = \sum_k [M_{kj} \exp(-\alpha_{kj}^F t)] * Y_k(t) + I_j, \quad (2)$$

respectively where W_{kj} and M_{kj} are the synaptic gain strengths, or weights, for the kth synapse of the linking and the feeding receptive fields, respectively. $Y_k(t)$ is the input pulse, α_{kj}^L and α_{kj}^F are the time constants, and $*$ is the convolution operator. I_j is an analog feeding input to the jth neuron. It is shown here as a distinct single term, but in general it can be a weighted sum as is the case for the pulsed inputs. If the inputs $Y_k(t)$ are permitted to be arbitrary functions of time, then I_j can be included in the weighted sum over the Y values as a step function in time, step$(t - t_0)$.

The linking modulation consists of adding a constant positive bias to the linking input and multiplying this by the feeding input. The bias is taken to be unity. The product is the total internal activity U_j of

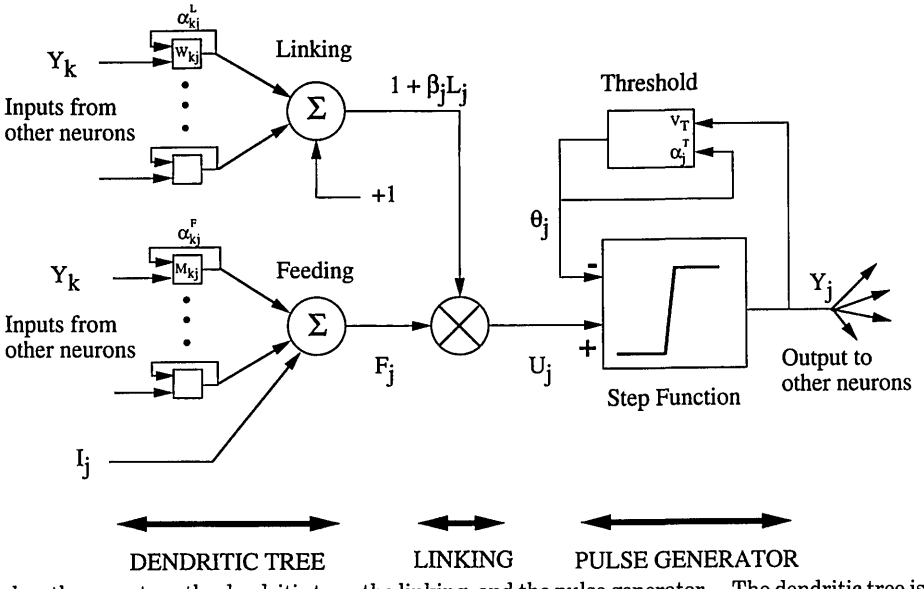


Fig. 1. Model neuron has three parts: the dendritic tree, the linking, and the pulse generator. The dendritic tree is subdivided into two channels, linking and feeding. All synapses are leaky integrator connections. The inputs are pulses from other neurons, and the output is a pulse. The linking input modulates the feeding input. When a pulse occurs in the linking input, it briefly raises the total internal activity U_j and can cause the model neuron to fire at that time, thus synchronizing it with the neuron transmitting the linking pulse.

the neuron:

$$U_j = F_j(1 + \beta_j L_j), \tag{3}$$

where β_j is the linking strength. U_j is a function of time. Under the above assumption that the feeding input has a smaller time constant than that of the linking input, the general behavior of U_j is that the linking inputs appear as spikelike modulations riding on a quasi-constant carrier formed by the feeding input. The internal activity is thus briefly raised to more than the feeding input level whenever a linking input occurs (Fig. 2). This effect is responsible for the synchronous activity found in the network as a whole. Equation (3) also establishes a correspondence between the linking-field model and higher-order networks. If Eqs. (1) and (2) are inserted into Eq. (3) there will be product terms of the form $M_{kj}W_{lj}Y_kY_l$ within a double sum. This is a second-order network.[10] This implies that if pulses rather than average firing rates are used in second-order nets, time-synchronous behavior should be observed. The research on adaptive second-order nets[11] may be applicable to the Eckhorn linking-field model as well.

The pulse generator uses a threshold discriminator followed by a pulse former, and a variable threshold that is dependent on the prior pulse output of the generator itself. When the neuron emits a pulse, it feeds back to the threshold, which is yet another leaky integrator, as shown in Fig. 1. One or more output pulses recharge the threshold to a high level. This quickly raises it to more than the current value of the internal activity U_j, which in turn causes the threshold discriminator to turn off the pulse former, and the neuron stops emitting pulses. The recharged threshold then decays exponentially according to its time constant and amplitude-gain factor until it drops to less than the internal activity again, triggering a new output pulse or a pulse burst from the neuron (Fig. 2). This is the pulse-generator model illustrated in Refs. 1 and 6–8 and is termed the exact model here. One important result of the exact model is that the pulse former produces a train of uniformly spaced pulses. Appendix 1 describes one way to do this. The spacing represents the refractory period, τ_r, of the neuron within which time a new pulse cannot occur. This gives an upper saturation limit to the maximum output pulse frequency. In the case of single output pulses, a simpler model can be used in which the threshold discriminator and the pulse former are replaced by a step function. This model[12] is used in this paper. Equations (4) and (5), given below, describe the threshold and the step function used in this model. Causality must be preserved in this approximation and is indicated in Eq. (4) by the use of t_{ret} to show that the step function output is evaluated at a slightly retarded time. The amount of retardation is bounded by τ_r, and is shown in Eq. (4) only to emphasize that the recharging of the threshold θ_j is causally driven by the step function output Y_j [the exact model does not require a retarded time (see Appendix 1)]:

$$\theta_j = [V_T \exp(-\alpha_j^T t)] * Y_j(t_{\text{ret}}) + \theta_0, \tag{4}$$

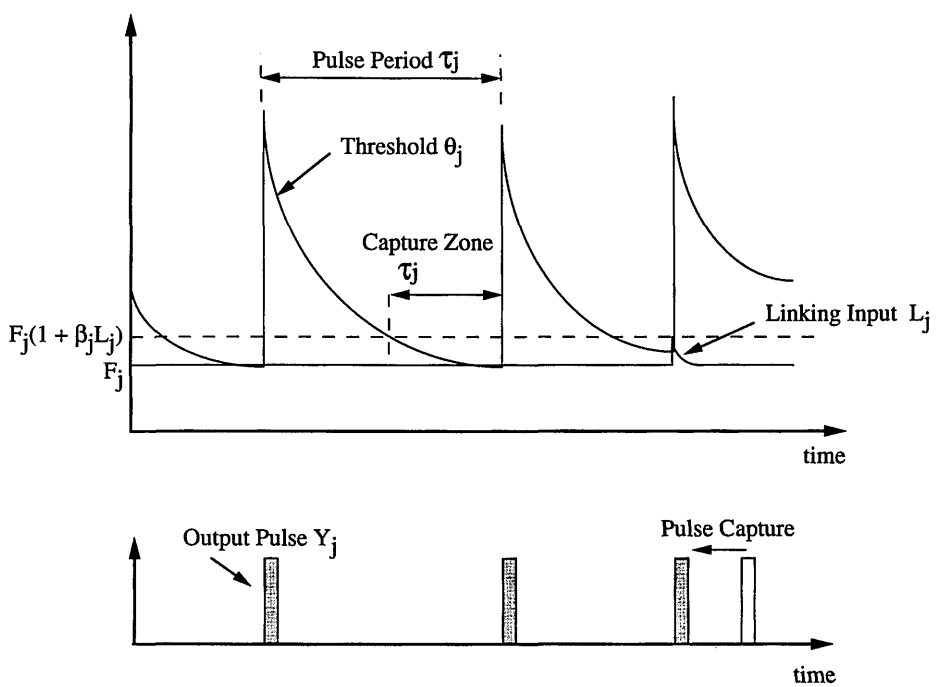


Fig. 2. Pulse generation and linking: The threshold is recharged when it decays to less than the internal activity $U_j = F_j(1 + \beta_j L_j)$. The output pulse is formed as the threshold turns the step function of Eq. (5) on and then off as the threshold goes below U_j, starts recharging, and then rises above U_j. If a linking pulse occurs in the capture-zone time, it causes the threshold to recharge sooner than otherwise, and the neuron fires a pulse synchronized with the arrival of the linking pulse.

$$Y_j(t) = \text{step}[U_j(t) - \theta_j(t)], \tag{5}$$

where V_T and α_j^T are the amplitude gain and the time constant of the leaky integrator threshold, θ_0 is a threshold offset, and step denotes a step function. The threshold time constants used by Eckhorn are intermediate in value between the linking and the feeding time constants. The simple pulse generator corresponds to a two-cell oscillator[13,14] in which the threshold acts as an inhibitor cell with a slow response and the step function acts as an excitatory cell with a fast response. The three parts of the model, the dendritic tree, the linking, and the pulse generator, act together to weight and to sum inputs in the receptive fields, to modulate one input channel with a second input channel, and to form the output pulses, which in turn are received by other neurons through their receptive fields. In the remainder of this paper the same threshold time constant is used for all neurons, the same linking time constant is used for all linking fields, the same feeding time constant is used for all feeding fields, and the same linking strength is used for all neurons, unless otherwise stated.

The firing rate of a single neuron is a sigmoidal function of the feeding input. This is shown by the pulse period τ_j of a neuron. It is the time required for the threshold to decay from its recharged initial height down to the internal activity level (Fig. 2). Consider Eq. (4), when the threshold is recharged with a single pulse by an amount V_T. For a constant feeding input F and a zero linking input the decay time back down to F is

$$\tau_j = \frac{1}{\alpha^T} \ln\left(1 + \frac{V_T}{F_j - \theta_0}\right). \tag{6}$$

The refractory period is added[8] to the pulse period at this point in the simple model. In the exact model it is implicitly included. The pulse firing rate f_j is then

$$f_j = (\tau_j + \tau_r)^{-1}. \tag{7}$$

It is shown in Fig. 3. It is a sigmoidal function.[15] Its monotonically increasing behavior shows that the original input feeding distribution can be approximately recovered at any time by an average being taken over many pulse periods because the pulse frequency is faster for stronger (more intense) feeding inputs. The sigmoidal nonlinearity will cut off values less than θ_0 and act as a squashing function near saturation; thus the overall function is a sigmoidal mapping of the internal activity to the output when pulse averaging is done.

When linking pulses are present, their strength defines a capture zone in the neuron receiving the linking pulse. From Fig. 2 the capture-zone time interval is, for $\theta_0 = 0$,

$$\tau_c = \frac{1}{\alpha^T} \ln(1 + \beta L), \tag{8}$$

where β is the linking strength. If a linking pulse is

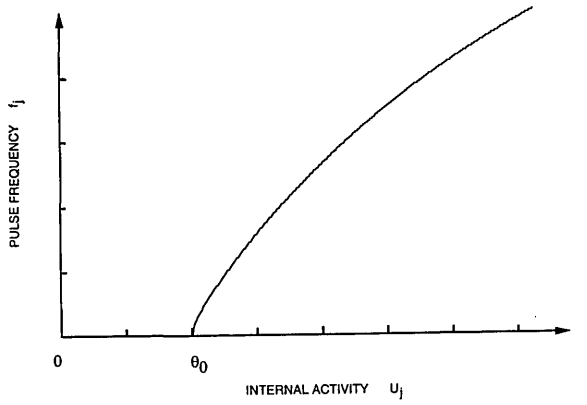


Fig. 3. Pulse frequency f_j as a function of the internal activity U_j: The pulse frequency is a sigmoidal function of the internal activity. Addition of a refractory time period τ_r makes the frequency saturate at the refractory frequency. A bias offset θ_0 shifts the curve's origin to that bias point.

received in this interval, it will briefly raise the internal activity level and cause the receiving neuron to fire at the arrival time of the linking pulse (Fig. 2). The receiving neuron will frequency lock to the transmitting neuron if their pulse rates τ_1 and τ_2 are similar. If the neurons have the same frequency ($\tau_1 = \tau_2$), they will phase lock when their phase difference ϕ is within the capture-zone time period:

$$|\tau_2 - \tau_1| < \tau_c, \tag{9a}$$

$$|\phi| < \alpha^T \tau_c, \tag{9b}$$

for the frequency lock and the phase lock, respectively. The pulse-coupled linking-field model contains synaptic weights [Eqs. (1) and (2)] but does not require a specific learning law. Any learning law can be used. Since the frequency function Eq. (7) gives the desired nonlinear response in the limit of averaging over many pulses, this model reduces to the usual nonpulsed networks in that limit. It has the weighted interconnects, the internal sums, and the sigmoid nonlinearity.

3. Time Series

Consider a group of weakly linked neurons. Suppose at time zero that all the neurons fire together. As time goes by, they will occasionally link in different combinations, as illustrated in Fig. 4. Each neuron has its own basic firing rate owing to its particular feeding input. Suppose further that at time T the neurons' combined firing rates and linking interactions cause them to again all fire together a second time. This duplicates the initial state at time zero. Then everything that happened during time T will happen again in the same order, and all the neurons will fire together again at time $2T$. This will continue, resulting in periodic behavior of the group with period T. The assumption of a single repetition of a given state (all the neurons firing together, for example) leads to the conclusion that everything that happened between the repetitions must necessarily

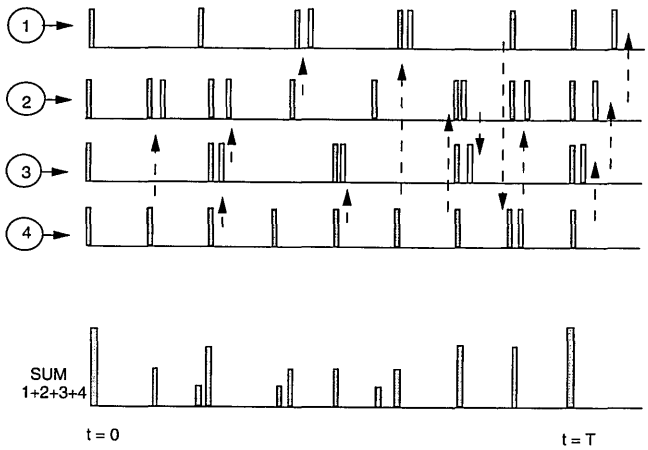


Fig. 4. Formation of a periodic time series: Neurons 1–4 all fire together at $t = 0$. As time passes, they occasionally link in various combinations. If at time T they link again so as to fire together, the situation will be the same as at $t = 0$. The system will repeat its behavior, generating a time series. The linear sum of the group's outputs is the periodic time signature of the input distribution to neurons 1–4.

happen again in the same order, in a permanently periodic way, for every neuron in that group. If all the outputs of the group are summed linearly, the result will be a single periodic time series that is the signature of that spatial input distribution. This is the time series $S(t)$. The length of time required for periodicity is governed primarily by the ratio τ_c/τ_{typ}, where τ_{typ} is the characteristic pulse period of the input image. For large β the ratio can be much greater than one, in which case the group links on every pulse and is completely periodic. Two other factors that promote periodicity in a two-neuron system are linking in quasi-harmonic ratios and linking on the decay tail of the linking pulses. For quasi-harmonic pulse rates such that

$$|m\tau_2 - n\tau_1| < \tau_c, \tag{10}$$

where (m, n) are integers, the two neurons will periodically link roughly every $m\tau_2$ seconds. When two mutually linked neurons link on the decay tails of the linking pulses, the cycle-to-cycle behavior is that they actively expel each other from this region into the leading-edge linking region. While both effects promote periodicity, they do not guarantee it. The time required to achieve periodicity and the overall period length can be large for large, weakly linked slabs.

The following interpretation of the time series relates it to the input image's intensity histogram. The network maps intensity to frequency. The size of an isointensity image patch determines how many neurons fire at that corresponding frequency; thus patch size maps to amplitude. The image's intensity histogram gives the number of pixels with a given intensity, while the amplitude of a given frequency gives the number of neurons firing at that rate. The frequency spectrum of the time signal is the intensity histogram of the input image as mapped through the sigmoidal response. This is true for discrete pulse models and for continuous oscillator models and also for any other model in which the output frequency is porportional to the input signal strength.

For a linked slab the coherent periodicity of the time signal suggests that there must exist phase factors as well as frequency and amplitude. Suppose that the time signal $S(t)$ is expressed as a sum of delta function pulses:

$$S(t) = \sum_n \sum_{k=1}^{K} a_k \delta(t - nT - \varphi_k), \tag{11}$$

where T is the periodicity, a_k is the amplitude of the kth subgroup, and φ_k is the time offset of the subgroup of cells with amplitude a_k. The time offset is between zero and T, and there are K subgroups that are linked into the overall repetition period T. If a Fourier transform (FT) is taken, it factors into a sum of complex phases and a sum representing the repetition period:

$$\mathrm{FT}(S) = \left[\sum_{k=1}^{K} a_k \exp(i\omega\varphi_k)\right]\left[\sum_n \exp(i\omega nT)\right]. \tag{12}$$

The corresponding histogram must in some form include the phases as well as the amplitudes. This is an open problem, as is the derivation and the evolution in time of a periodic time signal for a given feeding-input distribution. Other transforms may be more appropriate; the Fourier transform was used here for illustrative purposes. This discussion shows that the geometrical content of an image as well as its intensity is encoded in the time signal, and it shows that distance-dependent linking action provides a way to include syntactical information. The time signals are object specific. They are a signature, or an identification code, that represents a two-dimensional image as a time-domain signal that can be processed by the neural net. The signatures have some degeneracy, but as shown in Section 4, this can be an advantage rather than a drawback, since certain classes of degeneracy can also be viewed as invariance.

4. Invariances

If there are symmetries in a receptive-field weight pattern such that any two parts of it are the same, then an exchange of the corresponding parts of an input-image distribution will not change the weighted sum of the field. The neuron's output is determined by its internal activity U_j, which is a function of the feeding and the linking inputs. If the image changes in a way that fits the symmetries of the receptive fields so that the internal activity does not change, then the neuronal output will be invariant against these changes. The symmetries of the receptive fields correspond to invariances of the time signal generated by the input image. Section 3 showed that the pulse-coupled network produces time series that encode in their phase structure the two-

dimensional spatial input distribution, including its geometrical connectivity relationships. The time series were observed both in simulations and indirectly in a hybrid optical implementation to be generally periodic, and some mechanisms were identified in Section 3 that promote periodicity. Symmetries can be introduced in the receptive fields to make the time signature of an image invariant against translation, rotation, and scale. Simulation results also show that there is a significant invariance against scene illumination and distortion; further, there is some limited invariance against changes in the overlying patterns (shadows) on a given image shape.

The design objective is to make the internal activity invariant by introduction of geometrical symmetries into the receptive-field weight pattern: (1) For translational invariance, use the same receptive field weight pattern at every neuron. (2) For rotational invariance, make the receptive field patterns circularly symmetric. A given image patch of a translated and rotated image then sees the same receptive-field patterns as before but on a different set of neurons. The time signal is a sum over all the neurons; thus it does not matter which neurons are used, but rather how they are used. (3) For scale invariance, use an inverse-square radial falloff. This does not make the internal activity invariant against distances r, but rather against scale changes as represented by the factor k in the rescaled distance kr. Consider an optical image that is rescaled by a change in the object distance (Fig. 5). In this case the intensity per image patch is constant. The number of cells affected by the rescaled patch is changed, but not the output pulse frequency. A neuron receiving the input at the recaled location of the original image patch is driven by the same intensity as the neuron at the original location. For a rescaling factor of k,

$$Y(k\mathbf{R}) = Y(\mathbf{R}).$$

The linking input to this neuron, obtained with an inverse-square kernel, is

$$L(k\mathbf{R}) = \int_0^{2\pi}\int_{\rho_0}^{\infty} \frac{1}{(k\rho)^2} Y[k(\mathbf{R}+\boldsymbol{\rho})]k\rho k\mathrm{d}\rho\mathrm{d}\theta$$

$$= \int_0^{2\pi}\int_{\rho_0}^{\infty} \frac{1}{\rho^2} Y(\mathbf{R}+\boldsymbol{\rho})\rho\mathrm{d}\rho\mathrm{d}\theta + L(\mathbf{R}). \tag{13}$$

This removes the scale-factor dependence k from the integrand. The lower integration limit ρ_0 is fixed and does not scale, so the above relation is not an exact equality in some cases. This is discussed below.

If the feeding field is a single pixel (this is not essential and is done here for simplicity only), then

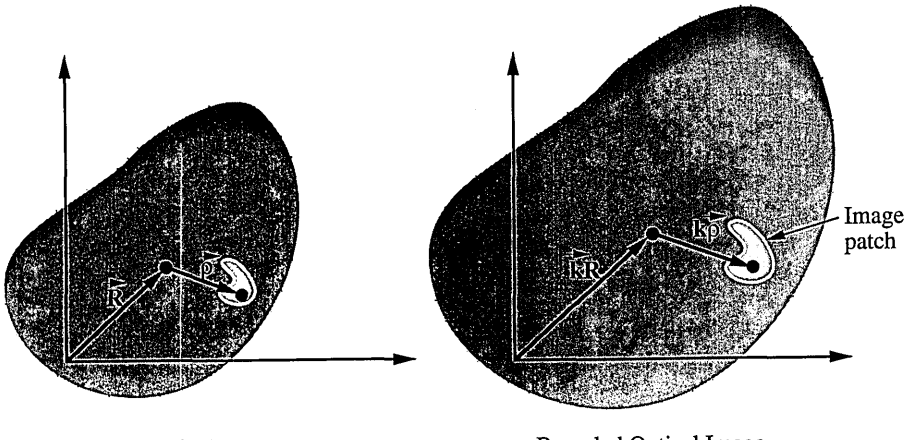
$$F(k\mathbf{R}) = f(\mathbf{R}).$$

The internal activity of the rescaled image is thus the same as that for the unscaled image:

$$U(k\mathbf{R}) = F(k\mathbf{R})[1 + \beta L(\mathrm{k}\mathbf{R})]$$

$$= F(\mathbf{R})[1 + \beta L(\mathbf{R})] = U(\mathbf{R}). \tag{14}$$

There is a problem that must be resolved before complete scale invariance is achieved. It appears to be less important for large images on a fine grid of cells, but when the isointensity patch sizes cover less than approximately 10 × 10 cells, it has some effect. The problem is that the local group around a neuron also changes in scale. The linking input that is due to the local group varies accordingly with scale, making the internal activity change as well. The cause is the fixed inner edge ρ_0 of the linking field. It does not scale. External groups do not have this property because all their boundaries shift accordingly as the image scale is changed. For simplicity, consider a neuron at the center of its local patch,



Original Optical Image Rescaled Optical Image

Fig. 5. Geometry for scale invariance: A neuron at $\mathbf{R}$ receives a linking contribution from a neuron at $\boldsymbol{\rho}$. When the image is rescaled, the image patch at $\mathbf{R}$ goes to $k\mathbf{R}$ and the patch at $\boldsymbol{\rho}$ goes to $k\boldsymbol{\rho}$. Only the latter patch is shown. For the case of an optical image rescaled by a change in the object distance the intensity per image patch is constant. The objective is to design a linking receptive field such that $L(k\mathbf{R}) = L(\mathbf{R})$.

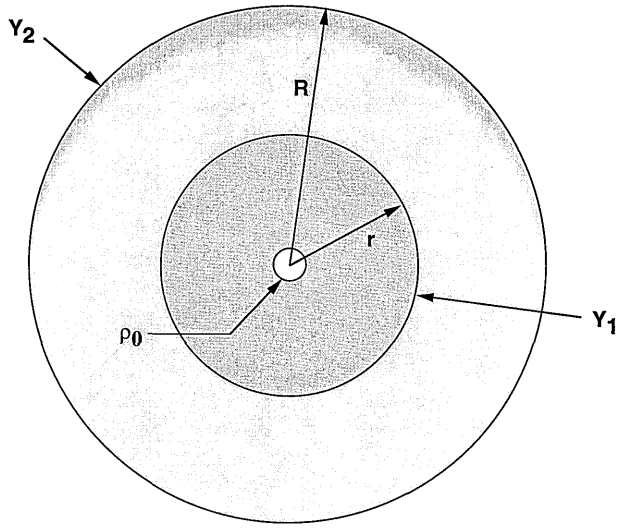


Fig. 6. Geometry used to show that the fixed inner radius ρ_0 of the local group L_1 causes a dependency on the rescaling factor k. The external group L_2 is in the annulus from r to R, while L_1 extends from ρ_0 to r.

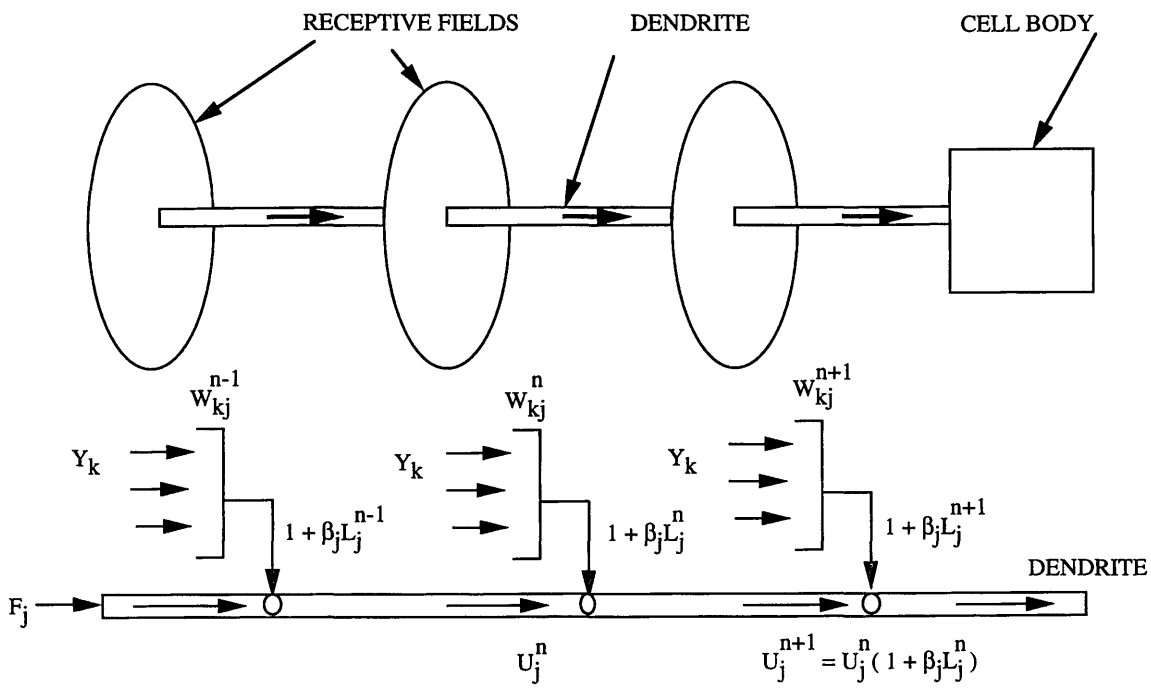
which is surrounded by an external patch, making two concentric circles, as shown in Fig. 6. Let ρ_0 be the fixed inner edge of the local patch and Y_1 and Y_2 be the pulse activities in the local and the external patches, respectively. Then

$$L = 2\pi \int_{\rho_0}^{r} \frac{Y_1}{\rho^2}\,\rho d\rho + 2\pi \int_{r}^{R} \frac{Y_2}{\rho^2}\,\rho d\rho$$

$$= 2\pi Y_1 \ln \frac{r}{\rho_0} + 2\pi Y_2 \ln \frac{R}{r}. \qquad (15)$$

Under a scale change, r and R become kr and kR, but ρ_0 is fixed. The linking input to the center neuron then has a scale-factor dependence porportional to $Y_1 \ln(k)$.

The solution is to make the internal activity distinguish between the local and the external group and also to make both scale invariant. The local group can be made independent of scale by use of a nearest-neighbor receptive field with a fixed outer limit. To distinguish between local and external groups, however, one needs to use the generalized linking-field model with multiple linking fields as well as excitatory and inhibitory dendritic inputs. While Eckhorn's diagrams show only one linking input per dendrite, multiple linking inputs are clearly implied in his discussions.[1,6,7] A dendrite can receive many linking inputs along its length, as shown in Fig. 7. Each input modulates the signal activity flowing toward the cell body, and the total dendritic signal



EACH DENDRITE IS A LOGICAL RULE

RF weighted sums ⟷ "OR" Linking product ⟷ "AND"

Fig. 7. Linking field model neuron as a multirule logical system: A dendrite receives inputs from many receptive fields along its length. Each input modulates the dendritic signal by the factor $(1 + \beta_j L_j^n)$ for the nth input. The receptive fields can give the same signal for more than one input distribution and thus correspond to a logical OR. The product term in the modulation factors corresponds to a logical AND. These logic-gate correspondences are not exact but can be used effectively, as shown by the example discussed in Section 4.

into the cell is

$$U_{\text{tot}} = F(1 + \beta_1 L_1)(1 + \beta_2 L_2) \cdots (1 + \beta_n L_n). \quad (16)$$

This can be viewed as a general higher-order net of order n owing to the various linking product terms. The dendritic signals are summed in the cell body and can be either excitatory or inhibitory. The weighted sums in the receptive fields correspond to fuzzy OR gates, while the products from the linking modulation correspond to fuzzy AND gates. A single neuron is a large logical multirule system in this view. This is used to construct a semiexclusive OR to let the neuron distinguish between the local and the external linking input. Use two dendrites, each having two linking inputs. One dendrite is excitatory, the other is inhibitory. The same linking inputs L_1 and L_2 are used on both, and both are fed by the same feeding input F, but the linking strength coefficients are all different:

$$U_{\text{exc}} = +a_1 F(1 + \beta_1 L_1)(1 + \beta_2 L_2), \quad (17a)$$

$$U_{\text{inh}} = -a_2 F(1 + \beta_3 L_1)(1 + \beta_4 L_2), \quad (17b)$$

$$U_{\text{tot}} = U_{\text{exc}} + U_{\text{inh}}. \quad (17c)$$

Choose the a and the β values so that they are all positive and so that the total internal activity has the form

$$U_{\text{tot}} = F(1 + \beta L_1 + \beta'\{1 - L_1/[L_{1(\max)}]\}L_2). \quad (18)$$

For the values $\beta = 0.2$, $\beta' = 0.3$, and $L_{1(\max)} = 40$ used in the simulations, one possible set of coefficients is $a_1 = 2$, $a_2 = 1$, $\beta_1 = 1$, $\beta_2 = 219/640$, $\beta_3 = 1.8$, and $\beta_4 = 123/320$. $L_{1(\max)}$ is the maximum possible value of the local-neighborhood linking input L_1, and L_2 is a linking input from a larger and more extended receptive field, such as the inverse-square field. L_1 gives the input from the local group, and L_2 gives the input from external groups that do not contain the neuron being linked. When the entire local group fires, $L_1 = L_{1(\max)}$ and the neuron sees only its nearest neighbors. When the local group is quiet, $L_1 = 0$, and the neuron can receive the L_2 linking from the external groups. Suppose the rescaled image patch now makes several new adjacent groups out of the local group, all with the same frequency. If they are in phase, the neuron's local group will mask them. If they are not in phase, then they will link with the local group through the second linking input and they will be captured by the local group. Then they will be in phase, and the local group effectively enlarges to include them but without altering the internal activity seen by a given neuron. When the outer limit of L_1 is chosen to overlap the inner limit on L_2, the inner boundary of the external group is always the outer boundary of the composite local group. The system's architecture has translation, rotation, and scale invariance. It is a third-order network, which has been shown[16] to be the minimum order necessary for achieving these invariances all at the same time.

This model was simulated on a personal computer. The array size was 33 × 33, and the images were made of five blocks, each with its own intensity, and the blocks were rearranged to form the different test images. A cross shape and a T shape were used. They differed only in their geometrical arrangement. Each block contained from 5 to 11 cells on a side, depending on the scale factor, and the background was set to zero in all cases. No noise was added. Analysis of the grid size indicated that reasonable results could be expected down to a 5 × 5 block size for rotation, and the scale increments were chosen so that the blocks varied in size by 5, 7, 9, and 11 cells on a side. The nearest-neighbor linking field for L_1 was a 3 × 3 square (center excluded), while the outer radius of the inverse-square linking field for L_2 was fixed at 10 and the inner radius at 1. Figure 8 schematically represents the simulation model. The simulation's equations were written for discrete time steps with the digital filter form from Ref. 1. They are

$$F = \text{image}(j, k)/255, \quad (19a)$$

U_{tot} as given by Eq. (18),

$$L_{\text{loc}}(t + 1) = A_1 L_{\text{loc}}(t) + V_L L_1(t), \quad (19b)$$

$$L_{\text{ext}}(t + 1) = A_1 L_{\text{ext}}(t) + V_L L_2(t), \quad (19c)$$

$$\theta(t + 1) = A_2 \theta(t) + V_T Y(t), \quad (19d)$$

$$Y(t) = \text{step}[U_{\text{tot}}(t) - \theta(t)]. \quad (19e)$$

The parameter values were $A_1 = \exp(-1/t_1)$, $A_2 = \exp(-1/t_2)$, $t_1 = 1$, $t_2 = 5$, $V_L = 5$, $V_T = 20$, $\beta = 0.2$, $\beta' = 0.3$, and $L_{1(\max)} = 40$, and image(j, k) was the input image. The results are shown in Figs. 9–13. The most important result was that the time signatures were object specific. Each test image generated a distinct periodic time signal. This showed that the pulse-coupled net encoded the images in accordance with their geometrical configuration, because both images were built of the same five blocks arranged in different geometrical configurations. Good invariance was achieved for translation, rotation, and scale. The time signatures of the two test images were easily distinguished in all cases except for the smallest rescaled T shape (Fig. 10). Its patch size was 5 × 5. A grid coarseness analysis indicated that, at less than 7 × 7 size, the grid effects would be significant. The rotated T images, likewise, were sensitive to these effects, but their signatures were still distinct from those of the cross image (Fig. 9) for patch sizes greater than 5 × 5. The rotated T images were translated, as well, to fit in the small slab grid of 33 × 33 cells; thus Fig. 10 also indicates translational invariance.

The images were tested with different scene-illumination levels. It was found that their time signatures (Fig. 11) were invariant over a factor of 200 in illumination. This was not expected, as the ratio of the capture-zone time to the neuronal period

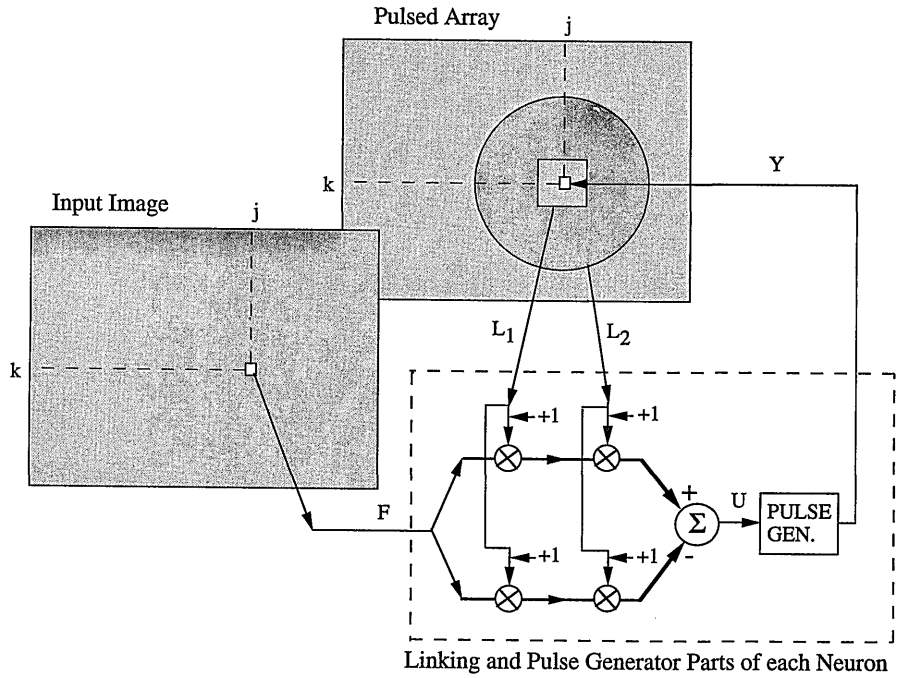


Fig. 8. Schematic of the model used in the simulations: An image pixel feeds the neuron. Two receptive-field inputs, a nearest-neighbor 3×3 field L_1 and an inverse-square field input L_2, modulate the feeding input on both an excitatory and an inhibitory dendrite. This system is a third-order neural network. The total internal activity is the sum of the two dendritic signals. The pulse generator sends the output pulse to the neuron located in the pulsed array at the same relative location as the feeding input pixel in the input image.

changes in this case. What happens is that the signature period varies, as expected, but the signature itself remains the same. Detailed examination of these runs, after the fact, gives a possible explanation: The signatures reflect the propagation of linking waves (see Section #5) through the scene object. These waves follow gradients, and changes in the overall scene illumination did not change the relative gradient patterns. There was less variation in the signatures caused by scene-illumination changes than from other image changes.

Figure 12 shows the effect of image distortion. A coordinate transform of the form ($x' = x + 0.01xy$, $y' = y + 0.01xy$) was used to approximate an out-of-plane rotation of $\sim 30°$ with some perspective added. The signatures retained their characteristic forms sufficiently for the cross and the T images to still be correctly classified by their signatures. Again, this suggests a close relationship between the image morphology and the time signature. The insensitivity to distortion is caused by the signature generation being more of an area effect rather than an edge or an angle effect.

Image-intensity overlays were investigated next. The 9×9 scale T image was altered by transposition of the two lower blocks. This would correspond to a shadow across the image, for example. The result, shown in Fig. 13, is not invariant, but shows a distinct correspondence of the new signature to the original. Figure 14 shows the effect of combined image changes. Translation, rotation, scale, and scene illumination and distortional changes were made, as indicated in the figure. The new signatures were similar enough to the originals for the altered images to be classified correctly as a cross or as T by use of only the signatures. They are clearly not strictly invariant, but show a substantial insensitivity to the geometrical changes while retaining their object-specific character.

There are distinct time-scale regimes in the pulse-coupled net. The size of the discrete time steps determined how quickly the linking interactions could propagate in the simulation. In a real system the linking inputs are transmitted much faster than the times required for the threshold to decay, and the pulse generator can respond quickly to the linking inputs. The linking can propagate as a wave (Section 5), with a high velocity. This defines two distinct time scales, one in which linking waves sweep quickly over uniform regions (uniform is dependent on the network parameters as well as the underlying image) and one that defines when the next such wave will start out over the region. If there is a sufficiently large and/or abrupt change in intensity between two regions, each region will have its own linking wave propagating through it at a repetition rate different from that of the other region's linking wave. This gives a way to segment pixels into large areas by grouping them according to the periods and the relative phases of their linking waves. This also discretizes the input intensity into gray levels. The above simulation results are in the fast time-scale (linking-wave) regime. The signatures reflect how the linking waves flowed through the images. If they are considered from the view of the slower time scale, then the signatures would be highly com-

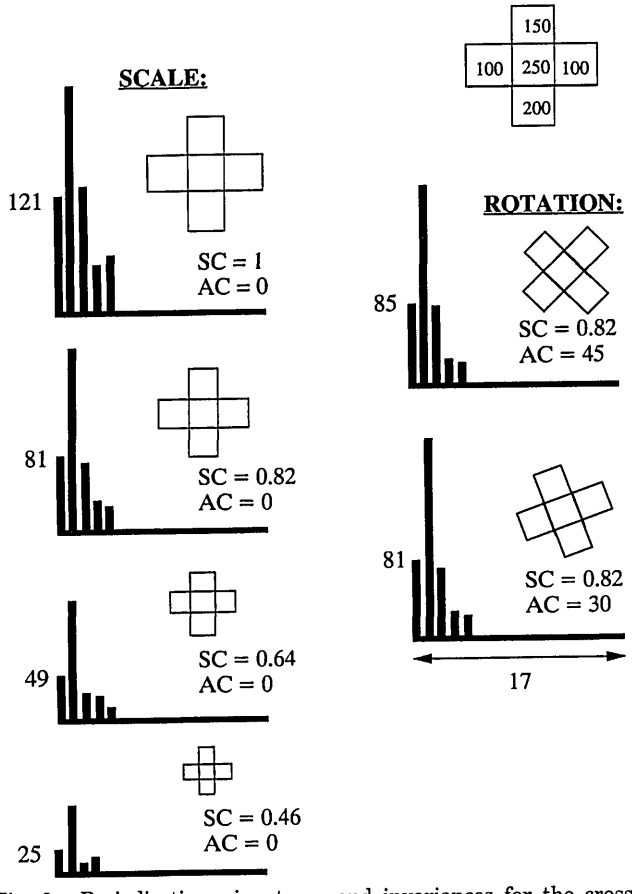


Fig. 9. Periodic time signatures and invariances for the cross image: The signatures are the periodic part of the total-output time signal of the pulsed array. SC is the scale factor, and AC is the rotation angle in degrees. Good scale invariance was found for scales over 1:0.46 and for large rotations of 30° and 45°. The five blocks arranged to form the image were scaled from 11 × 11, to 9 × 9, to 7 × 7, to 5 × 5 block sizes. The 33 × 33 slab had a background intensity level of zero. Grid coarseness effects were expected for 7 × 7 and smaller block sizes in scale and for 14 × 14 block sizes in rotation. Grid effects were not severe in this image.

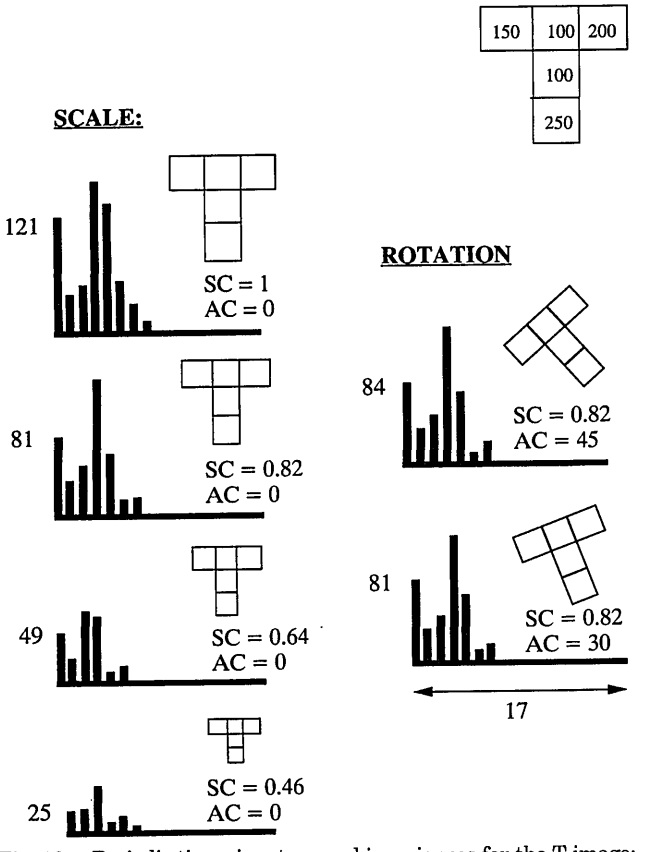


Fig. 10. Periodic time signature and invariances for the T image: The same setup as for Fig. 11 is used, but with the five blocks rearranged to form a T. The signature was very distinct as compared with the first case, showing that the net makes unique time signatures for different images even when they are rearrangements of the same components. The scale invariance was good down to the 7 × 7 block size. The rotated images' signatures still followed the overall T signature shape in contrast to the cross signature. Their variation from ideal is strictly caused by grid effects.

pressed in time, and each signature period would represent the segmentation of that image region in time. If the time scale is further enlarged, one has large segmented groups pulsating synchronously. The network can link these groups and define yet another still-larger time scale using groups of cells rather than individual cells. This leads in principle to groups of groups, large-scale group linking waves, signatures for groups, and segmentation of groups of groups. However, in these regimes there are large bursts of output pulses and large composite linking pulses, and so the exact model neuron should be used rather than the simple step-function approximation (Section 2).

5. Optical Implementation

The network was implemented[5] as a hybrid optical system with a framegrabber, a computer, and a defocused spatial light modulator (SLM). The step function and the threshold decay were performed in the computer, while the SLM formed the linking field and the modulation product. The system used real images, and each pixel was the single feeding input to each corresponding neuron. The amount of defocusing defined the extent of the linking receptive field, and was ~10 pixels in diameter. The output was a 512 × 512 image of all the pulses occurring in each processing cycle. Its defocused image on the SLM made the linking input. The light passing through the SLM multiplied the input image by the defocused SLM image to produce the linking modulation. The cycle time was ~10 s. This gave time to observe in detail all the changes from one cycle to the next and to follow the linking interactions.

The degree of linking was in the weak-to-moderate regime, as evidenced by the lack of multiple bursts of pulses in the output images. While the system was used primarily in the weak linking regime, it was possible to obtain multiple pulses on very bright input images (Fig. 15). This image was segmented on the first pulse cycle [Fig. 15(b)], and as time passed, the output pulses showed a tendency to outline the segments as well [Fig. 15(c)]. Even with linking, a

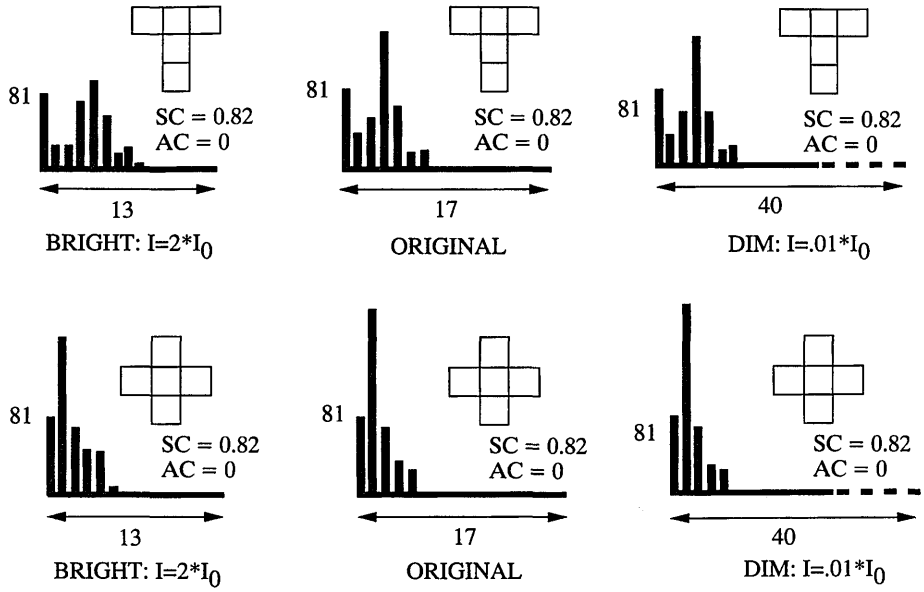


Fig. 11. Intensity invariance: The 9 × 9 block-size images were multiplied by an intensity factor I_0 corresponding to a change in scene illumination. From $I_0 = 2$ to $I_0 = 0.01$ the signature was invariant in its shape, though the period of the signature varied from 13 to 40 time units.

time average of the outputs [Fig. 15(d)] gave a good reconstruction of the original input. This is significant because the net is trying to obey two distinct requirements at the same time: it must fire at a frequency given by Eq. (7), and it must link according to Eq. (3). Figure 15 demonstrates that it meets both requirements: the time average shows that the cells are firing at a frequency porportional to the input intensity, while the observed segmentation and edge enhancement are properties of the linking activity.

Figure 16 shows the same net operating on a dim, low-contrast image. Here the major effects were edge and contrast enhancement. Since even small intensity differences map to different pulse frequencies, there will eventually be an output pulse image in which one region is bright and an adjacent region is black, giving a momentary contrast enhancement. The edge enhancement is due to the linking waves accumulating on a steep edge gradient. Figure 3 from Ref. 5 shows the output of the same net for an image with moderate gradients. The image was a light square (lower right) and a light spot (lower left) on a light background. The dominant behavior was

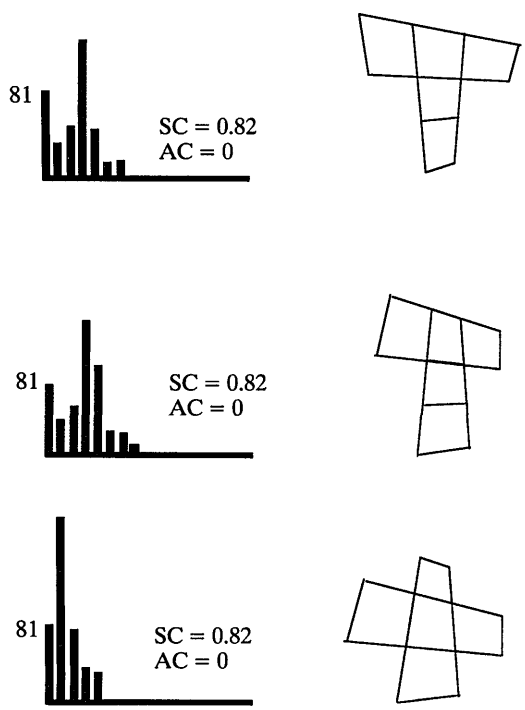


Fig. 12. Image distortion: A coordinate transform approximating a 30° out-of-plane rotation was used for both test images. Their signatures were still distinct and recognizable as belonging to the correct image classification.

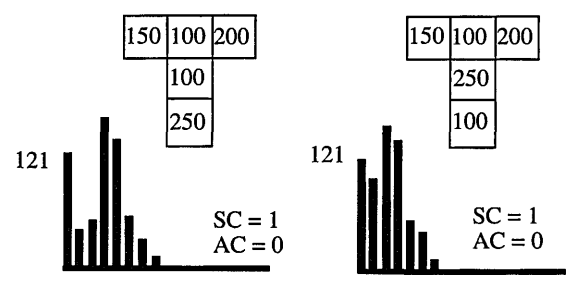


Fig. 13. Signature of the T image with two blocks interchanged: The two lower blocks of the full-scale unrotated T image were interchanged, simulating the effect of a shadow moving down the image. The new signature is similar to that of the 7 × 7 block-size T image and still has an initial peak followed by a valley and then a higher peak. In contrast the cross image's second peak was lower than its first peak, so this signature would still be classified as a T and not as a cross.

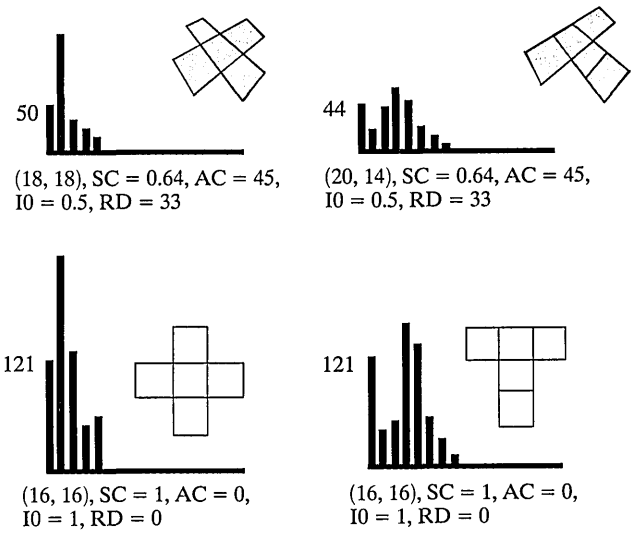


Fig. 14. Effect of combined image changes: The original images were located at coordinates (16, 16) with scale factors sc = 1, they were unrotated, and they had no distortion (RD is the approximate out-of-plane rotation). The signatures were sufficiently insensitive to the combined changes for the images to still be correctly classified.

dynamic periodic traveling waves. As with all the test images, when the linking was turned off, the result was chaotic in appearance because the relative phases increased linearly in time. With linking, waves were formed. Inspection of Fig. 3 in Ref. 5 shows that the wave patterns are locally periodic. When two waves collided, they destroyed each other because immediately behind the wave fronts the thresholds were reset to a high level. The waves have characteristics found in scroll waves[17] and autowaves.[18]

Figure 17(a) shows a horizontal gradient against a dark background. The waves from the background initially move downward across the gradient strip [Fig. 17(b)]. As time passed, the gradient's waves established themselves [Fig. 17(c)] until, after 300 time cycles, they had removed the initial waves caused by the background from most of the gradient strip [Fig. 17(h)]. The waves on the strip moved from right to left. It was observed that, on the strip's top boundary, its gradient waves set up new waves propagating upward into the background that

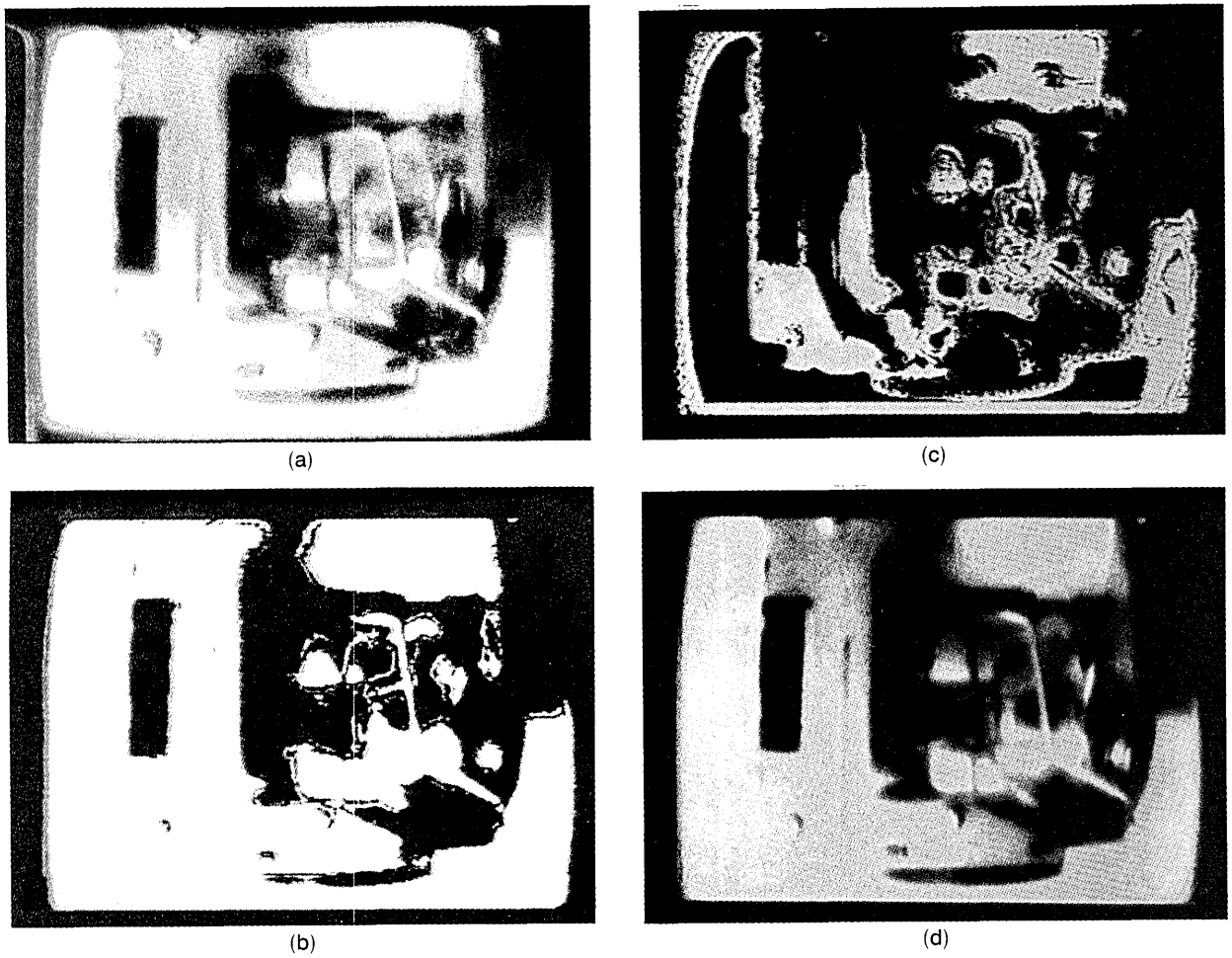


Fig. 15. Optical pulse images showing segmentation and linking: This object was chosen at random in the laboratory. The pulse-coupled optical network was operated at high illumination levels so as to produce multiple pulses in the bright areas of the image. (a) The input image, (b) the pulse output image after one cycle, and (c) the output after several cycles. Here the tendency of the network to outline object regions is apparent. (d) The sum of 30 pulse images, all similar to (b) or (c). The original image is recovered. Its intensity is remapped according to the sigmoid function. There was some noise reduction owing to the averaging inherent in the summation.

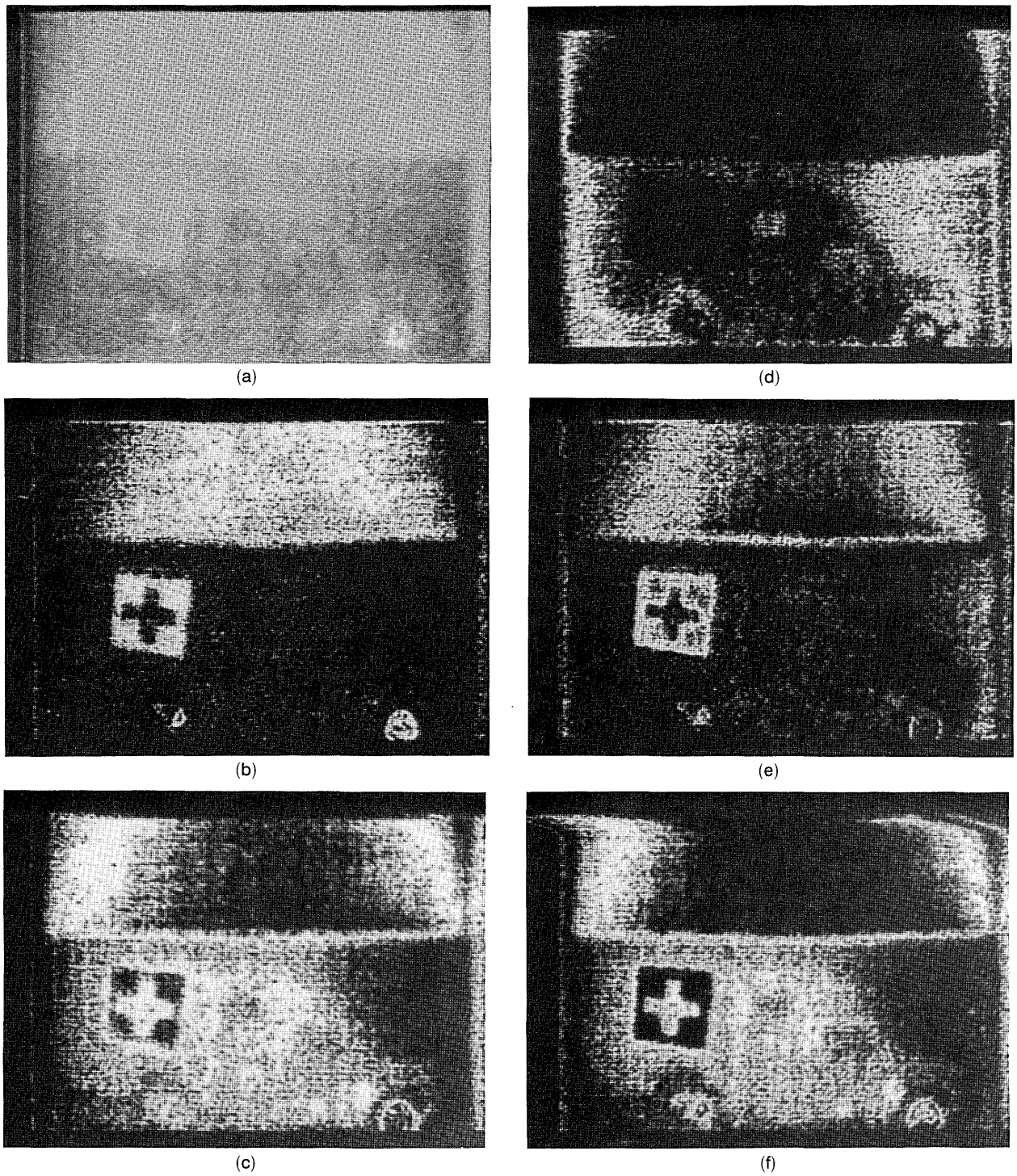


Fig. 16. Optical pulse images of a model truck: A low-intensity low-contrast image was both edge and contrast enhanced by the optical network. Even small intensity differences caused adjacent regions to pulse at different times, giving a momentary contrast enhancement. The network also enhanced edges because the linking waves tended to accumulate in regions of larger gradients.

interacted with the downflowing waves in that region. Within the gradient strip, disturbances were occasionally generated and were swept along the strip in the direction of its waves [Fig. 17(d)–17(g)]. It moved at a slower rate than the waves. This shows that the waves can carry information with them beyond the limits of the receptive fields. The receptive field size in this case was ~10 pixels in diameter, and the disturbances were carried ~50 pixels before dissipating.

In general, two types of waves were observed. The first type was a gradient-following wave. Their apparent velocity, wavelength, and duty cycle varied with the threshold decay constant, the linking

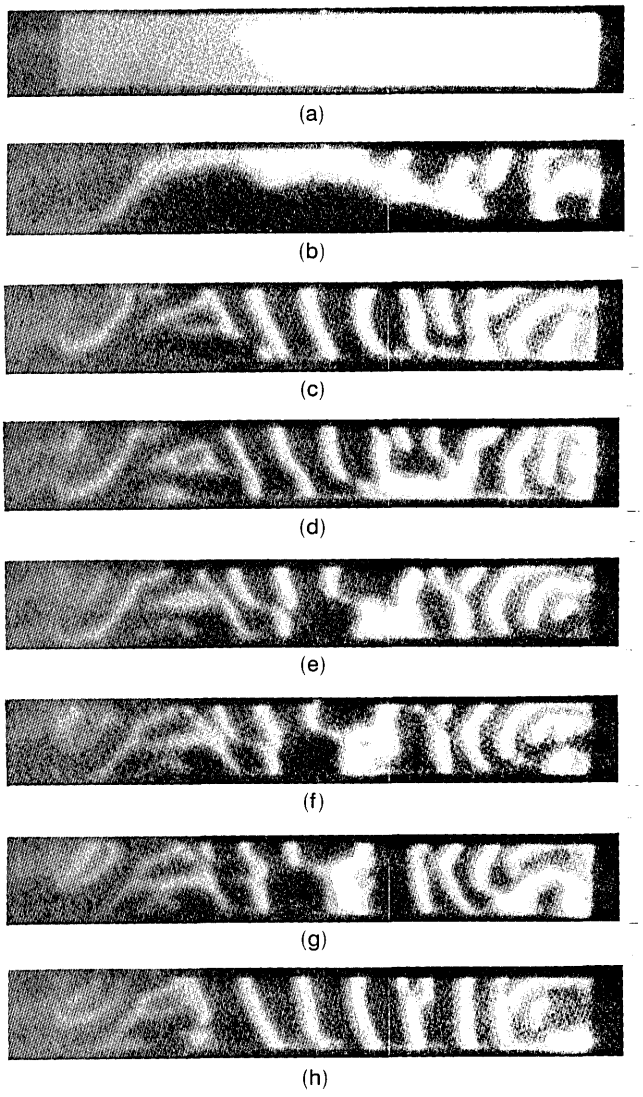


Fig. 17. Horizontal gradient-wave dynamics: These display the internal activity rather than the pulse images, showing what the neurons actually receive at the pulse generator. (a) The horizontal gradient strip against a dark background. (b) After 20 cycles, the gradient waves are beginning to push the initial pulse pattern to the left as they establish themselves. (c) The established pattern. (d), (e), (f), and (g) The propagation of a disturbance. It originated from a boundary interaction and was carried by the gradient waves from its point of origin to the left. It moved more slowly than the carrier waves but was taken over distances beyond the linking receptive field's interconnection distance. This shows that the waves can carry information with them. (h) The gradient waves after the disturbance was pushed off the left end of the strip. The system was run for 300 time cycles.

strength, and the underlying image gradient. The second type of wave could propagate without a gradient or even against a gradient. They originated at edges or at bright spots and moved out in patterns that were visually similar to water waves moving out from a source. Waves incident upon a border did not reflect, but waves passing from one intensity region to another intensity region with a different gradient would alter their direction of travel accordingly, giving an appearance similar to refraction.

6. Summary

It has been shown that the linking-field model is a higher-order network. It produces a periodic time series for two-dimensional images in the limit of weak linking strength. The time series are object specific and serve as signatures for scene objects. The signatures encode in their phase structure some of the geometrical content of the images. They are related to the histogram of the image and to the morphological structure of the image. By choice of receptive-field weight patterns with certain symmetries, the signatures become insensitive to translation, rotation, scale, intensity, and distortion changes in the input images, while retaining their object-specific character. An optical hybrid laboratory demonstration system showed that locally periodic traveling waves were produced. These waves carried waveform disturbances in the direction of propagation over distances beyond the interconnect distance of the receptive fields.

Appendix 1

The threshold discriminator and pulse former can be modeled by a train of pulses modulated by an envelope[19,20]:

$$Y_j(t) = \left(\frac{1}{1 + \exp\{-A[U_j(t) - \theta_j(t)]\}}\right)P(t), \quad \text{(A1)}$$

$$P(t) = \sum_n P_0(t - n\tau_r), \quad \text{(A2)}$$

$$P_0(t) = \frac{1}{\tau_w}\int_{-\infty}^{t} [\delta(t') - \delta(t - \tau_w)dt']. \quad \text{(A3)}$$

$\delta(t)$ is a delta function, τ_w is the pulse width, τ_r is the pulse repetition rate, and A is the scale factor of the argument of the hyperbolic tangent function used as the envelope of the pulse train $P(t)$. The constraints are that $1/\alpha_T \gg \tau_r \gg \tau_w$ and that the hyperbolic tangent go from low to high in a time much shorter than τ_r but much longer than τ_w. This lets one pulse quickly recharge the threshold. It can be put in terms of the scale factor A by the requirement that the full width at half-maximum (FWHM) of the derivative of the hyperbolic tangent, expressed in terms of time as the threshold passes through $U(t)$ at $t = t_0$, obey the condition

$$1/\alpha_T \gg \tau_r \gg \tau_{\mathrm{FWHM}} \gg \tau_w, \quad \text{(A4)}$$

where τ_{FWHM} is approximately

$$\tau_{\mathrm{FWHM}} = \mathrm{FWHM}/[U(t_0)\alpha_T]. \quad \text{(A5)}$$

This is obtained by expansion of the decaying threshold around $t = t_0$. The FWHM of the derivative, a hyperbolic cosecant squared, is easily found to be

$$\mathrm{FWHM} = \frac{2}{A}\ln\left(\frac{1}{3 - 2\sqrt{2}}\right). \quad \text{(A6)}$$

The author acknowledges the many improvements and clarifications made to this paper as a result of the valuable criticisms given by the reviewers. It was a difficult paper to review. They are to be commended, and I do so here.

References and Notes

1. R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "Feature linking via synchronization among distributed assemblies: simulations of results from cat cortex," Neural Comput. **2,** 293–307 (1990).
2. R. Eckhorn, R. Bauer, M. Rosch, W. Jordan, W. Kruse, and M. Munk, "Functionally related modules of cat visual cortex shows stimulus-evoked coherent oscillations: a multiple electrode study," Invest. Ophthalmol. Vis. Sci. **29,** 331–343 (1988).
3. R. Eckhorn, "Stimulus-evoked synchronizations in the visual cortex: linking of local features into global figures?" in *Neural Cooperativity,* J. Kruger, ed., Vol. 2 of Springer Series in Brain Dynamics (Springer-Verlag, New York, 1989), pp. 267–278.
4. J. L. Johnson, "Waves in pulse-coupled neural networks," in *Proceedings of the World Congress on Neural Networks* International Neural Network Society, Hillsdale, N.J., 1993), Vol. 4, pp. 299–302.
5. J. L. Johnson and D. Ritter, "Observation of periodic waves in a pulse-coupled neural network," Opt. Lett. **18,** 1253–1255 (1993).
6. R. Eckhorn, H. J. Reitboeck, M. Arndt, and P. Dicke, "A neural network for feature linking via synchronous activity: results from cat visual cortex and from simulations," in *Models of Brain Function,* R. M. J. Cotterill, ed. (Cambridge U. Press, Cambridge, 1989), pp. 255–272.
7. R. Eckhorn and T. Schanze, "Possible neural mechanisms of feature linking in the visual system: stimulus-locked and stimulus-induced synchronizations," in *Self-Organization, Emerging Properties and Learning,* A. Babloyantz, ed. (Plenum, New York, to be published).
8. P. W. Dicke, "Simulation dymanischer Merkmalskopplungen in einem neuronalen Netzwerkmodell," Ph.D. dissertation (Phillips University, Marburg, Germany, 1992).
9. A. S. French and R. B. Stein, "A flexible neural analog using integrated circuits," IEEE Trans. Bio-Med Electron. **BME-17,** 248–253 (1970).
10. C. Giles and T. Maxwell, "Learning, invariance, and generalization in high-order neural networks," Appl. Opt. **26,** 4972–4978 (1987).
11. C. Giles, C. Miller, D. Chen, H. Chen, G. Sun, and Y. Lee, "Learning and extracting finite state automata with second-order recurrent neural networks," Neural Comput. **2,** 393–405 (1992).
12. Appendix 1 of Ref. 1 shows that a step function for the pulse generator was used in the simulations there.
13. S. Grossberg and D. Somers, "Synchronized oscillators during cooperative feature linking in a cortical model of visual perception," Neural Networks **4,** 453–466 (1991).
14. N. Farhat and M. Eldefrawy, "The bifurcating neuron," in *Annual Meeting,* Vol. 17 of 1991 OSA Technical Digest Series (Optical Society of America, Washington, D.C., 1991), p. 10.
15. S. Grossberg, *Studies of Mind and Brain* (Reidel, Dordrecht, The Netherlands, 1982), pp. 49–50.
16. C. Giles, R. Griffin, T. Maxwell, "Encoding geometrical invariances in higher-order neural networks," in *Proceedings of the IEEE First International Neural Information Processing Systems Conference,* D. Anderson, ed. (Institute of Electrical and Electronics Engineers, New York, 1988), pp. 301–306.
17. A. V. Holden, M. Markus, and H. G. Othmer, eds. *Nonlinear Wave Processes in Excitable Media* (Plenum, New York, to be published).
18. A. B. Medvinsky, A. V. Panfilov, and A. M. Pertsov, "Autowave process characterization in heart tissue," in *Self-Organization, Autowaves and Structures Far From Equilibrium,* V. I. Krinsky, ed. (Springer-Verlag, New York, 1984), pp. 195–199.
19. J. L. Johnson, "Pulse-coupled neural networks," in *Critical Reviews of Optical Science and Technology,* S. Chen and H. J. Caulfield, eds., Proc. Soc. Photo-Opt. Instrum. Eng. **CR55:** 47–76 (1994).
20. J. L. Johnson, H. Ranganath, and H. J. Caulfield, "Pulse-coupled neural networks," in *Temporal Dynamics and Time Variant Pattern Recognition,* J. Dayhoff, ed., (Ablex, Norwood, N.J., to be published).