A Multitask Visual Information Processor with a Biologically Motivated Design

M. M. GUPTA AND G. K. KNOPF

Intelligent Systems Research Laboratory, College of Engineering, University of Saskatchewan, Saskatoon, Saskatchewan, Canada S7N 0WO

Received January 25, 1991; accepted August 15, 1991

1. INTRODUCTION

The successful operation of any autonomous intelligent machine depends upon its ability to cope with a variety of unexpected events. These machines must independently perceive, memorize, and comprehend the constantly changing real-world environment in which they operate. By virtue of its remoteness from the physical scene and its diverse informative nature, vision is considered the richest of all the biological sensory processes. Our own visual experience is relatively quick, effortless, and highly robust. At present, this sophisticated form of perceptual vision is within the exclusive domain of higher-order biological organisms such as human beings. These basic attributes are absent in all existing computer vision algorithms and hardware systems which are developed for machine applications. In general, a machine vision system is not very robust because it is very slow and intolerant to stimulus variability.

Biological vision employs millions of very slow and largely unreliable processors called neurons to achieve complex computational operations that are far beyond the capabilities of even the most powerful digital computers in existence. A single neuron requires about 2 ms (or an equivalent bandwidth of about 500 Hz) to fire and transmit a response to all other interconnected neurons. Ironically, the entire process of recognizing a complex scene with these relatively slow neural processors takes only 70–200 ms [30]. In sharp contrast, the basic digital processor employed by a computer is up to a million times faster than a biological neuron, and yet a computer vision system takes several minutes, if not hours, to process a single static image. Furthermore, many aspects of early biological vision are achieved in only 18 to 46 transformation steps [29, 30], far fewer than the millions of transformation steps needed by a sequential algorithm implemented on a digital processor. Clearly, an understanding of the neural structures involved in biological vision will be of great help in designing more robust machine vision systems for industrial, medical science, and space applications.

The incredible speed and flexibility of human vision is both the necessary proof that robust vision is possible and the framework from which a comparable machine vision systems can be developed. From the perspective of a vision system engineer it is not necessary to emulate the precise electrophysiological behavior of biological vision. Rather, it is desirable to replicate some of the computational operations involved in storing and processing spatio-temporal visual information. In this way, the scientific principles derived from vision physiology are used to design and develop a more effective “engineered” machine vision system. The term neuro-vision is used to refer to any artificial or machine vision system that embodies the computational principles exhibited by biological neural circuits. The process of designing artificial neuro-vision systems based upon biological analogies is termed reverse bioengineering [18, 29, 30] or inverse biomedical engineering [13].

The objective of this paper is to introduce the computational architecture of a multifunctional neural informa-
tion processor, called the Positive–Negative (PN) neural processor, and examine briefly how it can be used to perform a variety of tasks associated with early vision. The computational operations employed by the PN neural processor are loosely based upon the functional dynamics exhibited by certain cortical and thalamic nervous tissue layers situated in the brain [31, 32]. The role of the PN neural processor extends beyond the scope of most existing artificial neural networks because it provides a plausible hardware structure for realizing diverse aspects of early vision. This neural processor can be programmed to generate various steady-state and temporal phenomena that may be employed for short-term visual memory (STVM), spatio-temporal filtering (STF), and pulse frequency modulation (PFM). In this context, a robust neuro-vision system may be constructed from a parallel and hierarchical architecture of numerous individually programmed PN neural processors.

2. THE COMPUTATIONAL ARCHITECTURE OF THE PN NEURAL PROCESSOR

The biological visual pathway is composed of numerous distinct anatomical regions such as the retina, lateral geniculate nucleus (LGN), primary visual cortex, secondary visual cortex, and other cortical regions of the brain. All visual information which is sensed by the eye is projected through these different anatomical regions such that the relative spatial and temporal relationships found to exist on the photoreceptive surface of the retina are maintained throughout each advancing region. Each anatomical region or subregion may, therefore, be envisioned as being a functionally two-dimensional nervous tissue layer. All aspects of neural information processing that occurs within such a nervous tissue layer exist as spatio-temporal patterns of neural activity.

The majority of neural network models described in the existing literature consider the behavior of a single neuron as the basic computing unit for describing neural information processing operations. Often each computing unit in the network is based on the concept of an idealized neuron. An ideal neuron is assumed to respond optimally to the applied inputs. However, experimental studies in neurophysiology show that the response of a biological neuron is a variable, and only by averaging many observations is it possible to obtain predictable results [10, 11, 25, 26, 31, 32]. This observed variability in the response of a neuron is a function of both the uncontrolled extraneous electrical signals that are being received from activated neurons in other parts of the nervous system and the intrinsic fluctuations of electrical membrane potential within the individual neuron.

In general, a biological neuron is an unpredictable mechanism for processing information. Mathematical analysis has shown that the random behavior of individual neurons can transmit reliable information if they are sufficiently redundant in numbers [32]. It is postulated, therefore, that the collective activity generated by large numbers of locally redundant neurons is more significant in a computational context than the activity generated by a single neuron [22, 25, 26, 31, 32].

Each nervous tissue layer may be conceptualized as being a two-dimensional array of radial columns, whereupon each column contains one or more classes of locally redundant neurons. The total neural activity generated within a radial column results from a spatially localized assembly of nerve cells called a neural population. Each neural population may be further divided into several coexisting subpopulations. A single subpopulation is assumed to contain a large class of similar acting neurons that lie in close spatial proximity. The individual synaptic connections within any subpopulation are random, but dense enough to ensure at least one mutual connection between any two neurons. Furthermore, these neurons within a subpopulation are assumed to receive a common input and exhibit a common output.

For analytical simplicity only two subpopulations are assumed to coexist within each radial column. The first subpopulation of this column contains only excitatory neurons which project a positive influence when they are active. The second subpopulation that coexists within this column contains only inhibitory neurons which project a negative influence when they are active. Each excitatory or inhibitory subpopulation may be further composed of one or more types of nerve cells.

The dynamic neural activities exhibited by a nervous tissue layer is a result of immense coupling interactions that occur among the radially and laterally distributed subpopulations. In this context, the basic neural computing unit for describing the functional operations performed by a nervous tissue layer corresponds to the total neural activity generated by an individual subpopulation. Two antagonistic neural computing units are assumed to coexist at each spatial location on the tissue layer surface. These positive (P) and negative (N) neural units reflect the activity generated within the respective excitatory and inhibitory subpopulations as illustrated in Fig. 1.

The computational architecture of a generalized nervous tissue layer, as shown in Fig. 2, is represented by a two-dimensional PN neural processor composed of densely interconnected antagonistic positive and negative neural units. Each neural unit receives inputs from an external source as well as all other neural units in the processor. All steady-state and temporal phenomena associated with neural information processing stem from the coupled interactions of these antagonistic neural
units. The basic neural information processing element in the PN neural processor is, therefore, a spatially localized pair of antagonistic positive and negative neural computing units. A mathematical model of the PN neural processor is described in the following section.

3. THE MATHEMATICAL MODEL OF THE POSITIVE AND NEGATIVE NEURAL UNITS

The two state variables that describe the dynamic activity exhibited by the neural computing units at spatial location \((s, r)\) are defined as the proportion of positive (excitatory) neurons active, \(x_P(s, r, k)\), and the proportion of negative (inhibitory) neurons active, \(x_N(s, r, k)\), at the spatio-temporal location \((s, r, k)\). A neuron is assumed to be active if it is firing a sequence of action potentials and the measured axon potential at time \(k\) is greater than the rest potential. Since the response activities of the individual nerve cells in a neural unit are asynchronous, the proportion of neurons active at time \(k\) is composed of nonrefractory neurons that received new inputs and refractory neurons whose action potentials have not fully decayed.

The functional dynamics exhibited by a neural computing unit is defined by a first-order nonlinear state equation. For a discrete-time model of the neural unit dynamics, it is assumed that the sampling period is less than the time constant of the action potential generated by any constituent neuron. The state variables \(x_P(s, r, k+1)\) and \(x_N(s, r, k+1)\) generated at time \((k+1)\) by the positive and the negative neural units at spatial location \((s, r)\) of a two-dimensional PN neural processor are modeled as

\[
x_P(s, r, k + 1) = \alpha x_P(s, r, k) + \beta y_P(s, r, k) - \gamma x_P(s, r, k) \cdot y_P(s, r, k)
\]

\[
x_N(s, r, k + 1) = \alpha x_N(s, r, k) + \beta y_N(s, r, k) - \gamma x_N(s, r, k) \cdot y_N(s, r, k)
\]

The three most important factors that affect the dynamic properties of a neural unit are:

(i) The proportion of neurons active at time \((k+1)\) that were active during the previous sampling period at time \(k\). This factor is given by the term \(\alpha x_P(s, r, k)\) in Eqs. (1) and (2), where the subscript \(\psi\) indicates either a positive (P) or a negative (N) state, and \(\alpha < 1\) is the rate of decay in the proportion of neurons that are still active after one sampling period.

(ii) The proportion of neurons in the neural unit that is receiving inputs greater than an intrinsic threshold, \(y_P(s, r, k)\). This factor is given by the term \(\beta y_P(s, r, k)\), where \(\beta\) is the rate of growth in the proportion of neurons that become active during one sampling period. The rate of growth in neural activity is often assumed to be \(\beta = 1 - \alpha\).

(iii) The proportion of neurons in the neural unit that are refractory but still receiving inputs greater than an intrinsic threshold is given by the expression \(\gamma x_P(s, r, k) \cdot y_N(s, r, k)\). The rate of growth in refractory neurons during one sampling period is \(\gamma\), where \(\gamma < \beta\).

The proportion of neurons in a neural computing unit that receive inputs greater than a threshold value is given as

\[
x_N(s, r, k + 1) = \alpha x_N(s, r, k) + \beta y_N(s, r, k) - \gamma x_N(s, r, k) \cdot y_N(s, r, k)
\]
by a nonlinear function of the total applied inputs $u_\phi(s, r, k)$. This nonlinear input transformation function, $f_\phi(u_\phi(s, r, k))$, is related to the distribution of the neural thresholds, $g_\phi(u_\phi(s, r, k))$, within the neural unit. If the probability distribution of these neural thresholds about an aggregate value $\theta_\phi$ is given by a unimodal distribution, as shown in Fig. 3a, then the nonlinear input transformation is sigmoidal as shown in Fig. 3b. Thus, the proportion of neurons in a neural unit receiving inputs greater than an intrinsic threshold is given by the expression

$$y_\phi(s, r, k) = f_\phi(u_\phi(s, r, k)) = \int_{-\infty}^{u_\phi(s, r, k)} g_\phi(u_\phi(s, r, k)) du_\phi(s, r, k),$$

where the parameter pair $(v_\phi, \theta_\phi)$ determines the transformational properties of the function $f_\phi(\cdot)$. The parameter $v_\phi$ is defined as the maximum slope of the sigmoidal relationship at the point of inflection given by the aggregate value $\theta_\phi$. In other words, for a particular distribution of neural thresholds it is possible to determine the proportion of neurons receiving supra-threshold inputs by integrating the neural threshold distribution over the total applied inputs, Eq. (3).

The physiological significance of the sigmoidal curve in Fig. 3b is that for low levels of excitation most neurons within a neural unit will not be excited, whereas for very strong levels of excitation nearly all neurons become excited. An important assumption in deriving this function is that each neural unit is composed of only one "type" of neuron. This enables us to define the distribution of neural thresholds as a unimodal function, Fig. 3a. If more than one type of cell exists, then the nonlinear input transformation function, $f_\phi(\cdot)$, becomes more complex. A unimodal threshold distribution is given as

$$g_\phi(u_\phi(s, r, k)) = \frac{1}{2} (v_\phi \operatorname{sech}(v_\phi(u_\phi(s, r, k) - \theta_\phi)))$$

and the corresponding monotonically increasing nonlinear input transformation function is

$$f_\phi(u_\phi(s, r, k)) = \frac{1}{2} [1 + \tanh(v_\phi(u_\phi(s, r, k) - \theta_\phi))],$$

where the subscript $\psi$ is either P or N.

The input incident on a positive neural unit is defined as the proportion of neurons receiving supra-threshold inputs, and is given as

$$y_\psi(s, r, k) = f_\psi(u_\psi(s, r, k)).$$

Correspondingly, the input incident on the negative neural unit is given as

$$y_N(s, r, k) = f_N(u_N(s, r, k)).$$

The total inputs $u_p(s, r, k)$ and $u_N(s, r, k)$ applied to the corresponding neural unit at location $(s, r)$ and time $k$ are given by

$$u_p(s, r, k) = \omega_\psi \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} n_\psi(i - s, j - r) x_p(i, j, k)$$

$$+ S_p(s, r, k)$$

$$u_N(s, r, k) = \omega_N \sum_{i=0}^{I-1} \sum_{j=0}^{J-1} n_N(i - s, j - r) x_N(i, j, k)$$

$$+ S_N(s, r, k)$$

FIG. 3. Unimodal threshold distribution and the corresponding sigmoidal input transformation function for determining the proportion of neurons with total inputs, $u_\phi(s, r, k)$, greater than the threshold value. In this example $\theta_\phi = 1.5$ and $v_\phi = 1.5$, where the subscript $\phi$ is either P or N. (a) A unimodal probability distribution $g_\phi(u_\phi(s, r, k))$ of the neural thresholds about $\theta_\phi$. (b) The nonlinear sigmoidal function $f_\phi(u_\phi(s, r, k))$ arising from the neural threshold distribution given in (a).
and

\[
   u_n(s, r, k) = \omega_{pn} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{pn}(i - s, j - r)x_p(i, j, k) - \omega_{nn} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{nn}(i - s, j - r)x_n(i, j, k) + S_n(s, r, k),
\]

where \( S_p(s, r, k) \) and \( S_n(s, r, k) \) are the external visual stimuli, \( \omega_{pn} \) is the mean synaptic weight of all the neural connections from the neural unit \( \psi \) to \( \psi' \), and \( n_{\psi\psi'}(i - s, j - r) \) is a normalized spatial distribution function that describes the relative strength of the connections between the neurons in these different neural units with respect to the lateral distance \((i - s, j - r)\) which separates them.

If it is assumed that the spatial distribution function is isotropic, then the strength of the synaptic weights will decrease monotonically with distance from the origin at location \((s, r)\). This condition of isotropy ensures that the relative strength of the connections between the neural units are

\[
   n_{\psi\psi'}(i - s, j - r) = n_{\psi\psi'}(r - s, q - r)
\]

for any \( \psi \) neural unit located at \((i, j)\) or \((\rho, q)\) of equal distance from the \( \psi' \) neural unit at location \((s, r)\), regardless of direction.

The isotropic assumption enables the PN neural processor to exhibit the lateral excitatory on-center and inhibitory off-surround interactions commonly found in the nervous system. These neural interactions are most commonly defined as a circular receptive field [4, 10, 19–21] from which the inputs converge to a single sensory neuron. Neural network models with extensive lateral inhibition are often employed to describe these receptive field structures. However, a simple network model containing only lateral inhibition [10] is incapable of describing the complex computational operations associated with neural information processing because most nervous tissue layers, including those found in the visual cortex, contain densely interconnected populations of excitatory and inhibitory neurons. These more complex lateral on-center off-surround field structures are found in various nervous tissue layers associated with the hippocampus, neocortex, and cerebellum [12].

A normalized spatial distribution function incorporating this isotropic assumption is described as a two-dimensional Gaussian function written as

\[
   n_{\psi\psi'}(i - s, j - r) = \frac{1}{(2\pi \sigma_{\psi\psi'}^2)^{\frac{1}{2}}} \exp\left(-\frac{(i - s)^2 + (j - r)^2}{2\sigma_{\psi\psi'}^2}\right),
\]

where \( \sigma_{\psi\psi'} \) is the standard deviation or spatial spread of the synaptic connections with respect to the lateral distance. A block diagram of Eqs. (1) to (11) for the spatially isolated antagonistic neural units is shown in Fig. 4.

A variety of spatial diffusion characteristics associated with neural field structures can be realized by employing different normalized spatial distribution functions. For example, if the spatial distribution function is assumed to be symmetric but directionally dependent, then the neural units will respond to regional orientations in the stimulus pattern. This response behavior is similar to the activity of the simple cells situated in the primary visual cortex [19, 21]. The numerous different possibilities for the spatial distribution functions are beyond the scope of the present paper, and therefore, this discussion is restricted to the isotropic case.

4. A QUALITATIVE DESCRIPTION OF THE PN NEURAL PROCESSOR DYNAMICS

The PN neural processor is not designed to emulate any specific anatomical region in the biological vision system. Rather, it represents a simplified mathematical model of a generalized nervous tissue layer. By modifying the parameters of the PN neural processor it is possible to generate a variety of different dynamical modes that can be correlated with biological neural activity [1–3, 22, 31, 32]. From a computational perspective, this neural processor is capable of generating a variety of steady state and temporal phenomena that can be used to perform tasks such as STVM, STF, and PFM to name but a few.

The spatio-temporal stimulus, \( S_p(k) \), used for the following simulation studies is defined as a scaled spatial pattern \( b \cdot \lambda_p \) that varies with respect to a temporal function \( \Phi(k) \) combined with a time-varying pseudo-random noise component \( b \cdot \eta(k) \). Thus, the stimulus is written as

\[
   S_p(k) = b(\lambda_p \cdot \Phi(k) + \eta(k)),
\]

where

\[
   S_p(k) = \bigcup_{i=0}^{l-1} \bigcup_{j=0}^{l-1} \{S_p(i, j, k)\},
\]

\( b \) is a scalar that defines the upper bound of the stimulus intensity, \( S_p(k) \), over the interval \([0, b]\) for \( b \leq 1 \), and \( \cup \) is the union operation.

Unless it is stated otherwise, the spatial pattern \( \lambda_p \) is defined as a \((200 \times 200)\) pixel array with individual intensity values distributed over the range \([0, 1.0]\) and is written as

\[
   \lambda_p = \bigcup_{i=0}^{l-1} \bigcup_{j=0}^{l-1} \{\lambda_p(i, j, k)\}, \quad I = J = 200.
\]
The temporal function is defined as a pulse wave over the period \( k \in [k_a, k_b] \),

\[
\Phi(k) = \begin{cases} 
0 & \text{for } k < k_a, k > k_b \\
1.0 & \text{for } k_a \leq k \leq k_b .
\end{cases}
\]  

(14)

where \( k_a \) corresponds to the instant in time when the spatial pattern is suddenly applied to the PN neural processor and \( k_b \) is the instant when the pulse ceases.

The noise component \( \eta(k) \) is a pseudo-random variable bounded over the range \([0, 0.2]\). This term is written as

\[
\eta(k) = \sum_{i=0}^{l-1} \sum_{j=0}^{j-1} \{ \eta(i, j, k) \} .
\]  

(15)

The stimuli for the negative neural units are zero; that is, \( S_N(k) = 0 \), for all time \( k \).

A direct analytical solution for determining the steady-state and temporal behavior exhibited by the PN neural processor is not possible because of the inherent nonlinearities in Eqs. (1) and (2). However, these nonlinear equations can be analyzed qualitatively by obtaining phase trajectories in the \( x_p - x_N \) phase plane [9, 31]. These trajectories enable the system characteristics to be observed without solving the nonlinear equations. The locus of points where the phase trajectories have a given slope is called an isocline curve.

The steady-state activity exhibited by the various neural units of the PN neural processor is investigated by determining the isocline curves for \( \Delta x_p(s, r) = 0 \) and \( \Delta x_N(s, r) = 0 \) in the \( x_p - x_N \) phase plane. From Eqs. (1) and (2) these isocline curves are written as

\[
\Delta x_p(s, r) \triangleq x_p(s, r, k + 1) - x_p(s, r, k) = 0 \quad (16a)
\]

and

\[
\Delta x_N(s, r) \triangleq x_N(s, r, k + 1) - x_N(s, r, k) = 0 . \quad (16b)
\]
From the assumption of an isotropic spatial distribution function, Eq. (11), the isocline curve for $\Delta x_p(s, r) = 0$ in terms of the state variable $x_N(s, r, k)$ is given by

$$
\Delta x_N(s, r, k) = \frac{2\pi \sigma_{xp}}{\omega_{np}} \left\{ \omega_{pp} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{pp}(i - s, j - r) x_p(i, j, k) - \omega_{np} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{np}(i - s, j - r) \cdot x_N(i, j, k) - f_p^{-1} \left( \frac{1 - \alpha}{\beta} x_p(s, r, k) \right) \cdot S_p(s, r, k) \right\}
$$

and the corresponding isocline curve for $\Delta x_p(s, r) = 0$ in terms of the state variable $x_p(s, r, k)$ is given by

$$
\Delta x_p(s, r, k) = \frac{2\pi \sigma_{xp}^2}{\omega_{PN}} \left\{ -\omega_{PN} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{pp}(i - s, j - r) x_p(i, j, k) + \omega_{NN} \sum_{i=0}^{l-1} \sum_{j=0}^{l-1} n_{nn}(i - s, j - r) \cdot x_N(i, j, k) + f_N^{-1} \left( \frac{1 - \alpha}{\beta} x_N(s, r, k) \right) \cdot S_N(s, r, k) \right\}.
$$

In order to determine the shape of the isocline curves for $\Delta x_p(s, r) = 0$ and $\Delta x_N(s, r) = 0$ it is necessary to first establish equilibrium conditions throughout the entire PN neural processor for a given stimulus pattern. A stimulus pattern with uniform intensity is used to initially determine the parameters necessary for a PN neural processor to generate a particular mode of response activity.

4.1. Multiple State Attractors: Short-Term Visual Memory

The intersections arising from the isoclines for $\Delta x_p(s, r) = 0$ and $\Delta x_N(s, r) = 0$, Fig. 5, represent the steady-state values known as state attractors. For neural units with only one type of positive or negative neuron, the number of state attractors is generally one or three. If more than one intersection exists then these state attractors alternate between regions of stability and instability [22, 31]. Linearized stability analysis about each state attractor can be used to determine local stability properties [22]. The effect of changing the magnitude of the external stimulus, $S_p(s, r, k)$, is to translate the isocline curve for $\Delta x_p(s, r) = 0$ parallel to the $x_N$ axis, thereby altering the position and possibly even the number of state attractors. This effect is shown in Fig. 6. Typical parameters that are used to generate three state attractors for STVM are given in Table 1.

The observation that the position and the number of state attractors can be altered by a stimulus, $S_p(s, r, k)$, suggests that multiple attractors may generate a hysteresis loop over a specific range of stimulus intensities as shown Fig. 7. An important property of this hysteresis phenomena is that in order for the state activity, $x_N(s, r, k)$, to switch from a rest state to an excited state requires
TABLE 1

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\alpha$</td>
<td>0.9</td>
</tr>
<tr>
<td>$\beta$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0.1</td>
</tr>
<tr>
<td>$\omega_p$</td>
<td>1.5</td>
</tr>
<tr>
<td>$\omega_\theta$</td>
<td>1.5</td>
</tr>
<tr>
<td>$\omega_{PN}$</td>
<td>2</td>
</tr>
<tr>
<td>$\omega_{PNN}$</td>
<td>4</td>
</tr>
<tr>
<td>$\omega_{NN}$</td>
<td>4</td>
</tr>
<tr>
<td>$\omega_{PN}$</td>
<td>0.4</td>
</tr>
<tr>
<td>$\omega_{NN}$</td>
<td>0.8</td>
</tr>
<tr>
<td>$\omega_{PN}$</td>
<td>0.8</td>
</tr>
<tr>
<td>$\omega_{NN}$</td>
<td>0.4</td>
</tr>
</tbody>
</table>

Typical Parameters Used to Generate Three State Attractors for Short-Term Visual Memory (STVM)

4.2.1. Transient behavior: Spatio-temporal filtering. A stimulus applied to a nervous tissue layer results in the formation of spatio-temporal patterns of neural activity. These patterns of neural activity arise from the nervous tissue time constants and the laterally distributed feedback that exist among the constituent neurons. This response behavior implies that the various neural units, or subpopulations, of a nervous tissue layer will simultaneously process the neural signals in both space and time [8, 10, 23, 32]. By integrating both spatial and temporal components into the same neural response activity, a nervous tissue layer will respond strongly to contiguous and coherent attributes in the stimulus and not to the random fluctuations contained therein.

One important task of early biological vision is to reduce the noise and unwanted details in the visual stimulus pattern without significantly degrading the embedded information. This role of visual information processing can be summarized as reducing the ambiguities inherent in the stimulus pattern while enhancing local discontinuities in an effort to better partition the stimulus pattern into coherent regions for additional analysis by the higher cortical levels. Although the visual stimulus pattern is initially received via the retina by a two-dimensional array of independent photoreceptors, the information contained within this pattern is visually perceived by humans as cohesive surfaces, boundaries, colors, and movements, and not as random variations in the incident inten-

FIG. 7. The hysteresis loop arising from the isocline curves for $\Delta x_p(s, r) = 0$ and $\Delta x_N(s, r) = 0$ when the stimulus intensity is varied over the range $[-1.0, 1.0]$. The term $\Omega_l$ is the lower threshold intensity and the term $\Omega_H$ is the upper threshold intensity necessary for the neural activity to switch states. For the parameters given in Table 1 these thresholds are $\Omega_l = -0.05$ and $\Omega_H = +0.4$. 

the stimulus intensity to be greater than some upper threshold value, $S_p(s, r, k) > \Omega_H$. Similarly, the state activity can return to its original rest state by a future stimulus with an intensity less than a lower threshold value, $S_p(s, r, k) < \Omega_L$. The hysteresis phenomenon exhibits an important noise immunity characteristic beneficial for storing features from a noisy stimulus [5, 14, 15, 17, 22, 27, 31]. The inherent upper and lower thresholds prevent small perturbations or fluctuations within the applied stimulus pattern from drastically altering the PN neural processor steady-state activity. A stimulus that varies significantly from its prior value is required to alter the neural unit response activity.

A PN neural processor with neural units that exhibit localized hysteresis phenomena can function as a form of STVM [14-17, 22, 31]. This type of memory is the result of ongoing neural activity that is maintained by the dense excitatory feedback which exists among the various neural units. The neural activity will reverberate or circulate within the closed-loop neural circuitry of the PN neural processor such that the overall output response activity will remain constant, even after the stimulus pattern is removed. In this way any recipient information can be recalled while this reverberation occurs, and it will continue to occur until a strong inhibitory influence destroys the reverberating activity. Figure 8 is an example of STVM where the individual neural units of the PN neural processor exhibit a single hysteresis loop. A paper describing a PN neural processor exhibiting multiple hysteresis phenomena is being prepared by the authors. Any additional complexities introduced by these neural units are compensated for by the overall simplicity in the network’s performance. For example, changes to the contents of visual memory occur without any physical modifications to the strength of the synaptic connections.

4.2. Temporal Phenomena: Transient Behavior and Limit Cycle Oscillations

The temporal behavior generated by a neural unit with multiple steady-state characteristics is primarily the switching action between different stable attractors. However, there are two classes of temporal responses generated by a PN neural processor which are very useful for information processing. Both classes of temporal behavior are generated around single attractors in the $x_p$–$x_N$ phase plane.
FIG. 8. Short-term visual memory (STVM) of a noisy space shuttle by the PN neural processor at times $k_1$ and $k_2$. The parameters used in this example are given in Table 1. The PN neural processor responses, $x_p(k)$ and $x_p(k_2)$, exhibit the important spatio-temporal noise suppression properties associated with the multiple state attractors. (a) The spatial pattern, $\lambda_p$, and the temporal component, $\Phi(k)$, of the visual stimulus. The upper bound of the stimulus intensity is $b = 0.5$. (b) The stimulus, $S_p(k_1)$, and the PN neural processor response, $x_p(k_1)$, at time $k_1$. (c) The stimulus, $S_p(k_2)$, and PN neural processor response, $x_p(k_2)$, at time $k_2$.

sities. Two simultaneous psychophysical manifestations of this neural information processing property mechanism that are commonly found throughout the early stage of human vision are constancy and contrast. Constancy is summarized as the grouping of similar spatio-temporal stimuli into a common response, whereas contrast is the segregation of dissimilar spatio-temporal stimuli.

Spatial constancy and contrast can be maintained in a
FIG. 9. Temporal properties of the spatio-temporal threshold associated with transient behavior arising from the parameters given in Table 2. In (b) the stimulus will generate a strong response only if it is applied for a duration of time greater than a temporal threshold, $k_b > k$. Even if the applied stimulus intensity is increased for a short temporal duration, as shown in (c), it will fail to generate a strong response. (a) The spatial pattern, $\lambda_p$, and the temporal component, $\Phi(k)$, of the stimulus. The upper bound of the stimulus intensity is $b = 1.0$. The spatial pattern used in this example is a $(100 \times 100)$ pixel array. (b) The temporal response of the positive neural unit at spatial location $(s, r)$ to different temporal profiles $[k_i, k_b]$ of the stimulus, where $i = 1, 2, 3, \text{and } 4$. (c) The temporal response of the positive neural unit at spatial location $(s, r)$ to a stimulus with an increasing intensity, $b = 0.5, 1.0, 1.5, 2.0, 2.5, \text{and } 3.0$, but applied for a short temporal duration from $k_i$ to $k_b$.

PN neural processor through the dense lateral feedback between the various spatially distributed neural units that generate localized transient behavior. Correspondingly, the temporal constancy and contrast are sustained by these same neural units responding in a dynamic fashion to both the stimulus and the lateral feedback inputs. The neural cells located in many of the sensory and cortical nervous tissue layers associated with biological vision are organized in such a way that the stimulation of a given neural unit will inhibit the activity of the neurons located in the neighboring neural units [4, 10, 19, 21]. The effect of this lateral inhibition is to enable the neurons to discriminate the information contained in the neural signals being received and processed.

The effect of interdependent spatial and temporal components of visual information processing is illustrated in Figs. 9 and 10. If a stimulus pattern occurs too briefly then the PN neural processor will exhibit a very weak response regardless of the stimulus intensity. For neurovision system applications this inherent spatio-temporal threshold will suppress any time-varying spatially distributed random intensity fluctuations associated with pixel noise. By suppressing these local disturbances the PN neural processor enforces regional constancy on the out-
put of each neural unit. Typical parameters that enable the neural processor to generate transient behavior for STF are given in Table 2.

An important attribute of any visual information processor is its sensitivity to the intensity contrast within the stimulus pattern. The sensitivity of the human visual system to the spatial spacing of a set of contrasting areas is known as contrast sensitivity [4, 19–21]. The input used

<table>
<thead>
<tr>
<th>TABLE 2</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Typical Parameters Used to Generate Transient Behavior for Spatio-Temporal Filtering (STF)</strong></td>
</tr>
<tr>
<td>parameter</td>
</tr>
<tr>
<td>( \alpha )</td>
</tr>
<tr>
<td>( \beta )</td>
</tr>
<tr>
<td>( \gamma )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{op}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{pp}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{pn}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{ppn}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{ppn}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{ppn}} )</td>
</tr>
<tr>
<td>( \bar{\omega}_{\text{ppn}} )</td>
</tr>
</tbody>
</table>

FIG. 10. The effect of the spatial dimensions of a stimulus on the steady-state response for a positive neural unit centered at \((s, r)\). In this example the spatial spread values are \( \sigma_{\text{pp}} = \sigma_{\text{PN}} = 3 \) and \( \sigma_{\text{ppn}} = \sigma_{\text{NN}} = 6 \). A very narrow stimulus applied for a time duration greater than the temporal threshold will still fail to trigger a strong steady-state response. The stimulus pattern of dimension \((50 \times 50)\) generates a lower steady-state response than \((10 \times 10)\) because of contrast effects at the boundaries. (a) The spatial patterns, \( \lambda_p \), and the temporal component, \( \Phi(k) \), of the stimuli. The upper bound of the stimulus intensity is \( b = 1.0 \). The various spatial patterns used in this example are \((100 \times 100)\) pixel arrays. (b) The cross-sectional profiles at instant \( k_b \) for the stimulus patterns given in (a). (c) The temporal response of the positive neural unit at spatial location \((s, r)\) to inputs with different spatial dimensions.
to determine contrast sensitivity is often given as a bar grating pattern in which the intensity along the horizontal profile varies in a sinusoidal manner. The units of spatial frequency of the bar grating patterns are expressed as cycles/millimeter. The contrast of the output pattern, $C_S(k)$, is measured by the maximum peak to minimum trough amplitude over the entire PN neural processor and is given by

$$C_S(k) = \left| \bigvee_{i} \bigwedge_{j} \{x_P(i, j, k)\} - \bigwedge_{i} \bigvee_{j} \{x_P(i, j, k)\} \right|,$$

where $\bigvee$ and $\bigwedge$ are the maximum and minimum operations, respectively. The spatial frequency of the PN neural processor response is given in terms of cycles/layer. This contrast sensitivity function, $C_S(k)$, enables the response of the visual information processor to be described as a function of different spatial frequencies.

The three contrast sensitivity curves shown in Fig. 11 illustrate that the PN neural processor can be tuned to different spatial frequencies by adjusting the spatial spread values used for the synaptic connections, $\sigma_{\phi'}$. The attenuation of both low and high spatial frequencies in the PN neural processor response is consistent with the contrast sensitivity observed in human vision [4, 19, 20].

The spatial and the temporal properties generated within a highly interconnected network of neural units which exhibit transient behavior can provide a plausible explanation for a variety of psychophysical observations.

In terms of visual perception, a sharp change in the contrast between an object and its background will cause two thin parallel bands of intensity to be perceived along the object contour. These bands are not physically present in the stimulus pattern but are caused by the neurons of the various nervous tissue layers inhibiting the activity of their neighboring nerve cells. This perceived enhancement of contrast is called the Mach Band effect [20]. The reflected light intensity of each strip in Fig. 12a is uniform over its entire width and differs from its neighbors by a constant amount. Pseudo-random noise is superimposed on the stimulus pattern. The visual appearance or perception of these gray scales is such that each strip is darker along its right side than its left. Figure 12b is a recreation of this Mach Band effect by the PN neural processor when it is programmed with the parameters in Table 2.

4.2.2. Limit Cycle Oscillations: Pulse Frequency Modulation. The second important temporal phenomenon that can be generated by the PN neural processor is limit cycle oscillations. The limit cycle oscillations are observed in response to a constant stimulation as shown in Fig. 13. There exists a threshold value, $\Omega_L$, for the stimulus intensity which must be exceeded in order to evoke such oscillatory behavior. This lower threshold $\Omega_L$ is regarded as a necessary characteristic for low-level noise immunity. For a supra-threshold stimulus the frequency of the oscillations is a monotonic function of intensity as illustrated in Fig. 14. For extremely large intensities the limit cycle activity is extinguished because the neural units become saturated with positive or excitatory activity. Thus, there also exists an upper threshold, $\Omega_H$, for generating limit cycle oscillations. Typical parameters that enable the PN neural processor to generate spatially localized limit cycle oscillations are given in Table 3.

The capability of generating limit cycle behavior enables the PN neural processor to uniquely encode the intensities of the analog visual information. Engineers know from communication theory that pulse frequency-modulated signals are less susceptible to channel noise than are amplitude-modulated (AM) signals and are, therefore, the preferred method of transmitting information over long distances.

| $\alpha$ | $\beta$ | $\gamma$ |
| 0.9 | 0.1 | 0.1 |
| $\omega_0$ | $\phi_1$ | $\phi_2$ |
| 1.5 | 2.0 | 1.5 |
| $\omega_{\phi'}$ | $\sigma_{\phi'}$ | $\omega_P$ |
| 8 | 8 | 8 |
| $\sigma_{\alpha\phi'}$ | $\sigma_{\phi}$ | $\sigma_P$ |
| 0 | 0 | 0 |

In terms of visual perception, a sharp change in the contrast between an object and its background will cause two thin parallel bands of intensity to be perceived along the object contour. These bands are not physically present in the stimulus pattern but are caused by the neurons of the various nervous tissue layers inhibiting the activity of their neighboring nerve cells. This perceived enhancement of contrast is called the Mach Band effect [20]. The reflected light intensity of each strip in Fig. 12a is uniform over its entire width and differs from its neighbors by a constant amount. Pseudo-random noise is superimposed on the stimulus pattern. The visual appearance or perception of these gray scales is such that each strip is darker along its right side than its left. Figure 12b is a recreation of this Mach Band effect by the PN neural processor when it is programmed with the parameters in Table 2.

4.2.2. Limit Cycle Oscillations: Pulse Frequency Modulation. The second important temporal phenomenon that can be generated by the PN neural processor is limit cycle oscillations. The limit cycle oscillations are observed in response to a constant stimulation as shown in Fig. 13. There exists a threshold value, $\Omega_L$, for the stimulus intensity which must be exceeded in order to evoke such oscillatory behavior. This lower threshold $\Omega_L$ is regarded as a necessary characteristic for low-level noise immunity. For a supra-threshold stimulus the frequency of the oscillations is a monotonic function of intensity as illustrated in Fig. 14. For extremely large intensities the limit cycle activity is extinguished because the neural units become saturated with positive or excitatory activity. Thus, there also exists an upper threshold, $\Omega_H$, for generating limit cycle oscillations. Typical parameters that enable the PN neural processor to generate spatially localized limit cycle oscillations are given in Table 3.

The capability of generating limit cycle behavior enables the PN neural processor to uniquely encode the intensities of the analog visual information. Engineers know from communication theory that pulse frequency-modulated signals are less susceptible to channel noise than are amplitude-modulated (AM) signals and are, therefore, the preferred method of transmitting information over long distances.

| TABLE 3 |
| Typical Parameters Used to Generate Limit Cycle Oscillations for Pulse Frequency Modulation (PFM) |
| $\alpha = 0.9$ | $\beta = 0.1$ | $\gamma = 0.1$ |
| $\omega_0 = 1.5$ | $\phi_1 = 2.0$ | $\phi_2 = 1.5$ |
| $\omega_{\phi'} = 8$ | $\sigma_{\phi'} = 8$ | $\omega_P = 8$ |
| $\sigma_{\alpha\phi'} = 0$ | $\sigma_{\phi} = 0$ | $\sigma_P = 0$ |
| $\sigma_{\phi} = 0$ | $\sigma_{\phi} = 0$ | $\sigma_{\phi} = 0$ |
However, for PFM applications it is necessary that the pulse-decoded response be a near duplicate of the original stimulus pattern. In order to prevent spatial diffusion characteristics from corrupting the frequency of limit cycle oscillations generated by the PN neural processor it is necessary to assume that the spatial spread of the synaptic connections is zero; that is, \( \sigma_{\omega f} = 0 \). In this way each spatially localized pair of antagonistic positive and negative neural units uniquely encodes the recipient stimulus intensity. If the physiological aspects of neural information processing are to be investigated \([10, 32]\) then the spatial diffusion characteristics are included in the PN neural processor such that \( \sigma_{\omega f} > 0 \).

Once a PFM signal is received at its destination it may be decoded back into an analog or AM signal for additional processing as shown in Fig. 15. A simple peak detector can be used to convert the limit cycle oscillations to pulses at the transmitter, and then a time-limited pulse counter with a nonlinear signal enhancer, Fig. 16, can be used to reconstruct the image at the receiver.

Figure 17 is an example of a stimulus pattern that is converted to limit cycle behavior by the PN neural processor and then reconstructed using the simple pulse frequency-to-analog converter described above.

5. THE PN NEURAL PROCESSOR: A RETROSPECT

The goal of this paper was to outline the computational capabilities of a multitask visual information processor, called the Positive–Negative neural processor, that may be incorporated into a real-time on-line robust neurovision system for engineering applications. The PN neural processor extended beyond the scope of most specialized networks in that it could perform a variety of diverse tasks associated with early biological vision. The PN neural processor was programmed to generate distinct steady-state and temporal phenomena in order to perform computational tasks such as storing, spatio-temporal filtering, and pulse frequency encoding two-dimensional time-varying signals. This paper did not address
the applicability of the PN neural processor to higher-level cognitive functions such as associative memory, learning, recognition, and reasoning. However, many of these computational operations may be possible because of the highly parallel and densely interconnected structure of this neural network. It was further postulated that this versatile network could simplify the design, development, and programmability of robust neuro-vision systems that must operate effectively in real-time.

In Section 2 the biological basis of information processing by nervous tissue layers was presented. The biological vision system was conceptualized as a shallow hierarchy of several distinct anatomical regions. Each anatomical region was assumed to be represented by one or more two-dimensional nervous tissue layers. A dynamic neural network model of a generic nervous tissue layer was proposed. This model was loosely based on a mathematical theory for describing the functional dynamics exhibited by cortical and thalamic nervous tissue layers as proposed by Wilson and Cowan [31, 32].

FIG. 12 Continued

The fundamental neural design and structure of a nervous tissue layer was generalized in the form of a PN neural processor. This dynamic neural processor was described as a two-dimensional array of densely interconnected positive and negative neural computing units. A neural computing unit was assumed to contain a large class of similar neurons that lie in close spatial proximity. Each neural unit was mathematically represented by a nonlinear first-order difference equation. The coupling interaction between individual positive and negative neural units resulted in a variety of dynamic activities which were associated with neural information processing. These distinct modes of dynamic neural activity were interpreted in Section 4 in terms of storing, filtering, and pulse frequency encoding visual information. This qualitative analysis was used to demonstrate the viability of the PN neural processor as a generic visual information processor for performing certain tasks of early vision. A PN neural processor with neural units that exhibited multiple steady states were shown to temporarily retain at-
tributes of the visual stimulus through the process of neural reverberation [21]. That is, the activity from all neural units circulated throughout the processor such that the overall response activity remained constant. No physical changes were imposed upon the synaptic weights that existed between the different neural units.

Psychophysiological experiments have shown that many perceptual aspects of early vision may be associated with transient neural activity. In this role, a PN neural processor was able to suppress small perturbations that corresponded to neural noise by simultaneously diverging and converging the neural activity in both space and time. Global manifestations of this process were given by the principles of constancy and contrast.

Finally, the ability of the PN neural processor to generate limit cycle oscillations suggested that it could also function as a pulse frequency encoder of the stimulus intensity. From practical considerations certain neurovision systems require that the information be transmitted over long distances through noisy channels. In this capacity, the pulse frequency-modulated signals are less
susceptible to channel noise then are analog or amplitude-modulated signals.

6. CONCLUSIONS

The process of early vision must occur in milliseconds and not the minutes or hours it takes all existing digital computer-based vision algorithms. To achieve these capabilities it is necessary to draw on our present knowledge of biological vision physiology in order to aid in the design and development of a more effective machine vision system. Within this framework a multitask visual information processor based loosely on a mathematical theory of the functional dynamics of cortical and thalamic nervous tissue [13–15, 22, 31, 32] was discussed. This neural processor was programmed to store, filter, and pulse frequency encode spatio-temporal visual information. Real-time perceptual vision may be possible by constructing a shallow parallel and hierarchical architecture from numerous PN neural processors that pre-process and store the visual information prior to performing high-level decision operations.

ACKNOWLEDGMENTS

This work has been supported in part by the Network of Centre of Excellence on Neuro-Vision Research (IRIS) and the Natural Sciences and Engineering Research Council of Canada. The authors are grateful to the reviewers for their helpful comments.

REFERENCES


