GaAs optoelectronic neuron arrays

Steven Lin, Annette Grot, Jiafu Luo, and Demetri Psaltis

A simple optoelectronic circuit integrated monolithically in GaAs to implement sigmoidal neuron responses is presented. The circuit integrates a light-emitting diode with one or two transistors and one or two photodetectors. The design considerations for building arrays with densities of up to 10^4 cm^{-2} are discussed.

1. Introduction

The two components required for the implementation of a neural network are neurons and connections. In an optical implementation the neurons are typically arranged as two-dimensional arrays that are interconnected via the third dimension.[1–4] The interconnections are realized with holograms or spatial light modulators. We describe neuron arrays fabricated monolithically on gallium arsenide (GaAs) substrates, with each neuron circuit consisting of a light-emitting diode (LED), one or two photodetectors, and one or more transistors. The use of optoelectronic circuits provides the flexibility of implementing complex neuron response functions and of fine tuning the properties of the neurons, as is required by the neural-network algorithm that is being implemented.

LED's and laser diodes are the two choices for on-chip light sources. Laser diodes have higher quantum efficiency and a more directed beam than LED's, which means higher light efficiency.[5,6] Unfortunately, electrical power dissipation is a limiting factor for high-density circuits. The maximum current that can be drawn to drive either an individual LED or a laser diode is

$$I = \frac{P_{\max} A}{N V_D}, \tag{1}$$

where V_D is the power supply voltage for each neuron, N is the maximum number of elements, $P_{\max}$ is the maximum power dissipation per unit area, and A is the area of the array. The optical power generated by each laser diode is

$$P_{\text{opt,LD}} = \eta_{\text{LD}}(I - I_{\text{th}}), \tag{2}$$

where η_{LD} is the external efficiency and I_{th} is the threshold current of the laser diode. Substituting in the expression for the maximum current, we obtain the following expression for the total optical power from the chip:

$$P_{\text{opt,LD}} N = \eta_{\text{LD}}\left(\frac{P_{\max} A}{V_{\text{D}}} - I_{\text{th}} N\right). \tag{3}$$

For typical values ($P_{\max} = 1$ W/cm^2, $V_{\text{D}} = 2$ V, and $I_{\text{th}} = 500$ μA) the total output power reduces to zero when $N = 10^3$ and $A = 1$ cm^2. This is clearly the maximum density of neurons we can achieve if we opt for laser diodes. Because of the absence of a threshold current, LED's can operate with small currents, permitting a density of up to 10^5/cm^2. Therefore the conclusion is that, if one is interested in high-density arrays that operate with relatively slow switching time, then LED's are preferred.

The other option for the optical output is an optical modulator.[7–9] The principle advantages of modulators derive from the fact that the light source is off-chip, so that the optical gain can be increased simply by making the source brighter without increasing appreciably the power consumption on the chip. This permits us to build a higher-density array or to obtain a faster switching time. Moreover, the external source makes a spatially coherent array possible, which is a necessary property for most adaptive neural-network architectures.[10] Light sources, such as LED's or laser diodes, on the other hand, have a high contrast ratio, require small driving voltages, have simpler epitaxial structure, and require fewer

When this work was performed the authors were with the Department of Electrical Engineering, California Institute of Technology, Pasadena, California 91125. S. Lin is now with Vitesse Semiconductor, 241 Calle Plano, Camarillo, California 93012.

Received 17 June 1992.
0003-6935/93/081275-15$05.00/0.
© 1993 Optical Society of America.

critical fabrication steps.[7] Moreover, with on-chip light sources it is generally much easier to build a system because no external light source, no accompanying beam splitter, and no beam-forming optics are required, nor is it necessary to tune the source wavelength to match the modulators. Therefore, in comparing quantum-well modulators with LED's for building optoelectronic neuron arrays, in principle, modulators outperform LED's in most respects. However, in practical terms, both for building a system and for fabricating a large array, LED's have strong advantages. Therefore LED-based neurons can make it possible in the near future to fabricate large dense neuron arrays for applications such as early vision processing, which do not require holographic adaptation.[11]

2. Design Considerations

The neuron performs two basic tasks. First, it calculates a simple nonlinear function, mapping the input signal it receives on its photodetector(s) to the output LED intensity. The most commonly used nonlinearity is the threshold function or its close relative, the sigmoidal response.[12] Other commonly used nonlinearities are polynomial mappings and bump functions. The neurons we describe perform the nonlinearity with simple transistor circuits that transform the photocurrent from the detector to a current that is drawn through the LED. The circuit shown in Fig. 1 implements a sigmoidal response. The LED is driven by a metal–semiconductor field-effect-transistor (MESFET) whose gate voltage is controlled by an input circuit consisting of a photodetector and a biasing element. The second function performed by the neuron is optical gain. This is necessary because optical interconnections are generally lossy, and therefore the signal level must be restored by the neural planes at each stage in order to make the implementation of multilayer networks possible. In this section we examine general trade-offs between the need to provide gain and the desire to have large dense neuron arrays. In Section 3 we describe specific methods for implementing the threshold function.

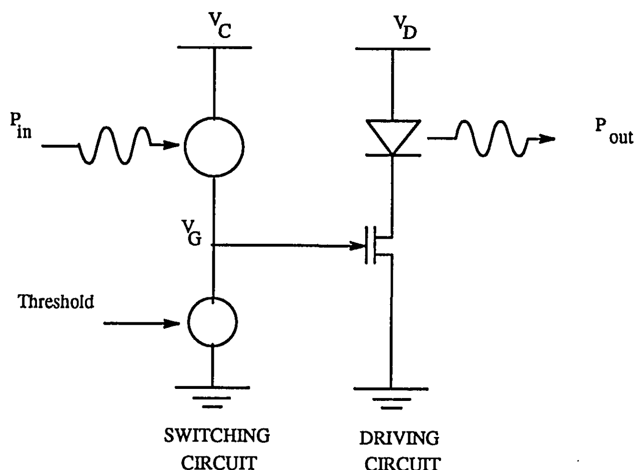

Fig. 1. Schematic diagram of the optoelectronic thresholding neuron circuit: V_D, power supply voltage; V_C, decector circuit voltage; V_G, LED-driving MESFET gate voltage.

A. Interconnection Loss

The interconnection loss is determined by a number of factors, including the type of light source used (coherent versus incoherent), the type of device used as the interconnection medium (e.g., holographic versus nonholographic), and the architecture of the network (e.g., the number of connections per neuron). The LED circuits that we describe produce spatially incoherent illumination. In general, incoherent systems are relatively inefficient because they radiate energy into a large cone angle, and only a portion of the radiated energy is captured by the numerical aperture of the optical system. To determine the dependence of the light efficiency on the number of connections per neuron, C, we write the strength of the optical interconnection between the ith and jth neurons, η_{ij}, as follows:

$$\eta_{ij} = \eta_0 H(C) w_{ij}, \tag{4}$$

where η_0 is the optical loss (in intensity) obtained when only two neurons are connected, which includes the numerical aperture loss mentioned above as well as reflections, the insertion loss of spatial light modulators, or the limitation on the diffraction efficiency owing to the holographic medium. $H(C)$ $[H(1) = 1]$ is a function that contains the dependence on C, and $0 \leq w_{ij} \leq 1$ is the normalized weight of the connection.[10]

The nonholographic interconnection system is shown in Fig. 2(a). Light with intensity x_i is emitted from the ith LED and is divided into C beams that impinge on C different spatial locations on the inter-

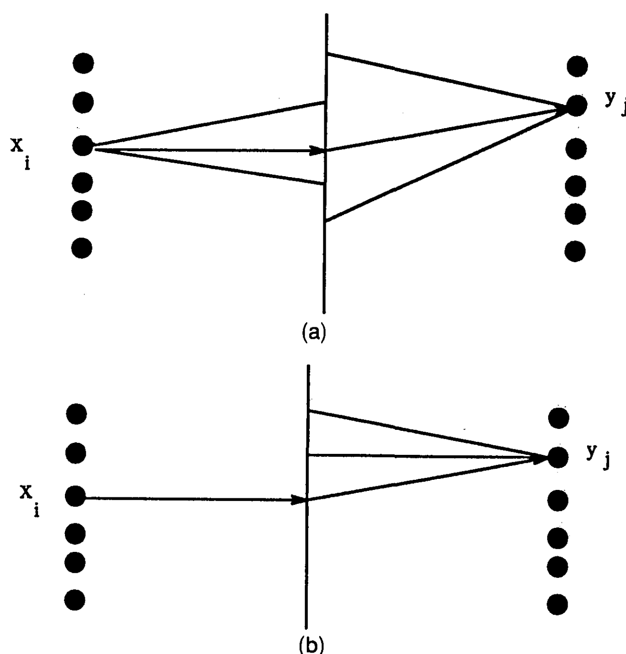

Fig. 2. (a) Nonholographic and (b) holographic interconnection optical neural-network implementations.

connection medium, one location for each neuron connected to the ith unit. The interconnection strength w_{ij} is recorded as the intensity transmittance of the medium at each location. The jth neuron collects light from C spatial locations on the interconnect medium to form its input as follows:

$$z_j = \sum_{i=1}^{C} \eta_{ij} x_i = \eta_0 H(C) \sum_{i=1}^{C} w_{ij} x_i. \quad (5)$$

In this case $H(C) = 1/C$, since the light from each LED is divided evenly. The output is at maximum when $x_i = P$ and $w_{ij} = 1$ for all i and j. Then $z_j = \eta_0 P$, which implies that the maximum efficiency of the optical connection is limited essentially by the loss owing to the finite numerical aperture of the system. The best known example of such an interconnection scheme is the vector–matrix realization of a neural network.[13]

A schematic diagram for holographic interconnections is shown in Fig. 2(b). Here, light from the ith input LED is collimated and illuminates a hologram on which G gratings are superimposed. The interconnection between the ith and jth neurons is realized by one of the gratings stored in the hologram by redirecting the light that is incident from the ith neuron towards the jth neuron. The interconnection weight w_{ij} is encoded in the strength of the grating. For a planar amplitude hologram, its effective amplitude transmittance as a function of position x is

$$t_H(x) = \sum_{G} [w_{ij} H(C)]^{1/2} \exp(j2\pi u_{ij} x), \quad (6)$$

where u_{ij} is the spatial frequency of the holographic grating that connects the ith and jth neurons. Since $0 \leq t_H \leq 1$, the amplitude of each grating must be small enough to enforce this constraint. Since the hologram is formed as an incoherent sum of G variables, t_H grows in proportion to $\sqrt{G}$.[2] The requirement that $t_H \leq 1$ is enforced if

$$H(C) \leq 1/G. \quad (7)$$

In the simplest case the total number of gratings G that are superimposed on the hologram equals C. This is the case of a completely shift-invariant interconnection pattern with each neuron connected to C others that have the same set of weights. In this case the efficiency of the connections is identical to the nonholographic efficiency. On the other hand, if each neuron connected by the hologram has a distinct receptive field, then we need to record a separate grating for each pair of neurons. Therefore $G = CN$, where N is the number of neurons. In this case, holographic interconnections are inferior to the nonholographic interconnections in terms of light efficiency by a factor N, which is typically in excess of 10^3. Since holographic interconnections are more efficient with coherent illumination,[2] the LED neuron circuits that we describe are best suited for holographic shift-invariant circuits (e.g., early vision tasks) or for nonholographic interconnection schemes for which spatial incoherence is actually preferable. For the remainder of the discussion we assume that $H(C) = 1/C$.

B. Density of Neurons

The most important consideration in designing the neuron arrays is the density N/A with which they can be fabricated, where A is the array area. One factor that limits the density is the geometric limit owing to lithography. For simple thresholding circuits this geometric limit can be in excess of 10^5 neurons/cm^2. In practice, the maximum permitted electrical power dissipation per unit area, $P_{\max}$, is the limitation. Thus it is important to reduce the power dissipation of each neuron in order to achieve high neuron densities. Since the electrical power dissipation is due mostly to current flowing through the LED, this means that we must be able to use low currents and hence low optical power from the LED. Consequently, the input-light sensitivity needs to be designed to be as low as possible. The minimum acceptable light level at the input of each neuron is determined by two factors: the noise level at the input (detector) part of the circuit and the nonuniformity in the setting of the threshold level owing to fabrication imperfections. In the following we present a simple statistical analysis that illustrates how the proper input light level is determined and that also provides us with an estimate for the density of neurons.

We estimate the minimum acceptable input power level by calculating the probability that a neuron makes an error, P_e, as a function of the optical power levels in the system. As the optical input power is reduced, we eventually reach the noise plateau of the circuit at which P_e increases beyond an acceptable level. We assume that the noise can be modeled as a Gaussian random variable with zero mean and standard deviation σ_n that is added to the signal photocurrent. To calculate P_e, we also need to characterize statistically the input signal. We assume that each element in the previous layer is on, with probability p and intensity P. The strength of a connection w_{ij} is 1 with probability q, and it is zero with probability $1 - q$. The input signal photocurrent $y_j = \eta_D z_j$ is the sum of C independent random variables. For large C, its distribution can be approximated by a Gaussian owing to the central limit theorem.[14] η_D is the detector efficiency in amperes per watt. Using Eq. (5), we can then calculate the mean μ_y and the variance σ_y^2 of the photocurrent as follows:

$$\mu_y = \frac{\eta_D \eta_0}{C} E\left(\sum_{i=1}^{C} w_{ij} x_i \right) = \eta_D \eta_0 P p q,$$

$$\sigma_y^2 = \left(\frac{\eta_D \eta_0}{C} \right)^2 E\left\{ \left(\sum_{i=1}^{C} w_{ij} x_i \right)^2 \right\} - \mu_y^2$$

$$= \frac{(\eta_D \eta_0 P)^2}{C} pq(1 - pq). \quad (8)$$

With these assumptions, P_e can be written as follows:

$$P_e = \frac{1}{2\pi\sigma_n\sigma_y}\left[\int_t^{\infty}\exp\left[-\frac{(n-t)^2}{2\sigma_n^2}\right]\right.$$
$$\times\int_t^{n}\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right]\mathrm{d}y\mathrm{d}n$$
$$+\int_{-\infty}^{t}\exp\left[-\frac{(n-t)^2}{2\sigma_n^2}\right]$$
$$\left.\times\int_n^{t}\exp\left[-\frac{(y-\mu_y)^2}{2\sigma_y^2}\right]\mathrm{d}y\mathrm{d}n\right], \tag{9}$$

where t is the threshold level for the photocurrent. We can find a closed form solution for Eq. (9) for the special case (which is incidentally the worst case) in which the threshold t is set at the mean of the input distribution μ_y. In this case the probability of error is

$$P_e = \frac{1}{\pi}\tan^{-1}\frac{\sigma_n}{\sigma_y}. \tag{10}$$

P_e is specified by algorithmic considerations, whereas σ_n is affected by the device fabrication process. Given P_e and σ_n, we can either solve Eq. (9) or use Eq. (10) to determine numerically the ncessary σ_y. From Eq. (8) we see that with η_0, η_D, p, and q given, we adjust σ_y so that it satisfies the required probability of error P_e by selecting the appropriate value for the quotient $P/\sqrt{C}$. The variables N and P are also constrained by the maximum permissable electrical power dissipation per unit area $P_{\max}$:

$$P_{\max}A = NPV_D/\eta_{\mathrm{LED}}, \tag{11}$$

where V_D is the power supply voltage in the LED circuit, A is the area of the array, and η_{LED} is the LED efficiency in watts per ampere. From Eqs. (8), (10), and (11) the maximum density of neuron circuits N/A is obtained by

$$N/A = \left\{\frac{\eta_0\eta_D\eta_{\mathrm{LED}}P_{\max}}{V\sigma_n\sqrt{C}}[pq(1-pq)]^{1/2}\tan(\pi P_e)\right\}. \tag{12}$$

For a fully connected feed-forward network, each neuron is connected to all neurons in the previous layer, i.e., $C = N$. In this case

$$N/A = \left\{\frac{\eta_0\eta_D\eta_{\mathrm{LED}}P_{\max}}{V\sigma_n\sqrt{A}}[pq(1-pq)]^{1/2}\tan(\pi P_e)\right\}^{2/3}. \tag{13}$$

The above equations let us determine the density of neurons that can be fabricated with LED circuits, given the device limitations (σ_n, $P_{\max}$, V_D, A, η_D, η_0, and η_{LED}) and the algorithmic specifications (P_e, p, and q). As an example, suppose that we want to build a network with $N = 10^4$ neurons and $C = 10^3$ connections per neuron in an area $A = 1$ cm^2. Suppose further that $\eta_0 = 0.1$, $\eta_{\mathrm{LED}} = 0.01$ W/A, $\eta_D = 0.5$ A/W, $P_{\max} = 10$ W/cm^2, $V = 2$ V, $p = q = 0.5$, and $P_e = 0.1$. Then we must have the capability to fabricate the neuron arrays with sufficient uniformity and low enough noise so that $\sigma_n \approx 1$ nA with a corresponding $P = 5$ μW.

C. Speed Considerations

The analysis in Subsection 2.B leads us to the conclusion that we must use as low an optical power as possible to maximize the density of neurons, in which the minimum acceptable power is determined by the noise in the circuit. The price that we must pay for the higher density is slower speed. This trade-off manifests itself in two ways. First, the time required to charge the various components of the circuit (the detectors and the gate of the MESFET) is inversely proportional to the available photocurrent.[15] Second, since the detector noise is proportional to the square root of the bandwidth and since we use small signal levels, we are forced to use slow speeds to reduce the noise to satisfy Eq. (9). We made preliminary temporal measurements, which indicate that the switching time of these devices is of the order of 10 μs for input optical powers that are consistent with the density requirement obtained by Eq. (11).

3. Optoelectronic Circuit

In this section we describe the fabrication and the performance of three different implementations of the circuit shown in Fig. 1. We begin with a description of the LED and MESFET fabrication and design before describing the complete circuit.

A. Light-Emitting-Diode Fabrication

The light source in our optoelectronic neurons is implemented with double Zn-diffusion LED's; the LED cross-sectional structure is shown in Fig. 3. The epitaxial layers consist of a double-heterojunction n-AlGaAs/p-GaAs/n-AlGaAs sandwiched between two n^+-GaAs contact layers. The p-n diode required for the LED is achieved by duffusing Zn selectively through the top n-AlGaAs layer to convert it from n-type to p-type. To improve the current

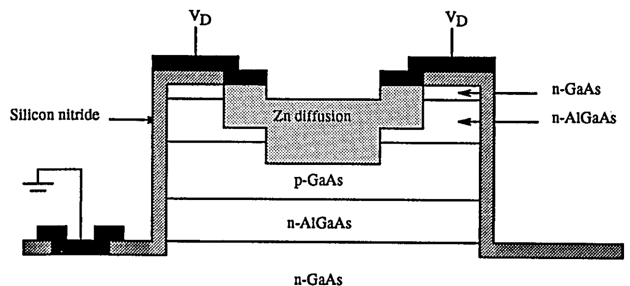

Fig. 3. Cross section of the LED structure, employing double Zn diffusion.

confinement so that carrier recombination takes place underneath the LED emitting window, a double-diffusion technique is employed. LED's fabricated by this technique have the advantage of separating spatially the top electrodes from the carrier recombination region, thus permitting a higher external efficiency for the LED's. This improvement is especially important when one considers that the critical angle for the photons impinging on the GaAs/air interface is only 16.7°, which accounts for most of the loss of photons. Despite this improvement, most of the generated photons suffer a non-negligible absorption loss because the minority carrier injection takes place at the bottom p-GaAs/n-AlGaAs heterojunction. This absorption can be minimized by decreasing the GaAs-layer thickness. However, interfacial recombination, which contributes to nonradiative recombination and reduces the overall quantum efficiency, becomes significant if the active-layer thickness is too small. Thus there is an optimal thickness for the active layer. Figure 4 is a theoretical plot of the dependence of external efficiency on the LED active-layer thickness for typical absorption lengths and interfacial recombination velocities.[16] The two curves show the improvement that can be obtained by adding an antireflection coating to the LED window.

The fabrication of the double-diffusion LED begins by first isolating each LED electrically using a conventional wet etchant that stops in the bottom n^+-GaAs layer, followed by a blanket deposition of Si_3N_4, which serves as the diffusion mask for Zn. A window for the Zn diffusion is defined by opening the nitride over the LED mesa by using a CF_4 dry-etching technique and removing the top n^+-GaAs layer to minimize absorption. Zn diffusion is performed by placing the sample and the $ZnAs_2$ in a quartz ampoule, which is then sealed under high vacuum. The ampoule is then inserted in a 640°C furnace for 15 min. to permit Zn to diffuse into the p-GaAs active layer. After the first diffusion is completed, a larger nitride window, which encloses the original nitride window defining the first diffusion area, is opened using the same dry-etch technique. The LED sample and the Zn-diffusion source are again sealed in the ampoule under high vacuum. The duration of the second diffusion is calculated such that the diffusion front is properly placed in the top n-AlGaAs layer. It is extremely critical that during this step the second diffusion front does not extend into the p-GaAs layer, nor does it stop in the top n^+-GaAs layer so that current can be channeled spatially through the middle of the LED. After the double Zn-diffusion process is completed, n- and p-type ohmic contacts are defined by first removing the nitride over the unexposed contact area and then by evaporating AuGe/Ni/Au and Ti/Au, respectively, to metalize the contacts.

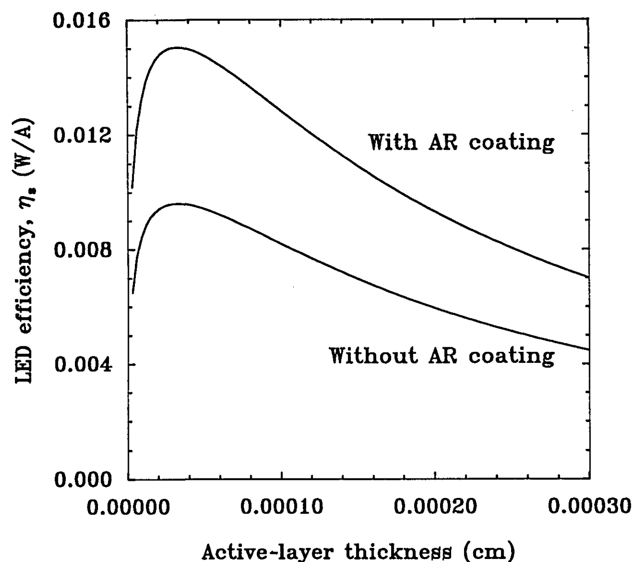

Fig. 4. LED external efficiency both with and without antireflection coating as a function of the active-layer thickness with a nonzero absorption coefficient and an interfacial recombination velocity.

The external efficiency of the LED fabricated by the double Zn-diffusion process was measured to be 0.01 W/A. This is a 55% improvement over the efficiency of LED's with a single Zn-diffusion process (Ref. 15, p. 73). This improvement indicates the effectiveness of channeling the current through the middle of the LED, thus permitting more photons to be emitted into the air.

B. Fabrication of Metal–Semiconductor Field-Effect Transistors

Figure 5 shows the epilayers for the recessed gate MESFET used in our circuits. The top n^+-GaAs layer is added to make ohmic contacts and to reduce the parasitic gate–source resistance. The channel region is etched through the n^+ layer to the n^- layer below it until the MESFET becomes an enhancement mode transistor; i.e., the channel is pinched off with 0 V on the gate and conducting when the gate voltage becomes positive. To fabricate the MESFET, we first deposit a uniform layer (10^2 nm) of Si_3N_4, and then we open two windows in the silicon nitride with a CF_4 plasma etcher for the source and the drain ohmic contact regions. AuGe/Ni/Au metals are evaporated onto the wafer for the source and the drain contacts using a standard lift-off technique for patterning. The ohmic contacts are alloyed at 430°C for 4 min. The next step defines the gate-recess region by opening a window in the silicon nitride. Using the silicon nitride as a mask, we use a wet-chemical etchant, consisting of NH_4OH, H_2O_2, and H_2O, to remove the n^+ layer in the gate region and the

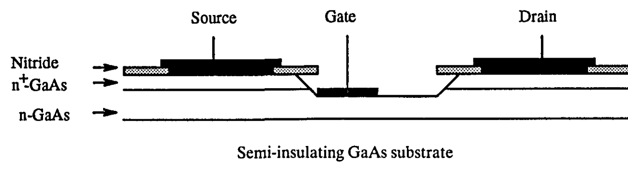

Fig. 5. Self-aligned and passivated MESFET with a recessed asymmetric gate.

recess of the n-channel. The exact etch depth at which the channel is pinched off at zero gate bias is determined by measuring periodically the source–drain current while etching. The final step is to evaporate Ti/Pt/Au to form the gate contact in a self-aligned manner with respect to the source. The dimensions of the gate are 7 μm × 100 μm.

Figure 6 shows the *I–V* curve of a MESFET in series with a double-diffusion LED, fabricated as described above. A transconductance of 30 mS/mm and a source–drain breakdown voltage of 4 V were measured. The initial offset in the drain–source voltage is due to the turn-on voltage of the LED. These results were consistent with expectations except for the relatively low breakdown voltage of 4 V. This was probably caused by surface-induced instead of bulk-induced breakdown, and perhaps it was a result of dirt particles in the vicinity of or underneath the gate. The transconductance of 30 mS/mm corresponds to a 1 mA current swing for a 0.3-V swing in the gate voltage. For the neuron circuits that we are building the required LED current is less than 1 mA per element, and detector circuits can be designed to produce a 0.3-V swing. Therefore a single MESFET is sufficient to drive the LED. Moreover, the transconductance of the MESFET can be increased easily by reducing the gate length and by increasing the width, if needed. For this reason, MESFET's are excellent candidates as LED drivers. When we consider a larger two-dimensional array, gain nonuniformity from device to device becomes an important consideration. The material structure for the complete circuit has the epilayers for the LED on top of the MESFET structure described above. Therefore we need to etch through 2.55 μm of material just to reach the n^+ layer of the MESFET. This deep etch introduces nonuniformity in the MESFET channel thickness, which in turn produces gain variations. This can be improved by controlling the etch more carefully or by inserting an AlAs stop-etch epilayer above the MESFET structure. With this layer, selective etchants that do not etch AlAs but do etch GaAs and vice versa can be used to define the top of MESFET accurately.[17]

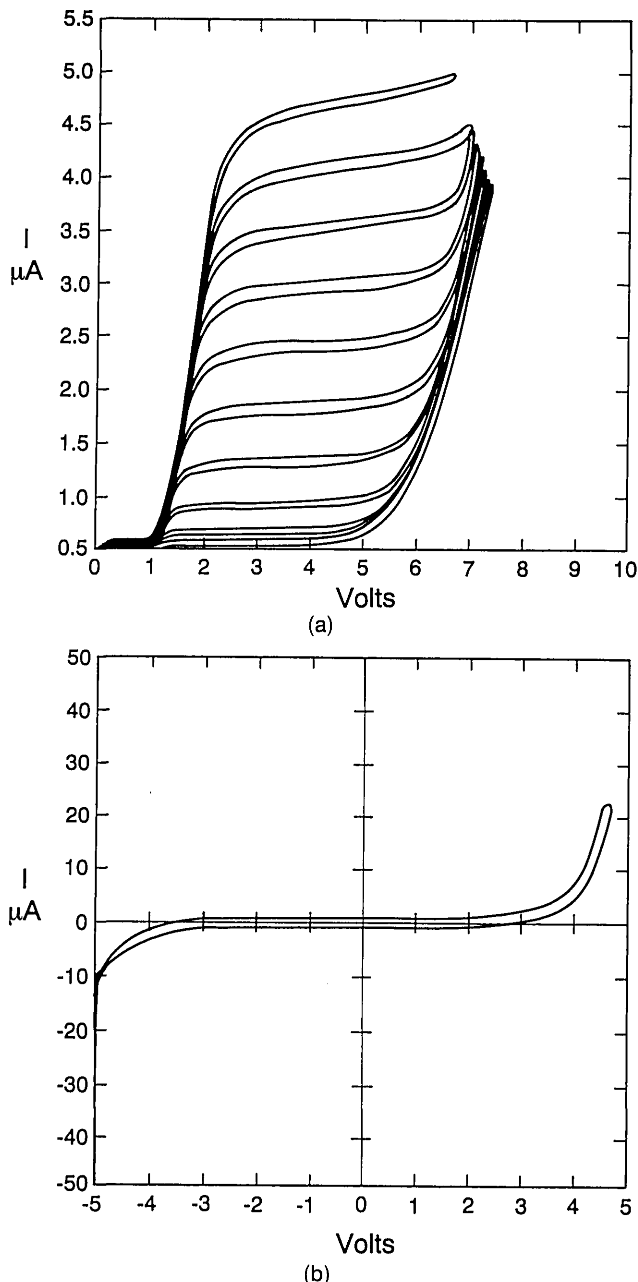

Fig. 6. (a) Common-source *I–V* characteristics of a MESFET showning a breakdown voltage of ~4 V. The initial turn-on voltage of 1 V is due to the LED, which is in series with the MESFET. The scales are 500 μA/division vertically and 1 V/division horizontally. (b) Reverse breakdown characteristics of the gate–drain Schottky diode (first quadrant) and the gate–source Schottky diode (third quadrant). The scales are 10 μA/division vertically and 1 V/division horizontally.

C. Switching Characteristics of the Neuron Circuit

The basic circuit for implementing a thresholding function with optical inputs and optical outputs is shown in Fig. 1. The gate voltage V_g, on the driving MESFET is controlled by the input circuit, which consists of a photodetector acting as an optical input port and a biasing element, which can be either a photodetector or a transistor. The switching characteristics of the circuit are determined by the *I–V* characteristics of the photodetector and the biasing element. Figure 7(a) shows the load-line curves for three different light intensities illuminating the photodetector and two different bias levels. The intersection point on the *I–V* curves for the photodetector and the biasing element determines the node voltage V_g. When the current in the photodetector exceeds the threshold current set by the biasing element, the gate voltage switches from ground to V_{dd}, which, in turn, switches on the driving MESFET. This causes current to flow through the LED and the output light intensity to increase to its high value. V_g as a function of input light intensity is shown in Fig. 7(b).

Since the current through the LED is roughly proportional to V_g, the nonlinear input–output relationship is determined almost exclusively by the

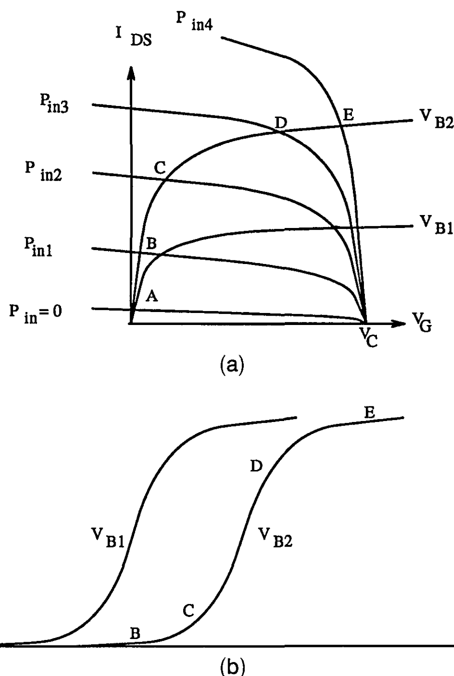

Fig. 7. (a) *I–V* characteristics of the input switching circuit in the MESFET-based neurons with a different biasing voltage applied to the gate of the loading MESFET, (b) the gate voltage on the driving MESFET as a function of the optical input, where I_{DS} is the current through the threshold MESFET or the photodetector, V_{B1} and V_{B2} are two different gate voltages on the threshold MESFET, and A–E in (a) and (b) indicate identical operating points.

input circuit. The sharpness of the threshold function shown in Fig. 7(b) is determined by the relative flatness of the *I–V* curves of the two devices in the saturation regime. The threshold becomes sharper as the output impedance of the devices in the saturation regime becomes larger. Specifically, the transition region of the threshold function is

$$\Delta V_g \approx \frac{R_D R_B}{R_D + R_B} \eta_D \Delta P_{\mathrm{in}}, \tag{14}$$

where ΔV_g (up to 0.5 V) is the voltage swing on the gate of the driving MESFET and R_D and R_B are the output impedances of the photodetector and the biasing element, respectively. The leakage current through the gate of the driving MESFET also affects the switching characteristics of the circuit. The extra current that is drawn through the gate needs to be supplied by extra photocurrent, which tends to increase the optical threshold level. The gate–source leakage current that we measure in our MESFET's (using Ti/Pt/Au Schottky contacts) is 10–100 nA for gate voltages up to 0.5 V. From our discussion in Section 2, noise considerations dictate that the minimum workable photocurrent be several microamperes. This becomes the limiting factor for circuits with low detector efficiency and high leakage currents.

In most neural-network implementations the neuron outputs or the weights are bipolar values. Since the LED's are incoherent-light sources, only positive values can be implemented directly with these circuits. In most cases it is possible to work with unipolar neuron activation functions. But it is necessary to have bipolar weights.[18,12] There are two ways to represent bipolar signals in an incoherent system. The first method is to add a constant bias to all the bipolar weights before they are recorded in the optical system. In this case the input signal to each neuron is a positive quantity with the desired bipolar signal riding on a bias. In our circuits the control signal (either optical or electronic) on the biasing element adjusts the threshold of the circuit. The second method for representing bipolar signals consists of separating spatially the recorded positive and negative weights. The inner product between the input vector to the neuron and each of the two sets of weights is formed separately, and the results are subtracted electronically before thresholding. The circuits we describe here can be used in this mode if the biasing element is also a photodetector, identical to the input detector. The positive signal P^+ is routed to the signal port, and the negative signal P^- is routed to the biasing port. The gate voltage V_g saturates at V_D (or ground) as $P^+ - P^-$ gets large and positive (or negative). When $P^+ = P^-$, $V_g = V_D/2$. Therefore the circuit implements a sigmoidal function on the difference $P^+ - P^-$, as desired.

In the following we describe three different monolithically integrated neuron circuits based on the LED as the optical output and a MESFET as the amplifying transistor. The circuits differ in the type of photodetector used. The three photodetectors that we investigate are a bipolar phototransistor, a metal–semiconductor–metal (MSM) detector, and an optical field-effect transistor (FET).

1. *Phototransistor-Based Neuron Circuit*

The phototransistor is a bipolar junction transistor without an electrical base contact.[19] The structure of the phototransistor incorporated in our circuits is shown in Fig. 8. It consists of a lightly p-doped GaAs base layer sandwiched between two higher-band-gap n-doped AlGaAs layers, which are the emitter and the collector of the transistor. Light incident upon the phototransistor window first traverses the transparent higher band-gap emitter region and then is absorbed in the lower band-gap base layer. The carriers generated in the base modify the base–emitter junction potential such that more electrons are injected from the emitter to produce a large collector current. The overall efficiency of the device is $\eta_D = \eta' \beta$, where η' is the quantum efficiency, or the fraction of electrons generated from each incoming photon in the base, and β is the transistor common-emitter current gain.[20] The base-layer thickness is a critical parameter in determining the overall detector efficiency. If the base layer is thin, the current gain β is large, but the amount of light absorbed is small. Likewise, if the base is thick, the amount of light absorbed is high, but the current gain is low. By lowering the doping concentration of the base layer,

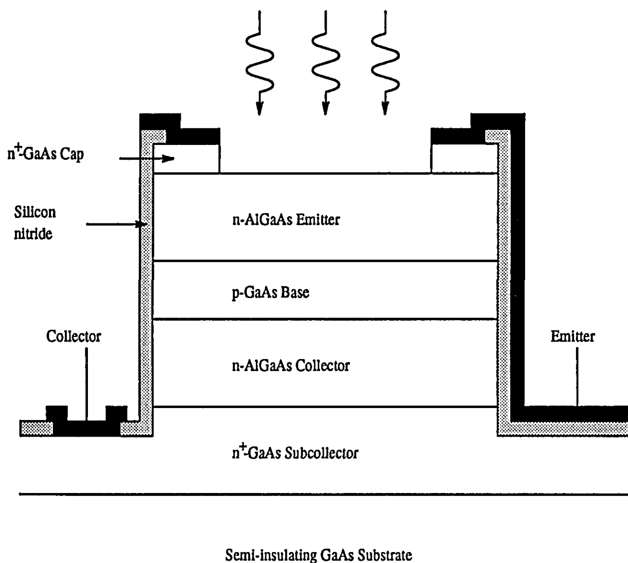

Fig. 8. Structure of a double-heterojunction phototransistor incorporating a p-doped GaAs layer as the base.

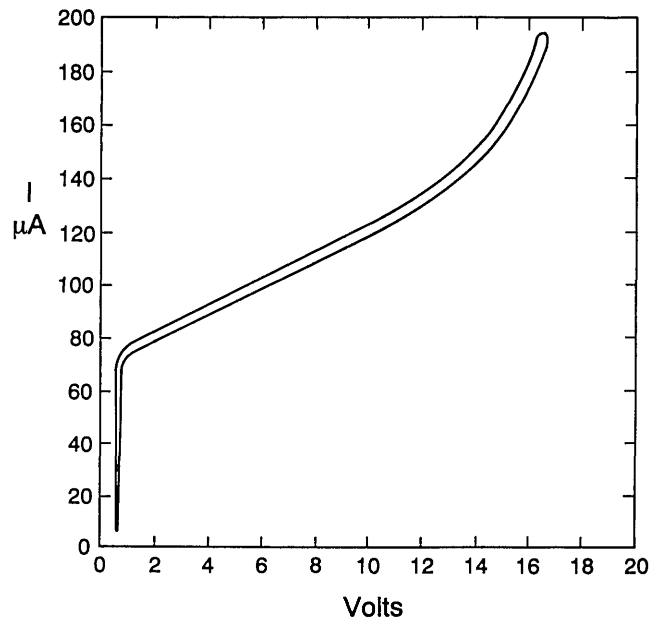

Fig. 9. *I–V* characteristics of the phototransistor. The intensity of the incoming laser beam is 90 μW. The scales for the vertical and horizontal axes are 20 μA/division and 2 V/division, respectively.

we can increase the thickness without decreasing the current gain significantly. However, one drawback of low base doping is that the early voltage is low, which corresponds to low output impedance. This decreases the sharpness of the switching curve.

The base layer was designed to be 1.5 μm thick and doped p-type to a concentration of 10^{16} cm^{-3}. The n^+-GaAs layers act as emitter and collector ohmic contacting layers. The side wall of the phototransistor was passivated with a Si_3N_4 dielectric film. In the window area for detecting the incoming photons the absorbing n^+ ohmic cap layer was removed by wet-chemical etching. The *I–V* characteristics of the phototransistor were obtained by illuminating the device with a beam from a GaAs laser diode, with known intensity, and by measuring the emitter–collector current. The result is shown in Fig. 9. The relatively low output impedance (175 kΩ) is due to the low base doping. The external quantum efficiency η_D was approximately 1 A/W, measured for a collector–emitter voltage of 4 V and a 90-μW incident optical power. The relatively low gain is because of the thick base layer. In the integrated circuit the base layer of the phototransistor is shared with the active layer of the LED. Thus the designed thickness for the base was a compromise between the efficiencies of the phototransistor and the LED.

Figure 10(a) shows the circuit diagram for the neuron circuit incorporating a bipolar phototransistor as the detector, a MESFET to set the threshold for the circuit, the LED-driving MESFET, and the LED. Figure 10(b) is the cross-sectional view of the monolithically integrated device.[21] The undoped layer on top of the semi-insulating substrate acts as a buffer. The MESFET's are fabricated in the two n-GaAs layers above the buffer layer, with the n^+-GaAs layer used for source and drain contacts, and the n^--GaAs layer used for the MESFET channel. The MESFET fabrication is described in Subsection 3.B. The N-p-N heterostructure that is grown above the MESFET layers is used to fabricate both the phototransistor and the LED. As described in Subsection 3.A, the LED is formed by a double Zn-diffusion process to convert selectively the top n^+-GaAs and n-AlGaAs layers to p-type. The processing of the complete integrated device involves nine masking steps. The first two steps are wet-chemical etches to define and isolate the individual components, followed by two $ZnAs_2$ diffusions for the LED and the metalizations for the ohmic and Schottky contacts. The connection between the LED cathode and the MESFET drain is made through the shared n^+-GaAs, whereas metalization was used for all other connections. Figure 11 is a photomicrograph of a single phototransistor-based neuron circuit.

The input–output relationship of the neuron, shown in Fig. 12, was obtained by measuring the LED current as a function of the laser power incident upon the phototransistor. The gate on the loading MESFET was left floating to reduce the gate–drain leakage current in the loading MESFET. The measured LED current was converted to optical output power by using the LED efficiency of 0.01 W/A, determined from measurements on the double-diffusion heterojunction LED experiments. Measuring the current through the LED rather than the emitted optical power eased the practical problems of simultaneously detecting the output optical signal and illumating the optical input detector. The optical threshold level was 1 μW, and a $\Delta P_{in} = 0.2$ μW increase in the input power resulted in an 8-μW increase in the optical output power. The LED

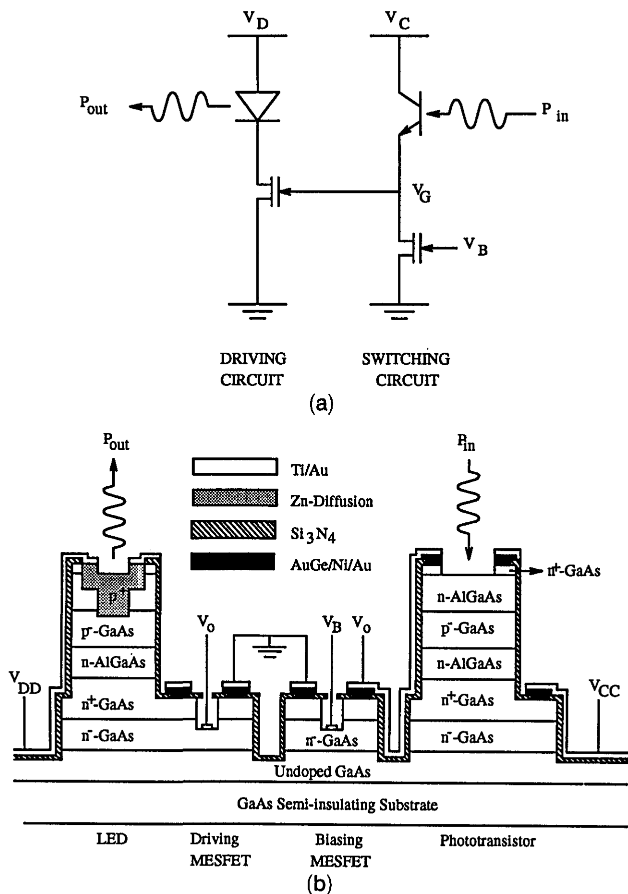

Fig. 10. (a) Schematic circuit diagram of the optoelectronic neuron that incorporates two MESFET's, a phototransistor, and an LED; (b) the cross-sectional view of the MESFET-based optoelectronic neuron monolithically integrating two MESFET's, a LED, and a phototransistor.

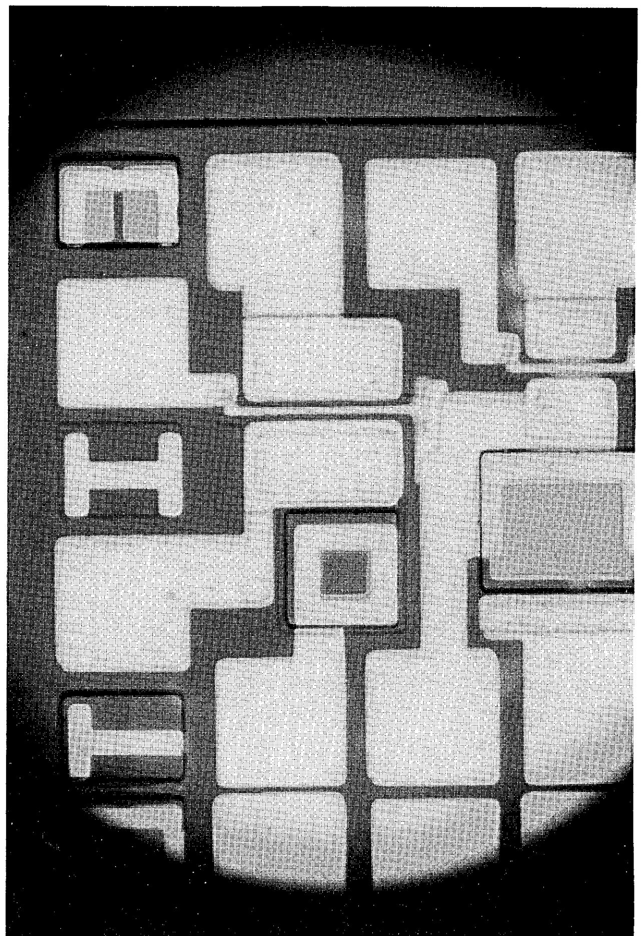

Fig. 11. Photomicrograph of a completely fabricated optoelectronic neuron. The input switching circuit is on the right side of the picture, and the output driving circuit is on the left side of the picture. The lower-left square is the LED, which is monolithically connected to the drain of the driving MESFET. The gate of the same MESFET is controlled by the combination of the phototransistor located at the lower-right corner of the picture and the loading MESFET, which is located immediately above the phototransistor. The windows of the LED and the phototransistor are 40 μm $\times$ 40 μm and 60 μm $\times$ 80 μm, respectively.

current during the on state of the neuron was measured to be 0.8 mA. Therefore the electrical power dissipation per neuron was 1.6 mW with a 2-V power supply. The switching time of the neuron was measured by applying an electrical square-wave signal to the laser diode. A rise time of 65 μs was obtained when the neuron was illuminated with a laser pulse of intensity slightly higher than 1 μW.

2. *Metal–Semiconductor–Metal-Detector-Based Neuron Circuit*

The disadvantage of the phototransistor is that it is fabricated in the same layers as the LED, thus making it difficult to optimize the circuit. We can overcome this difficulty by using MSM photodetectors, which we can fabricate in the undoped buffer layer.[22] The MSM's are fabricated by depositing Ti/Pt Au metal Schottky contacts on the undoped GaAs. This forms two Schottky diodes back to back. Optically generated electron-hole pairs in the depletion region of the reverse-biased Schottky diode are collected at the electrodes. Therefore the MSM detector operation is similar to that of a p-i-n diode. As is the case for the p-i-n diode, the external efficiency of the MSM detector cannot be larger than 100%. However, the MSM detector has an advantage in that the only epilayer required is the buffer layer, which is not shared by any other devices in the neuron. The electrodes are patterned as interdigitated fingers, 4 μm wide and 6 μm apart with an active area of 40 μm $\times$ 40 μm. With 3 V applied to the detector the measured external efficiency was 0.3 A/W.

Figure 13(a) shows the circuit diagram with the two MSM photodetectors, one for the optical input signal and the other for the optically controlled biasing element. The cross section of the processed epilayers is shown in Fig. 13(b). To fabricate the circuit we first defined the LED, the MESFET, and the MSM detectors through a series of wet-chemical etches. Then a layer of Si_3N_4 was deposited for surface passivation and insulation. The Si_3N_4 was removed selectively for the MESFET channel region, the LED window, and the ohmic contacts. AuGe/Ni/Au lay-

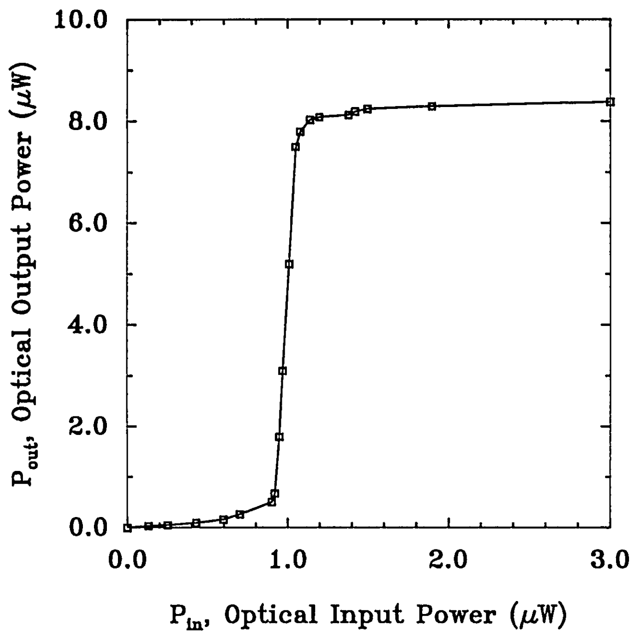

Fig. 12. Input–output relationship of the phototransistor–MESFET-based optoelectronic neuron.

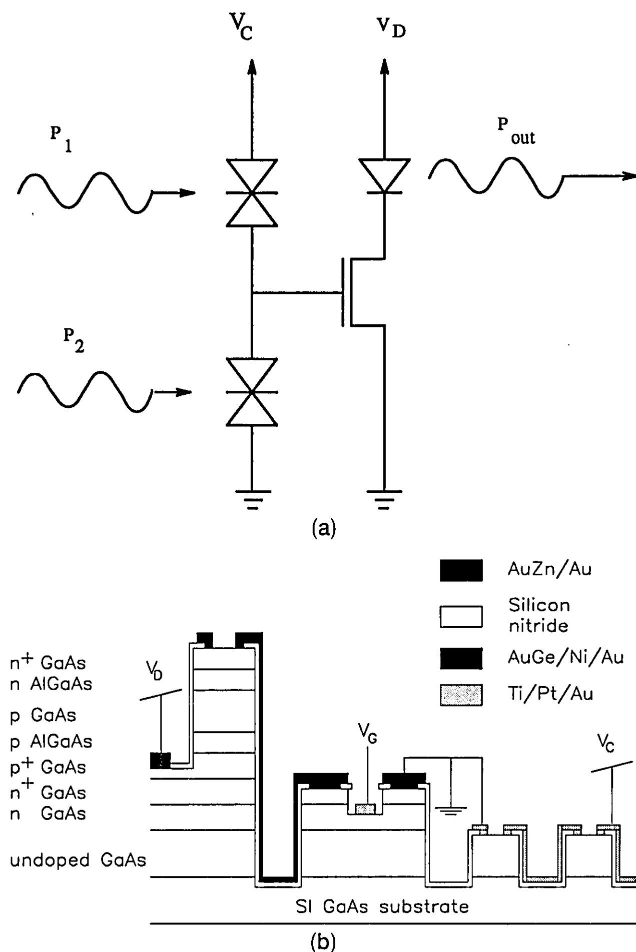

Fig. 13. (a) Circuit diagram for the MSM-based neuron circuit, (b) the cross section of the processed epilayers for the monolithic integration of the MSM-based neuron circuit.

ers were deposited for the n-ohmic contact, followed by a deposition of AuZn/Au for the p-ohmic contact. After deposition, the contacts were alloyed at 430°C. The gate region of the driving MESFET was recessed as described in Subsection 3.B so that the MESFET operated in the enhancement mode. The entire fabrication process requires nine masking steps. The LED in this MSM-detector-based neuron is fabricated directly on a double-heterojunction P-i-N structure without the double-diffusion process. The maximum efficiency obtained was 0.001 W/A owing to the lack of current confinement. Therefore we opted for the double-diffusion process in all subsequent LED devices. The size of the gate region was 6 μm × 60 μm, the LED window was 30 μm × 30 μm, and the overall neuron area was 200 μm × 150 μm, which included the area for probe pads.

To test the circuit, we measured the output as a function of one of the optical inputs while we held the other one constant. Figure 14 shows the results when P_1 and P_2 are the optical signals, as shown in Fig. 13(a). When the two optical input powers are equal, the output light levels switches. The differential optical input power P_{in} required to turn the LED on was 0.2 μW. The time response was not measured for the MSM neurons. But we think the time response is comparable with that of the phototransistor-based neuron. This is because the detector efficiencies are roughly the same, and the response of the neurons in both circuits is due mainly to the gate capacitance charging in the MESFET's that drive the LED's.

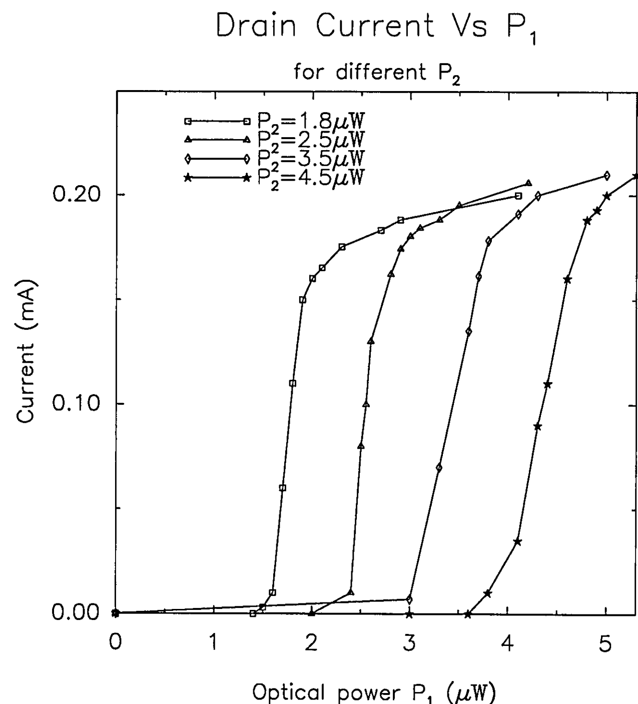

Fig. 14. Output current as a function of input optical intensity for different optically set thresholds.

3. Optical Field-Effect-Transistor-Based Neuron Circuits

The two detectors considered in Subsections 3.C.1 and 3.C.2 do not have gain. Thus the neurons require relatively high optical input intensities (approximately 1 μW). In order to increase the density, we need to be able to reduce the optical input light level. This can be accomplished by using a detector, such as the optical FET, that has gain. The optical FET can be fabricated easily in the same epilayers as the driving MESFET in the circuit, and as described below, the efficiency for such a device can be as large as 10^5 A/W.[23]

The mechanism of operation of an optical FET is based on photoconductivity.[24] Electron-hole pairs generated by the incident optical signal increase the conductivity of the channel. If the transit time for the electrons to cross the region is short compared with the lifetime of the holes, then the detector efficiency can be much greater than 1 A/W. The expression for drain–source current is

$$I = \frac{\tau_h}{\tau_t} \eta' P_{\text{in}}, \tag{15}$$

where τ_h is the hole lifetime, τ_t is the electron transit time, and η' is the quantum efficiency. The gain of the optical FET detector is maximized if the gap between the source and the drain is small. However, there is a trade-off since this same gap is the optical window of the device, and we want this to be large enough to permit sufficient input light to be detected.

Figure 15 shows the cross-sectional structure of an optical FET. Since it is similar to the conventional MESFET, its fabrication is identical to that of a MESFET, except that the gate metalization is not defined. The channel thickness, determined by the recessed etch, controls the sensitivity and the dark current. As the photoconducting channel gets thinner by etching, the photoconductivity starts to decreasing owing to the smaller absorption region. As a result, the optical gain decreases gradually. The dark current is proportional to the cross-sectional area of the channel, and therefore it also decreases as the channel is etched. Eventually, only the surface depletion region remains. Figure 16 shows the measured efficiency as a function of the input optical power for four different dark currents, which correspond to four different channel thicknesses.

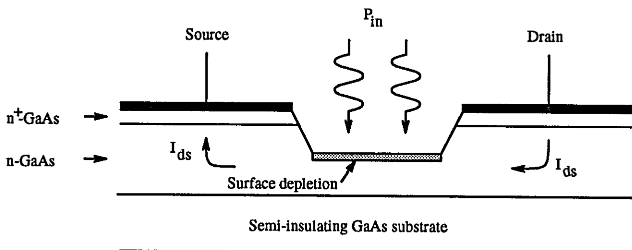

Fig. 15. Cross-sectional view of an optical FET, where I_{ds} is the drain–source photocurrent.

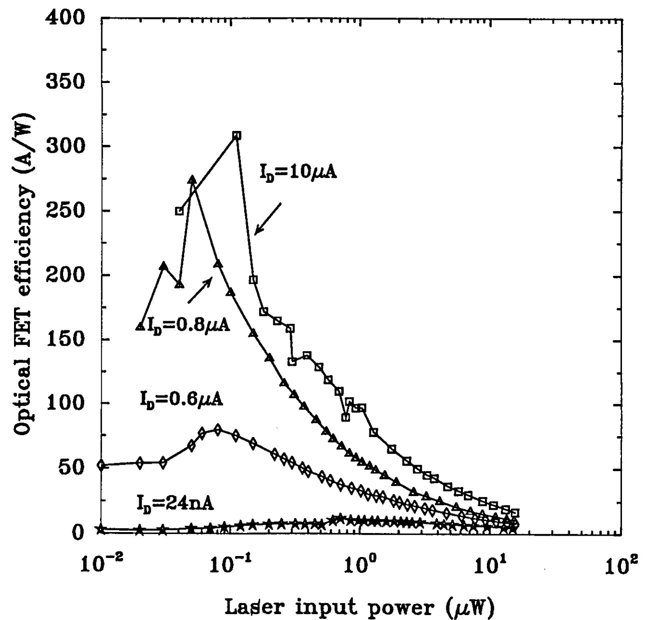

Fig. 16. Efficiency of the optical FET as a function of input light intensity. The four curves correspond to four different dark currents, which correspond to the four different recessed depths.

The fabrication steps incorporating the optical FET as the detector and a MESFET as the biasing element for this particular neuron circuit are similar to those of the circuits described above. The only difference is in the definition of the optical FET. The complete fabrication process requires nine masking steps, as shown in the cross-sectional view of the circuit in Fig. 17. Even though the MESFET's and the optical FET's share the same epilayers, the recess etch for each device was performed separately to ensure that the MESFET was correctly pinched off and that the optical FET had the proper dark current. The circuit used the double-diffusion heterojunction LED as the optical output. In testing the neuron circuit, we left the gate of the loading MESFET in the input circuit floating to minimize the gate leakage current. The optical input–output characteristics are shown in Fig. 18. Because of the insufficient recess in the gate of the output driving MESFET, this MESFET was not pinched off at zero gate bias. As a result, a current flowed in the output circuit with zero input power on the detector. This caused a nonzero LED output power at zero input power. In a properly fabricated device the channel of the LED-driving MESFET would have been etched further until the current was close to zero at zero gate bias. Nevertheless, the output rose by 4.3 μW over an input power swing of 54 nW. The measured rise time of the neuron was 700 μs. The minimum input power needed for turning the neuron on dropped significantly, from the 1-μW regime down to the 1-nW regime. This is attributed to the higher efficiency of the detector. During the on state of the neuron the total current drawn by the LED was 0.9 mA, which implied an electrical power dissipation of 1.8 mW/neuron with a 2-V power supply.

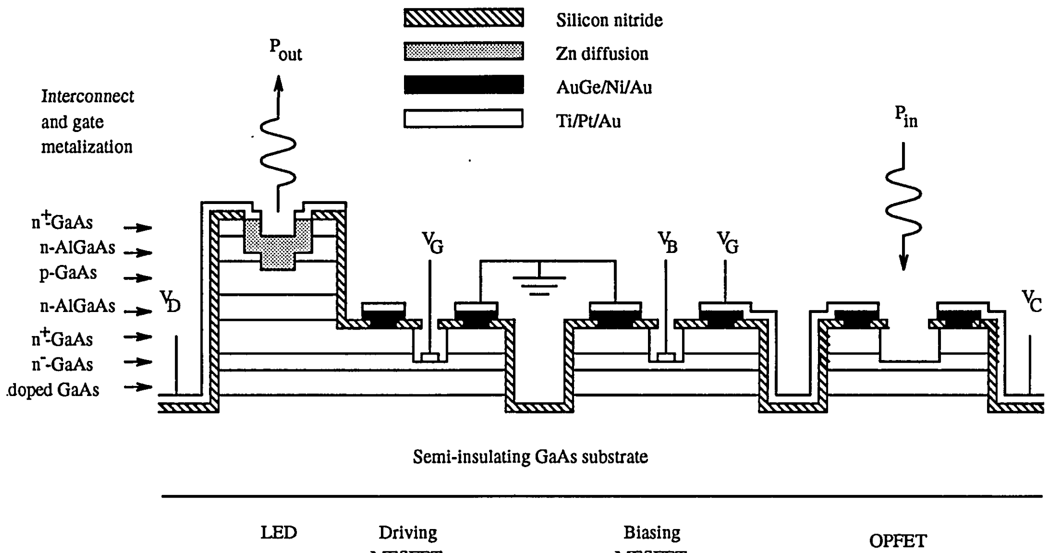

Fig. 17. Complete device cross-sectional view of the MESFET-based optoelectronic neuron incorporating the optical FET (OPFET) as the detector.

4. Bipolar Transistor-Based Neurons

Optoelectronic neuron circuits can also be fabricated with bipolar transistors. Bipolar transistors are amenable to this type of integration[25] because their material structure is similar to that of the LED and the phototransistor. Since the LED requires a double heterostructure for carrier confinement, the same material can be used to fabricate both double-heterojunction bipolar transistors and phototransistors. The requirements on the types of dopants for both devices can be simultanteously met by having the initial structure consist of a standard N-p-N DHBT structure and then by converting the top n-type high band-gap AlGaAs layer to p-type through Zn diffusion to form the LED structure. However, there is a performance trade-off between the LED and the transistors as a result of sharing the same epitaxial layers required for monolithic integration. This compromise arises because the p-GaAs is shared by all three devices.[26] For the photodetector this

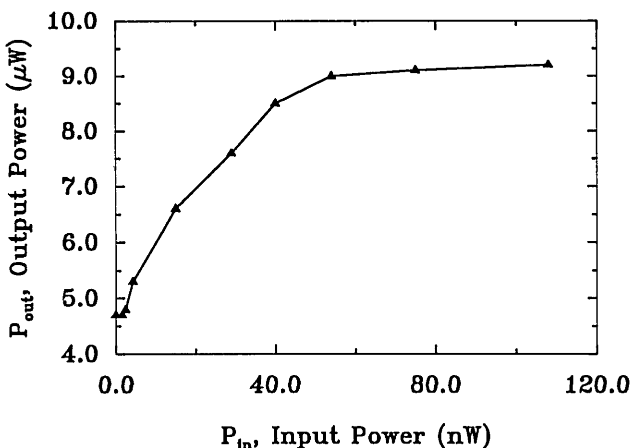

Fig. 18. Input–output characteristics of the MESFET-based optoelectronic neuron that incorporates the optical FET.

p-GaAs is the absorption layer, which needs to be thick to permit the complete absorption of incoming photons. However, for the transistor this layer is the base, which should be as thin as possible to maximize the current gain. Similarly, the LED cannot be too thick or too thin, as the self-absorption or the interfacial recombination will dominate the process. Thus the thickness of the p-GaAs layer needs to be chosen carefully so that the overall performance of the neuron, not of each device, is maximized.

Figure 19(a) shows the neuron thresholding circuit implemented in a Darlington-pair bipolar transistor configuration. The threshold is provided by applying a reverse-biased current I_{bb} on the base of the transistor such that the transistor does not turn on until the photogenerated current has exceeded the reverse-biased current. As the photocurrent is further increased, the transistors amplify the signal received to produce an output current that drives the LED. This process continues until the transistors saturate, which causes the neuron to saturate as well. Figure 19(b) shows the cross section of the epitaxial layers shared by the three devices. Isolation between devices is achieved by etching down to the semi-insulating substrate. For the integrated transistor, Zn diffusion is used to contact the base of the transistor. This also serves as a necessary step in converting the n-AlGaAs upper-cladding layer to p-AlGaAs, thus forming a P-i-N diode for the LED, as discussed in Subsection 3.A.

The current gain of the circuit has to be large enough to compensate for all losses, including those in the hologram, the detector, and the LED. In order to cascade layers without any attenuation, the following relationship is mandated:

$$(\eta_H)(\eta_D)(\eta_L)(\beta) \geq 1, \qquad (16)$$

where η_H is the efficiency of the hologram that

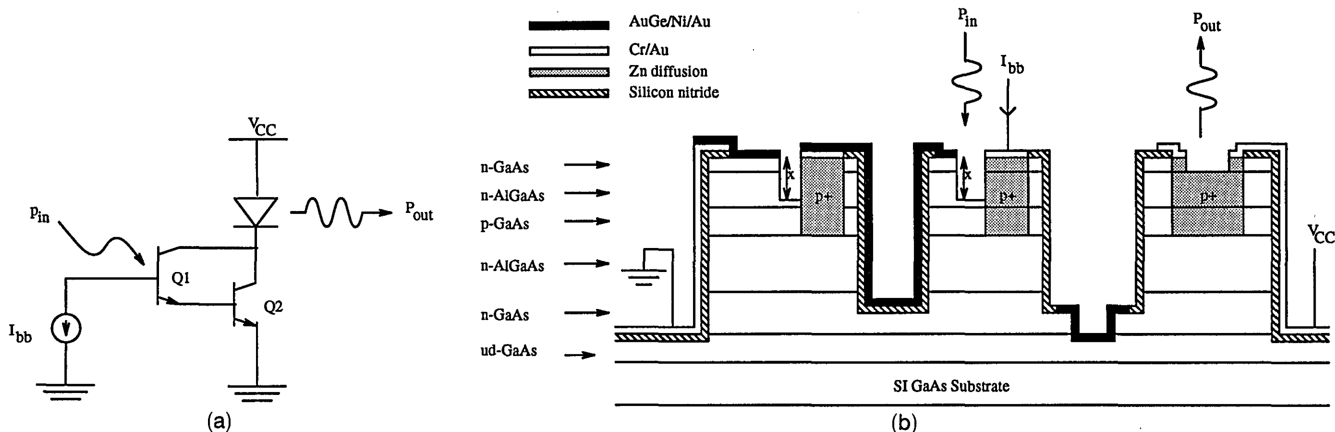

Fig. 19. (a) Schematic circuit diagram of an optoelectronic neuron incorporating two bipolar transistors (Q1 and Q2) to provide the gain needed to satisfy the loop gain requirement; (b) cross-sectional view of the monolithically integrated optoelectronic neuron that consists of two Zn-diffusion double-heterojunction bipolar transistors, which form a Darlington transistor pair, and an LED.

specifies the interconnection, η_D is the detector efficiency, and η_L is the efficiency of the LED. For $\eta_H = 0.1$, $\eta_D = 0.3$ A/W, and $\eta_L = 0.01$ W/A, the current gain β has to be at least 3333. A logarithmic plot of β for a single double-heterojunction bipolar transistor that we fabricated, as a function of the collector current, is shown in Fig. 20 for two different collector–emitter voltages. The maximum current gain obtained is in excess of 500, but it is obtained at $I_C = 10^2$ mA. The dependency of β on the collector current is due to the recombination in the depletion region of the transistor. Approximately, the relationship can be expressed as

$$\beta \sim I_c^{1-(1/n)}, \tag{17}$$

where n is the ideality factor for the base–emitter junction, which ranges from 1 for the ideal junction to 2 for the nonideal junction.[28] The case of $n = 2$ corresponds to the situation in which the base current is dominated by recombination taking place through deep-level traps in the space-charge region. The dependency of β on I_C that we measure experimentally conforms to the theoretical prediction with an ideality factor $n \approx 1.3$. With the Darlington pair we can obtain current gains in excess of the required 3000, since its current gain is the product of the gains of the two transistors. The maximum current gain that we were able to obtain experimentally with the Darlington pair was ~ 6000. This gain is marginally sufficient to permit cascading of such neurons. However, the current level at which this gain is obtained is 20 mA. With a 5-V power supply the electrical power dissipation is 10^2 mW/neuron, which severely limits the density of neurons. As a result, the MESFET-based circuits proved more suitable for the high-density low-power low-speed operation that is required for the neural-networks application.

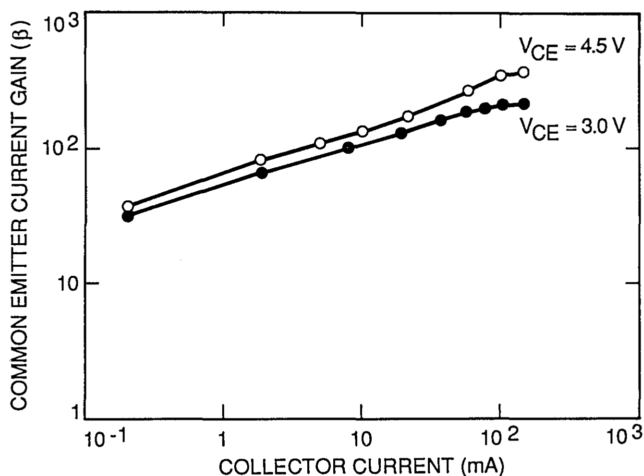

Fig. 20. Common-emitter current gain as a function of the collector current at $V_{CE} = 3$ V and $V_{CE} = 4.5$ V.

5. Discussion

In Section 2 we derived Eq. (12), which obtains the maximum density of neurons that can be built given the parameters of the circuit and the optical system and also the requirements of the system. Above, we examined specific optoelectronic neuron implementations; here, we re-examine Eq. (12) and discuss what would be a practical upper limit for the density with our approach. In the example used in subsection 2.B a density of 10^4 neurons/cm^2 and 10^3 connections per neuron could be achieved if $\eta_{LED} = 0.01$ W/A, $\eta_D = 0.5$ A/W, and $\sigma_n = 1$ nA. We saw in Subsection 3.A that it is possible to fabricate integrated circuits with LED's that have efficiencies equal to 0.01 W/A using the double-diffusion process. We also saw that $\eta_D = 0.5$ A/W was achievable with MSM detectors. Therefore, if the noise level in the MSM circuit is $\sigma_n = 1$ nA, then the density of 10^4 neurons/cm^2 can be achieved. However, we also saw that it is possible to obtain $\eta_D \gg 1$ with the optical FET's. Therefore, if the noise were the same when optical FET's are used, we could build arrays that are much more dense. Unfortunately, the situation is complicated by the fact that σ_n does depend on the choice of detector. The principal contributions to σ_n are thermal noise, shot noise, and nonuniformities in device sensitivity

across the array. If we increase the device density in proportion to the increase in detector efficiency, then the average photocurrent [μ_y in Eq. (8)] remains constant. Therefore the shot-noise component contribution to σ_n stays the same. However, the thermal noise and the variations owing to nonuniformities would be larger for the optical FET circuit. In our current optical FET-based neurons, the dominant source of noise is due to the nonuniformities in sensitivity. This is because we use a wet-etch process to define the channel thickness of the optical FET, which determines the FET's sensitivity. In this case the average noise is proportional simply to the average signal level, $\sigma_n = \alpha\mu_y$, where the proportionality constant α is determined by the uniformity of the etch. Combining this relationship with Eqs. (8) and (10), we can derive the following result:

$$C = \left[\frac{\tan(\pi P_e)}{\alpha}\right]^2 \frac{1 - pq}{pq}, \tag{18}$$

which obtains a direct estimate for the connectivity that can be supported, but not the density of neurons. To determine the density, we need to know the minimum σ_n directly, which in practice is given by the detector noise floor. Suppose that we want to implement an optical FET-based neuron with $N/A = 10^4/\mathrm{cm}^2$, $C = 10^3$, $\eta_D = 500$, and with the rest of the parameters the same as before; then the necessary σ_n is approximately 1 μA. The α necessary to support this connectivity is found from Eq. (18) to be equal to 0.02, which implies that the responsivity of the optical FET's should have 2% uniformity. To obtain an estimate for the uniformity that can be obtained with commercial GaAs processes, we measured the drain–source current of 20 MESFET's with the same gate voltage on a chip fabricated by Vitesse through metal oxide semiconductor implementations service (MOSIS). The average current was 250 μA, and the standard deviation was 15 μA, which corresonds to 6% uniformity for this particular chip. This does not correspond directly to α, but it does give an indication of the uniformity we can expect from the optical FET's. Thus the major challenge in fabricating high-density arrays that can be densely connected is not only minimization of the power consumption, but, equally important, the uniformity with which these arrays can be built.

The support of the Defense Advanced Research Projects Agency for this work is gratefully acknowledged. We thank the Jet Propulsion Laboratory (JPL) for growing some of the material used in our work. Annette Grot gratefully acknowledges support from a National Science Foundation graduate fellowship, and Steven Lin thanks JPL for support from a JPL fellowship. The authors also thank Jae Kim, Joe Katz, and Francis Ho for their contributions in the early phase of this work and Jean-Jacques Drolet for many helpful discussions.

References

1. D. Psaltis and N. H. Farhat, "Optical information processing based on an associative-memory model of neural nets with thresholding and feedback," Opt. Lett. **10,** 98–100 (1985).
2. D. Psaltis, D. Brady, and K. Wagner, "Adaptive optical networks using photorefractive crystals," Appl. Opt. **27,** 1752–1759 (1988).
3. D. Z. Anderson and D. M. Linninger, "Dynamic optical interconnects: volume holograms as optical two-port operators," Appl. Opt. **26,** 5031–5038 (1987).
4. B. H. Soffer, G. J. Dunning, Y. Owechko, and E. Marom, "Associative holographic memory with feedback using phase-conjugate mirrors," Opt. Lett. **11,** 118–120 (1986).
5. A. Yariv, *Optical Electronics,* 3rd ed. (Holt, Reinhart & Winston, New York, 1985), Chap. 15, p. 488.
6. M. Orenstein, A. C. von Lehmen, C. Changhasnian, N. G. Stoffel, J. P. Harbinson, L. T. Florez, E. Clausen, and J. E. Lewell, "Vertical-cavity surface-emitting InGaAs-GaAs-lasers with planar lateral definition," Appl. Phys. Lett. **56,** 2384–2386 (1990).
7. D. A. Miller, "Quantum wells for optical information processing," Opt. Eng. **26,** 368–372 (1987).
8. L. K. Cotter, T. J. Drabik, R. J. Dillon, and M. A. Handschy, "Ferroelectric liquid crystal silicon integrated circuit spatial light modulator," Opt. Lett. **15,** 291–293 (1990).
9. K. Kasahara, T. Numai, H. Kosaka, and I. Ogura, "Vertical to surface transmission electrophotonic device (VSTEP) and its application to optical interconnection and information processing," Inst. Electron. Inform. Commun. Eng. Trans. Fundamentals Elctron. Commun. Comput. Sci. **E75A,** 70–80 (1992).
10. D. Psaltis, D. Brady, X. G. Gu, and S. Lin, "Holography in artifical neural networks," Nature (London), **343,** 325–330 (1990).
11. D. Marr, *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information* (Freeman, New York, 1983), Chap. 2.
12. J. Hopfield, "Neurons with graded response have collective computational properties like those of two-state neurons," Proc. Natl. Acad. Sci. USA **81,** 3088–3092 (1984).
13. N. H. Farhat, D. Psaltis, A. Prata, Jr., and E. G. Paek, "Optical implementation of the Hopfield model," Appl. Opt. **24,** 1469–1475 (1985).
14. A. Papoulis, *Propability, Random Variables, and Stochastic Processes,* 2nd ed. (McGraw-Hill, New York, 1984), p. 194.
15. S. H. Lin, "Optoelectronic integrated circuits for optical neural network applications," Ph.D. dissertation (California Institute of Technology, Pasadena, Calif., 1991).
16. T. P. Lee and A. G. Dentai, "Power and modulation bandwidth of GaAs-AlGaAs high-radiance LED's for optical communication systems," IEEE J. Quantum Electron. **QE-14,** 150–159 (1978).
17. C. Juang, K. J. Kuhn, and R. B. Darling, "Selective etching of GaAs and $Al_{.3}Ga_{.7}As$ with citric acid/hydrogen peroxide solution," J. Vac. Sci. Technol. B **5,** 1122–1124 (1990)
18. D. Rumelhart, J. L. McClelland, and the PDP Research Group, *Parallel Distributed Processing: Explorations in the Microstructure of Cognition* (MIT Press, Cambridge, Mass., 1986), Vol. 1.
19. R. A. Milano, T. H. Windhorn, E. R. Anderson, G. E. Stillman, R. D. Dupuis, and P. D. Dapkus, "$Al_{0.5}Ga_{0.5}As$-GaAs heterojunction phototransistors grown by metalorganic chemical vapor deposition," Appl. Phys. Lett. **34,** 562 (1979).
20. N. Chand, P. A. Houston, and P. N. Robson, "Gain of a heterojunction bipolar phototransistor," IEEE Trans. Electron Devices **ED-32,** 622–627 (1985).
21. S. H. Lin, F. Ho, J. H. Kim, and D. Psaltis, "Monolithic integrated optoelectronic thresholding devices for neural net-

work applications," in *Conference on Lasers and Electro-Optics,* Vol. 10 of 1991 OSA Technical Digest Series (Optical Society of America, Washington, D.C. 1991), paper CTuD1.

22. W. Kośćielniak, J.-L. Pelouard, R. Kolbas, and M. A. Littlejohn, "Dark current characteristics of GaAs metal-semiconductor-metal (MSM) photodetectors," IEEE Trans. Electron Devices **37,** 1623–1629 (1990).
23. J. C. Gammel and J. M. Ballantyne, "The OPFET: a new high speed optical detector," in *Digest of International Electron Device Meeting* (Optical Society of America, Washington, D.C., 1978), pp. 120–121.
24. J. C. Gammel and J. M. Ballantyne, "High speed photoresponse mechanism of a GaAs-MESFET," Jpn. J. Appl. Phys. **19,** L273 (1980).
25. J. Katz, N. Bar-Chaim, P. C. Chen, S. Margalit, I. Uryand, D. Wilt, M. Yust, and A. Yariv, "Monolithic integration of a GaAlAs buried-heterostructure laser and a bipolar phototransistor," Appl. Phys. Lett. **37,** 211 (1980).
26. S. H. Lin, J. H. Kim, J. Katz, and D. Psaltis, "Integration of high-gain double heterojunction GaAs bipolar transistors with a LED for optical neural network applications," in *Proceedings of the IEEE/Cornell Conference on Advanced Concepts in High Speed Semiconductor Devices and Circuits,* (Institute of Electrical and Electronics Engineers, New York, 1989), pp. 344–352.
27. J. H. Kim, S. H. Lin, J. Katz, and D. Psaltis, "Monolithically integrated two-dimensional arrays of optoelectronic threshold devices for neural network applications," in *Laser Diode Technology and Applications,* L. Figueroa, ed., Proc. Soc. Photo-Opt. Instrum. Eng. **1043,** 44–52. (989).
28. S. M. Sze, *Physics of Semiconductor Devices,* 2nd ed. (Wiley, New York, 1981), Chap. 2, p. 92.