



CONTRIBUTED ARTICLE

Light-Adaptive Architectures for Regularization Vision Chips

HARUO KOBAYASHI,¹ TAKASHI MATSUMOTO,² TETSUYA YAGI,³ AND KOJI TANAKA²¹Yokogawa Electric Corporation, ²Waseda University, and ³Kyushu Institute of Technology

(Received 16 July 1993; revised and accepted 2 May 1994)

Abstract—Light-adaptive algorithms/architectures are proposed for regularization vision chips. The adaptation mechanisms allow the regularization parameters to change in an adaptive manner in accordance with the light intensity of given images. This is achieved by adaptively changing the conductance values associated with massively parallel resistive networks. The algorithms/architectures are inspired by the adaptation mechanisms of the horizontal cells in the lower vertebrate retina.

Keywords—Regularization, Vision chip, Image processing, Adaptation, Analog CMOS, Retina.

1. INTRODUCTION

Vision chips generally refer to massively parallel arrays of simple analog circuits together with parallel array sensors. Due to their algorithms/architectures, those chips perform signal processing in an extremely fast manner with relatively low power dissipation compared with their digital counter parts (Ruetz & Brodersen, 1987; Heatin, Blevins, & Davis, 1990; Maruyama et al., 1990). Many of the chip architectures/algorithms are motivated by visual information processing mechanisms found in vertebrates, in particular, the retina (Mead & Mahowald, 1988; Mead, 1989; Boahen & Andreou, 1992; Matsumoto et al., 1992; Shimmi et al., 1992; Kobayashi et al., 1993). Some of the chip algorithms/architectures are derived from the Tikhonov regularization theory developed for solving ill-posed problems.

This paper proposes regularization vision chip architectures that incorporate **light adaptation**. More precisely, the architecture enables one to change the **filter width** of a $\nabla^2 G$ -like filter in accordance with the input light intensity. Two adaptation architectures are proposed: global and local. In the former, global light intensity information controls a particular parameter in the information processing cells (pixels) altogether, whereas in the latter, local light intensity information

regulates a parameter in each individual pixel. CMOS circuits are also proposed to implement the adaptation mechanisms.

The problem is formulated in terms of (a discrete version of) the regularization theory, whereas the adaptation algorithm has been inspired by the adaptation mechanism of the horizontal cells in the lower vertebrate retina.

2. PHYSIOLOGICAL BACKGROUND

This section briefly describes a neural adaptation system found in the lower vertebrate retina together with our interpretations on previous physiological findings. Figure 1 shows a fundamental structure of the vertebrate retina. Although the retina is transparent, the figure is colored for an obvious reason. Light comes from the bottom side of the figure and passes through the transparent retina to reach the photoreceptors (gray). The electrical signal is transmitted to the *second-order* neurons, which are the horizontal cells (blue) and the bipolar cells (red). These three types of neurons interact via chemical synapses in the outerplexiform layer, which is a morphologically identifiable lamina seen in the cross section of the retina (indicated by arrow 1 in Figure 1b). It is well known that neighboring horizontal cells are coupled by electrical synapses and possess a large receptive field that sometimes covers almost the entire retina (Naka & Rushton, 1967; Yagi, 1986). Neighboring photoreceptors are also coupled electrically, but the size of receptive field is much smaller than that of horizontal cells (Baylor, Fuortes, & O'Bryan, 1971; Schwartz, 1973). In the present study, the network of photoreceptor and horizontal cell is described by an equivalent electrical circuit, as shown

Acknowledgements: We would like to thank S. Yasui of Kyushu Institute of Technology for valuable discussions. Thanks are also due to the reviewers for carefully reading the manuscript and the detailed comments.

Requests for reprints should be sent to H. Kobayashi, Yokogawa Electric Corp., 2-9-32 Naka-cho Musashino, Tokyo 180, Japan.

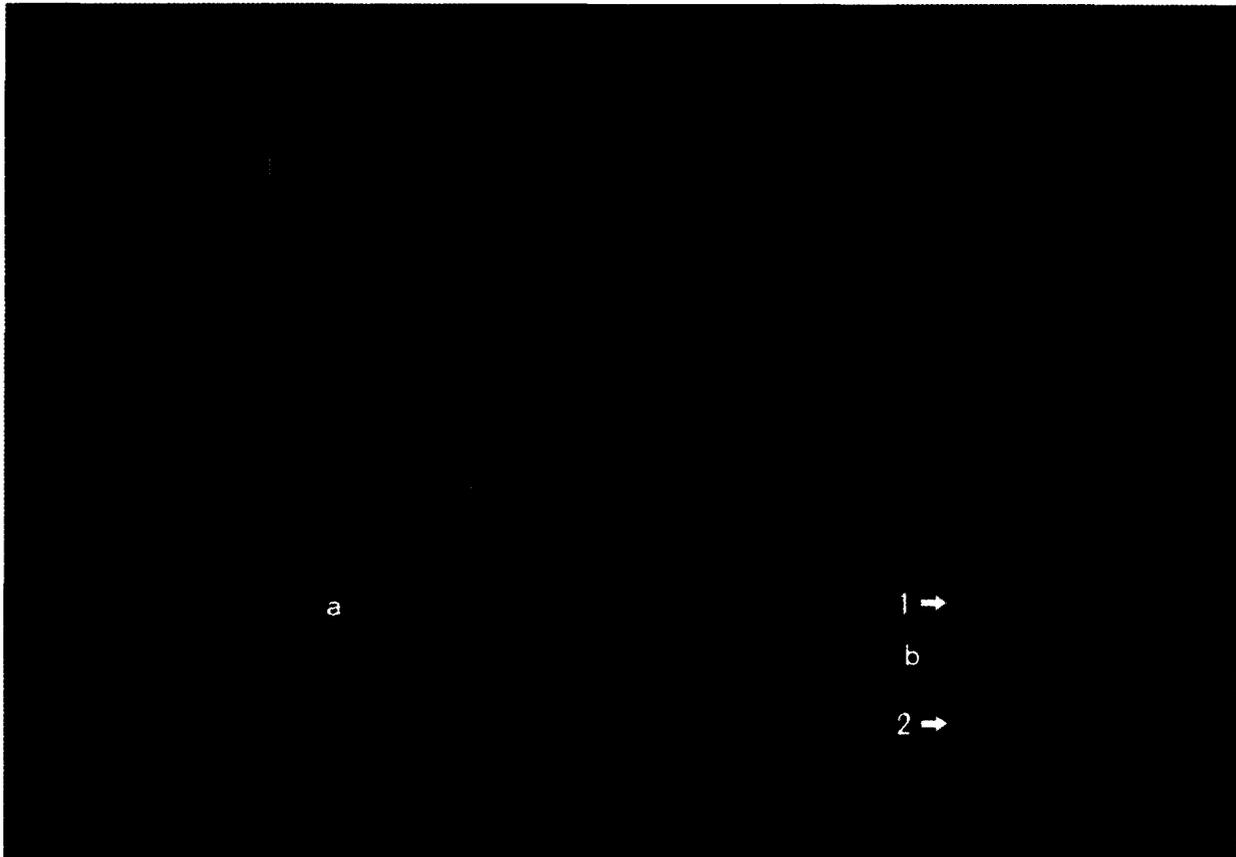


FIGURE 1. Schematic illustration of the vertebrate retina. (a) Overview of the retina. The retina is a very thin tissue (200 ~ 300 μm) that consists of six major types of cells. These major cell types are distinguished by different colors in the figure. Gray, photoreceptors; red, bipolar cells; blue, horizontal cells; white, amacrine cells; orange, ganglion cells; light gray, IP cells. The light comes from the bottom side of the figure. The photoreceptors, horizontal cells, and bipolar cells are located in the distal part of the retina. The amacrine cells and ganglion cells are located in the proximal part of the retina. The ganglion cells are the output neurons of the retina. The IP cell (light green) locates the cell body in the proximal part of the retina to send a long fiber (axon) to the horizontal cell. See text for details. (b) Cross section of the retina. Different types of cells are arranged in separate layers. The locations of the outer and the inner plexiform layers are indicated by arrows Nos. 1 and 2, respectively. Reprinted with permission from Matsumoto, T., Kobayashi, H., and Yagi, T. (1993), Vision chip [1]—analog image-processing neuro chips. *Proceedings of IEICE*, 76(7), 783–791.

in Figure 2. Each photoreceptor is represented by a conductance g_{m1} and each horizontal cell by g_{m2} . The conductance g_{s1} represents the electrical coupling between photoreceptors. The conductance g_{s2} represents the electrical coupling between horizontal cells. The significance of electrical coupling is thought to be relevant to the reduction of noise occurring in the retinal neural circuit. When cells are electrically coupled, the current generated in a single cell spreads into neighboring cells. Because the intrinsic noise in each cell is not correlated, the signal-to-noise ratio can be improved when the image has an appropriate size (Lamb & Simon, 1976; Tessier-Lavigne & Attwell, 1988). However, the electrical coupling blurs the image in return. Thus, the coupling strength between neighboring cells is a critical parameter to be determined by the trade-off between these conflicting two factors.

The bipolar cell is the first neuron that exhibits a ∇^2G -like receptive field in the vertebrate visual system.

In the bipolar cell, the response to a stimulus placed in the center region antagonizes the one placed in the surround region. Concerning this receptive field, it is widely believed that the center response is mediated by the direct input from the photoreceptor and the antagonistic surround response is mediated by the horizontal cell. As was demonstrated by Marr and Hildreth (1980), the ∇^2G -like receptive field is able to perform a smoothing and a contrast enhancement of the input image simultaneously. It is possible, under certain conditions, to identify edges of an object with the zero-crossings of a ∇^2G -like filter. It is important to notice that the size of receptive field should be different depending on the signal-to-noise ratio of the input image. The receptive field of the bipolar cell is controlled in the retinal neural circuit as described below.

The **interplexiform cell** (IP cell) is a unique neuron (colored light green in Figure 1) and is believed to be a centrifugal neuron innervating to the horizontal cell.

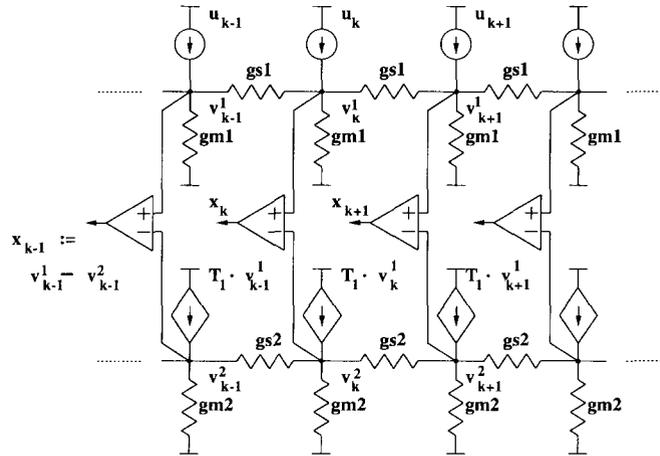


FIGURE 2. The double-layer network realizing a ∇^2G -like filter.

Its cell body is located near the inner plexiform layer (arrow 2 in Figure 1b) with ascending axons to horizontal cells (Dowling & Ehinger, 1978). Teranishi, Negishi, and Kato (1983) found that a physiologically active substance, **dopamine**, is released from the IP cell and reduces a receptive field size of the horizontal cell by decreasing the conductance of electrical synapses connecting neighboring cells. More recently, it was demonstrated that the effect of dopamine on the horizontal cell is mimicked by exposing the retina in the light-adapted state (Shigematsu & Yamada, 1988). These observations indicate that the receptive field size of the horizontal cell is reduced in the light-adapted state and consequently the receptive field of the bipolar cell becomes smaller. Based on these previous physiological findings, we *hypothesize* that the IP cell *adaptively controls the receptive field size of the horizontal cell* according to the signal-to-noise ratio of the image. If we assume that the intrinsic noise is constant regardless of the adaptation level of the retina, the relative magnitude of noise-to-signal is small in the daytime because the light intensity of signal image is high. In that situation, the size of the bipolar cell receptive field is to be reduced to **gain the spatial resolution**. This adaptation is likely to be carried out by the IP cell in the retinal neural circuit.

The adaptive architectures of the vision chip proposed in this paper have been inspired by the above physiological background. Namely, the coupling conductance between horizontal cells (g_{s2} of Figure 2) is set to be variable. The function and the mechanism of the proposed adaptive vision chip are described in the following sections.

3. EARLY VISION PROBLEMS VIA TIKHONOV REGULARIZATION

When a solution to an operator equation (not necessarily linear),

$$Av = d, \quad v \in X, \quad d \in Y, \quad (1)$$

loses existence or uniqueness or continuity in d , eqn (1) is called ill-posed. Ill-posedness typically arises when “data” d is noisy, whereas the solution v sought should be reasonably smooth. It can also result from the nature of A . The Tikhonov regularization (Tikhonov, 1963a,b, 1965) converts eqn (1) into a family of minimization problems:

$$G(v, d, \lambda) = \|Av - d\|^2 + \lambda\Omega(v) \quad (2)$$

where $\|\cdot\|$ denotes a norm (on Y), $\Omega: X \rightarrow R$ is continuous and strictly convex, $\lambda > 0$. If $Av^* = d^*$, then under reasonable conditions, eqn (2) regularizes eqn (1) in the sense that for any ϵ -neighborhood $N_\epsilon(v^*)$ of v^* (with respect to an appropriate topology), there is a δ -neighborhood $N_\delta(d^*)$ of d^* , such that if $d \in N_\delta(d^*)$, and if $\lambda(\delta) > 0$ is appropriate, then there is a unique $v(d, \lambda(\delta)) \in N_\epsilon(v^*)$ that minimizes eqn (2). It should be noted, however, that when d is noisy, choosing the best λ is another interesting, as well as difficult, problem because one needs to take into account the statistics of d (MacKay, 1991; Whaba, 1987), and it is outside of the scope of this paper. It is argued in Poggio, Torre, and Koch (1985) and Poggio, Voorhees, and Yuille (1985) that many of the early vision problems (edge detection, stereo, optical flow, etc.) can be formulated as Tikhonov regularization problems.

Now a typical “stabilizer” $\Omega(v)$ in eqn (2) is of the form

$$\Omega(v) = \sum_{r=1}^P \int_D C_r \left(\frac{d^r v(x)}{dx^r} \right)^2 dx \quad (3)$$

where $C_r \geq 0$ and $D = [a, b]$ is the domain of the problem. If eqn (2) with eqn (3) can be written as

$$G(v, d, \lambda) = \int_D F(v(x), v^{(1)}(x), v^{(2)}(x), \dots, v^{(P)}(x), x, d(x), \lambda) dx,$$

$$v^{(r)} = \frac{d^r v}{dx^r} \quad (4)$$

where F is “well-behaved,” then the variational principle gives Euler’s equation

$$\sum_{r=0}^P (-1)^r \frac{d^r}{dx^r} \frac{\partial}{\partial v^{(r)}} \times F(v(x), v^{(1)}(x), \dots, v^{(P)}(x), x, d(x), \lambda) = 0 \quad (5)$$

with natural boundary conditions:

$$\sum_{r=0}^P (-1)^r \frac{d^r}{dx^r} \frac{\partial}{\partial v^{(P-q-r)}} \times F(v(x), v^{(1)}(x), \dots, v^{(P)}(x), x, d(x), \lambda) = 0, \quad \text{at } x = a, b \text{ for } q = 0, 1, \dots, P.$$

Most of the vision chips implemented/proposed to date consist of massively parallel (analog) resistive networks (Mead & Mahowald, 1988; Mead, 1989; Boahen & Andreou, 1992; Harris, 1986, 1988; Harris et al., 1989; Hutchinson et al., 1988; Liu & Harris, 1989; Mathur, Lin, & Wang, 1985; Matsumoto et al., 1992; Shimmi et al., 1992; Kobayashi et al., 1993; Kobayashi, White, & Abidi, 1990, 1991). Thus, in a chip the space variable x takes finite discrete values. Implementations of Euler’s equation or the natural boundary conditions separately by parallel resistive networks are easy. It is rather difficult, however, to implement a resistive network that simultaneously fulfills eqns (4) and (5) when $P \geq 2$. To see the difficulty, first observe that because of the particular form of eqn (3), Euler’s equation (4) necessarily contains terms of the form

$$\left(\frac{d^{2r}v}{dx^{2r}} \right)(x), \quad r = 1, 2, \dots, P.$$

Namely, if the stabilizer (3) contains the r th order derivative, one needs to implement the **2rth order** derivative operation for solving the regularization problem. On the other hand, eqn (5) contains other orders of derivatives of $v(x)$. There are two more factors that make the problem even more difficult. First, eqns (4) and (5) are for a single space variable x , whereas in a vision chip there are two space variables x and y , which significantly complicates eqn (5) (see Courant & Hilbert, 1953; Kobayashi et al., 1993). Second, in most vision chips, resistive networks have a “hexagonal geometry” instead of a square geometry. It is rather involved to implement the two-dimensional versions of Euler’s equation together with its natural boundary conditions on a hexagonal grid.

This naturally leads us to a finite dimensional discrete formulation of the regularization problems. Namely, let

$$\mathbf{v} := (v_1, v_2, \dots, v_n) \in R^n$$

and replace the derivative operations by the difference operations, for example,

$$\left(\frac{dv}{dx} \right)(x) \rightarrow v_k - v_{k-1}, \quad \left(\frac{d^2v}{dx^2} \right)(x) \rightarrow v_{k+1} - 2v_k + v_{k-1},$$

which can be put in a vector form as

$$\left(\frac{dv}{dx} \right)(\cdot) \rightarrow \mathbf{D}\mathbf{v}, \quad \left(\frac{d^2v}{dx^2} \right)(\cdot) \rightarrow \mathbf{L}\mathbf{v},$$

respectively, where

$$\mathbf{D} = \begin{bmatrix} 1 & 0 & 0 & \cdot & \cdot & \cdot & 0 \\ -1 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 0 & -1 & 1 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & 1 & 0 \\ 0 & 0 & 0 & \cdot & \cdot & -1 & 1 \end{bmatrix},$$

$$\mathbf{L} = \begin{bmatrix} -2 & 1 & 0 & \cdot & \cdot & \cdot & 0 \\ 1 & -2 & 1 & \cdot & \cdot & \cdot & 0 \\ 0 & 1 & -2 & \cdot & \cdot & \cdot & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ 0 & 0 & 0 & \cdot & \cdot & -2 & 1 \\ 0 & 0 & 0 & \cdot & \cdot & 1 & -2 \end{bmatrix}.$$

Therefore, the stabilizer corresponding to eqn (3) is given by

$$\Omega(\mathbf{v}) = \sum_{r=1}^P \begin{cases} C_r \|\mathbf{L}^{r/2}\mathbf{v}\|^2 & r: \text{even} \\ C_r \|\mathbf{D}\mathbf{L}^{r-1/2}\mathbf{v}\|^2 & r: \text{odd} \end{cases}$$

so that the regularization problem on R^n corresponding to eqn (2) calls for the minimization of

$$G(\mathbf{v}, \mathbf{d}) = \|\mathbf{A}\mathbf{v} - \mathbf{d}\|^2 + \sum_{r=1}^P \begin{cases} \lambda_r \|\mathbf{L}^{r/2}\mathbf{v}\|^2 & r: \text{even} \\ \lambda_r \|\mathbf{D}\mathbf{L}^{r-1/2}\mathbf{v}\|^2 & r: \text{odd} \end{cases} \quad (6)$$

where \mathbf{A} is now a map between finite dimensional spaces and

$$\lambda_r := \lambda C_r$$

are called the **regularization parameters**. Differentiating eqn (6) with respect to \mathbf{v} and setting it to zero, one has

$$\frac{1}{2} \frac{\partial G}{\partial \mathbf{v}} = \mathbf{A}^T(\mathbf{A}\mathbf{v} - \mathbf{d}) + \sum_{r=1}^P (-1)^r \lambda_r \mathbf{L}^r \mathbf{v} = 0 \quad (7)$$

where

$$\mathbf{D}^T \mathbf{D} = -\mathbf{L}$$

was used. Note that if $\lambda_r \neq 0$, then the solution (7) necessarily contains the $\mathbf{L}^r \mathbf{v}$ term. This corresponds to the presence of the $(d^{2r}v/dx^{2r})(x)$ term in the infinite dimensional case. In this paper as well as in Matsumoto et al. (1992), Shimmi et al. (1992), and Kobayashi et al. (1990, 1991) $\mathbf{A} = \mathbf{I}$, the identity map. Other forms of \mathbf{A} are possible. When the data \mathbf{d} is sparse, for instance, \mathbf{A} is a projection operator onto a lower dimensional subspace, which gives rise to a rather interesting architecture that will be reported elsewhere.

Let us explain how eqn (7) can be naturally mapped into parallel resistive networks. To this end, consider the simplest case $P = 1$. Then eqn (7) reads

$$v_k - d_k - \lambda_1(v_{k-1} + v_{k+1} - 2v_k) = 0. \quad (8)$$

Equation (8) is naturally mapped into a parallel resistive network where each node k is induced by a current source u_k , and connected with g_0 to ground, and nodes $k + 1$ and $k - 1$ with g_1 with the identification

$$\lambda_1 = \frac{g_1}{g_0}, \quad d_k = \frac{u_k}{g_0}.$$

This structure is extremely popular in the analog vision chips (Mead & Mahowald, 1988; Mead, 1989). Now note that regularization parameter λ_1 specifies how much weight should be put on the first-order derivative penalty. Similarly, each λ_r in eqn (6) determines the weight on the r th derivative penalty.

4. LIGHT-ADAPTIVE ARCHITECTURE

In all the vision chip architectures implemented/proposed to date that we know of, the regularization parameters λ_r are **fixed**. Our architecture proposed below makes λ_r variable so that adaptation can be incorporated. Most generally, λ_r can depend on \mathbf{v} , \mathbf{d} , and k . The dependency of λ_r on \mathbf{v} makes eqn (6) nonquadratic, and general analytical form corresponding to eqn (7) can be nonlinear, which we do not pursue, at least in the present paper. Although the dependency of λ_r on k does not alter the quadratic nature of the problem, the generalization in this direction does not, so far, find interesting enough applications. Therefore, we will consider the minimization of eqn (6) where λ_r is now $\lambda_r(\mathbf{d})$. Although this requires only a straightforward modification in eqn (7), that is, λ_r should be replaced with $\lambda_r(\mathbf{d})$, it leads to rather interesting adaptation networks. Among many possible adaptive networks, the SCE (smoothing-contrast enhancement) filter network (Matsumoto et al., 1992; Shimmi et al., 1992; Kobayashi et al., 1993; Boahen & Andreou, 1992) has probably one of the most interesting structures suited for this adaptation. To explain what SCE filter network does, let us first state the following fact proved in Kobayashi et al. (1993).

Fact 1. Consider the double-layer network given in Figure 2.

- (i) The second layer voltage distribution v_k^2 solves the second-order regularization problem with

$$\lambda_1 = \frac{g_{s1}}{g_{m1}} + \frac{g_{s2}}{g_{m2}}, \quad \lambda_2 = \frac{g_{s1}g_{s2}}{g_{m1}g_{m2}}, \quad d_k = \frac{T_1}{g_{m1}g_{m2}} u_k.$$

- (ii) The first layer voltage distribution v_k^1 solves the first-order regularization problem with

$$\lambda_1' = \frac{g_{s1}}{g_{m1}}, \quad d_k' = \frac{u_k}{g_{m1}}.$$

- (iii) The difference

$$x_k := v_k^1 - v_k^2$$

enhances contrast of u_k after smoothing.

REMARKS.

- (i) Figure 3 shows the responses x_k to a narrow “slit” located at the center with two different g_{s2} s:

$$1/g_{s2} = 5M\Omega \quad \text{and} \quad 500k\Omega, \quad (9)$$

respectively, whereas other parameters are fixed at

$$1/g_{s1} = 30M\Omega, \quad 1/g_{m1} = 1/g_{m2} = 1G\Omega,$$

$$T_1 = 10^{-9} \text{ siemens}. \quad (10)$$

The parameter values given in eqns (9) and (10) are taken from the physiological studies on carp retina (Yagi & Kaneko, 1987; Ohshima, Yagi, & Funahashi, in press) so that parameter rescaling is necessary when one designs transistor circuits. Two facts are clear. First, the network response naturally approximates the well-known $\nabla^2 G$ filter. Second, with different g_{s2} -values, one can have different “filter width,” which roughly corresponds to changing “sigma” of the $\nabla^2 G$ filter. There is one feature associated with our network that is not associated with the $\nabla^2 G$ filter. Namely, **a larger g_{s2} results in a higher filter gain**. This is very natural if g_{s2} increases when the environment is darker, which is exactly what happens in vertebrate retina (see Section 2).

- (ii) This network is obviously temporally as well as spatially stable in the sense of Matsumoto, Kobayashi, and Togawa (1992).

The following fact is a straightforward consequence of Fact 1 and the argument preceding it.

Fact 2. Consider the double-layer network given in Figure 2, where the second layer horizontal conductance g_{s2} has an adaptation mechanism described by

$$g_{s2}(\mathbf{u}) := \frac{1}{G \sum_k u_k}, \quad G > 0, \quad (11)$$

where G is a constant and u_k is a photocurrent induced at node k . Then (i) the second layer voltage distribution v_k^2 solves the second-order regularization problem with

$$\lambda_1(\mathbf{u}) = \frac{g_{s1}}{g_{m1}} + \frac{g_{s2}(\mathbf{u})}{g_{m2}}, \quad \lambda_2(\mathbf{u}) = \frac{g_{s1}g_{s2}(\mathbf{u})}{g_{m1}g_{m2}}$$

so that the weight ratio is given by

$$\frac{\lambda_2(\mathbf{u})}{\lambda_1(\mathbf{u})} = \frac{1}{g_{m1}/g_{s1} + g_{m2}G(\sum_k u_k)}, \quad d_k = \frac{T_1}{g_{m1}g_{m2}} u_k. \quad (12)$$

Statements (ii) and (iii) of Fact 1 are still valid.

REMARKS.

- (i) When the total input current $\sum_k u_k$ gets larger, which amounts to the facts that the environment is **light**, the second-layer horizontal conductance g_{s2} decreases. Although the decrease of g_{s2} changes both $\lambda_1(\mathbf{u})$ and $\lambda_2(\mathbf{u})$, the ratio $\lambda_2(\mathbf{u})/\lambda_1(\mathbf{u})$ decreases [eqn (12)]. This means that when $\sum_k u_k$

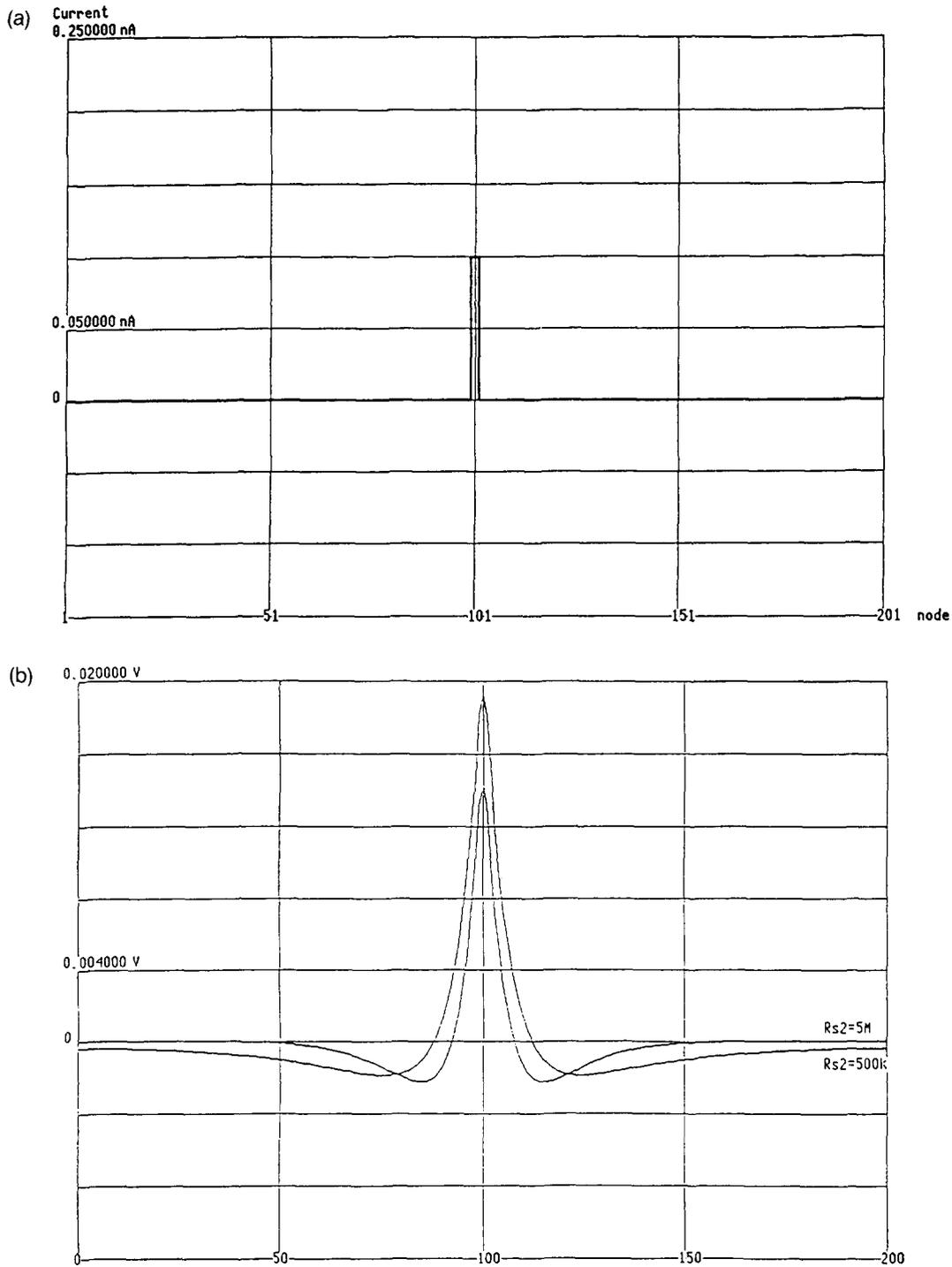


FIGURE 3. Responses of the double-layer network for a “slit” input. (a) A slit input with 3-pixel width. (b) Responses with $1/g_{s2} = 5M\Omega$ and $1/g_{s2} = 500k\Omega$, respectively.

is large, the emphasis of the network on the second-order derivative decreases. This adaptation mechanism has rather interesting implications. Suppose that $u_k = u_k^0 + \xi_k$, where u_k^0 is the noiseless image whereas ξ_k stands for noise. Suppose also that the mean of the noise has been absorbed into u_k^0 so that ξ_k has zero mean. If $\xi_{\min} \leq \xi_k \leq \xi_{\max}$ where ξ_{\min} and ξ_{\max} are independent of u_k^0 , then $\sum_k u_k$ large means that effect of noise is less

significant than when $\sum_k u_k$ is smaller. Thus, when $\sum_k u_k$ is smaller, noise is more significant and the network puts more emphasis on the second-order derivative penalty. Figure 4 shows the effect of the adaptation mechanism. Figure 4a is a (one-dimensional) rectangular “image”

$$u_k^0 = \begin{cases} 1\mu A & 61 \leq k \leq 141 \\ 0 & \text{otherwise} \end{cases}$$

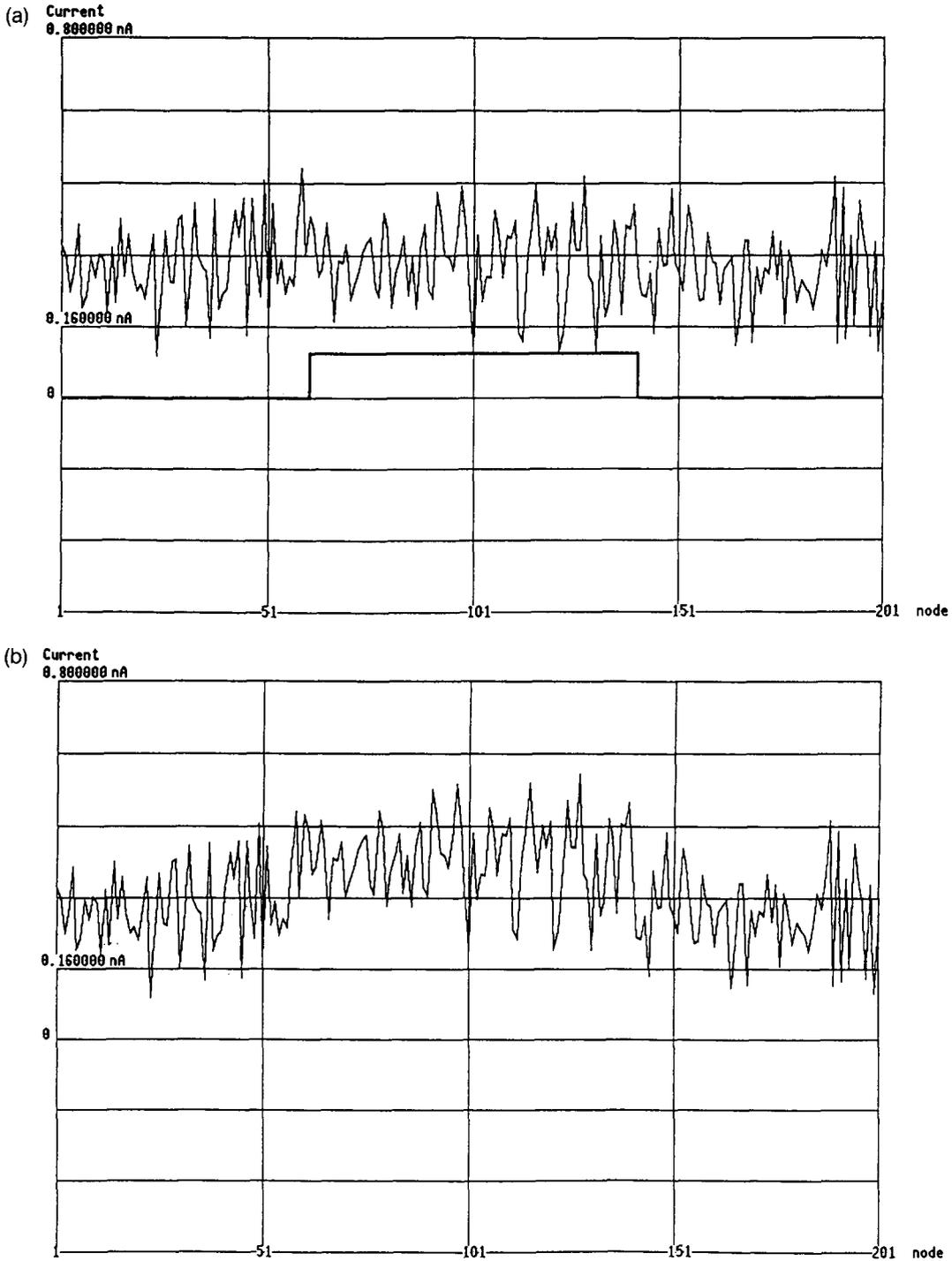


FIGURE 4. One-dimensional input image. (a) A rectangular image with 80-pixel width and noise with $m = 300pA$, $3\sigma = 600pA$. (b) Sum of the rectangular image and the noise.

and the Gaussian white noise with mean $300pA$, $3\sigma = 600pA$ whereas Figure 4b shows the sum of them. Figure 5a shows the network response x_k , where

$$\begin{aligned} 1/g_{s2} &= 5M\Omega, & 1/g_{s1} &= 30M\Omega, \\ 1/g_{m1} &= 1/g_{m2} = 1G\Omega, & T_1 &= 10^{-9} \text{ siemens.} \end{aligned} \quad (13)$$

A dramatic effect is discernible when the g_{s2} adaptation (11) is incorporated where

$$G = 1.0 \times 10^{13}.$$

It is known that $\nabla^2 G$ filter identifies edges of an object by its zero crossings even though not every zero crossing corresponds to an edge (Clark, 1989). Observe that although Figure 5a gives no information about the edges of the original object, Figure 5b, which is the network response with the g_{s2} adaptation given by eqn (11), correctly identifies the edge of the original image by its zero

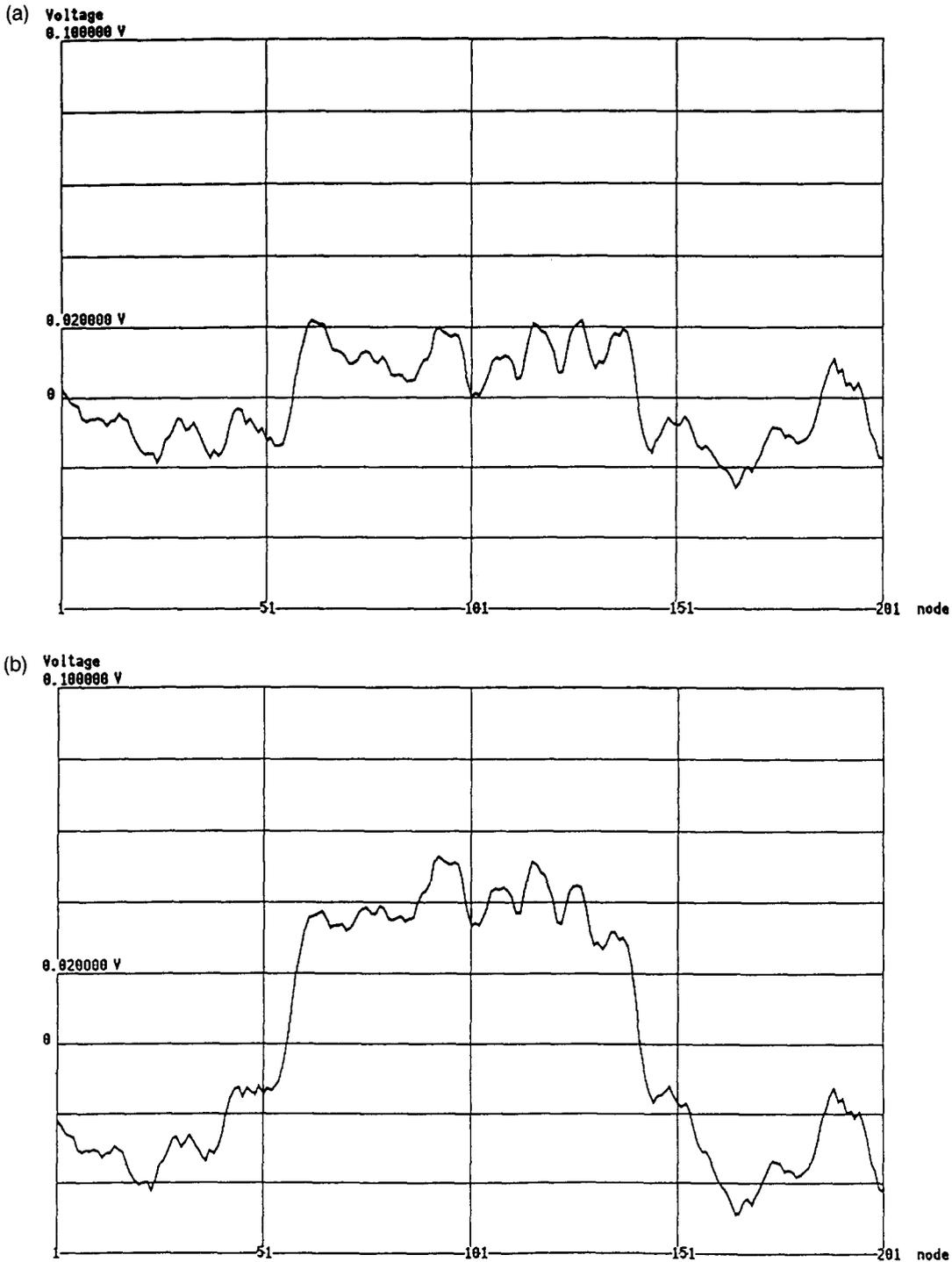


FIGURE 5. Responses of the network in Figure 2 for the input image in Figure 4a. (a) Adaptation is not incorporated ($1/g_{s2} = 5M\Omega$). (b) Adaptation of eqn (11) is incorporated with $G = 1.0 \times 10^{13}$.

crossings. Figure 6 gives the responses of the networks with eqn (13) and $G = 2.0 \times 10^{15}$ for a slit input.

- (ii) In Matsumoto et al. (1992) and Shimmi et al. (1992), the g_{s2} values are changed manually.
- (iii) Because the photocurrent u_k is always positive, one does not have to square it or one does not have to take the absolute value. In fact, v_k and

v_k^2 are also positive. The output $x_k = v_k^1 - v_k^2$, however, can be negative.

We remark that Boahen and Andreou (1992) proposed and implemented a vision chip for which the architecture was inspired by a physiological model of vertebrate retina (Yagi et al., 1989) whereas another one of our earlier vision chips reported in Kobayashi

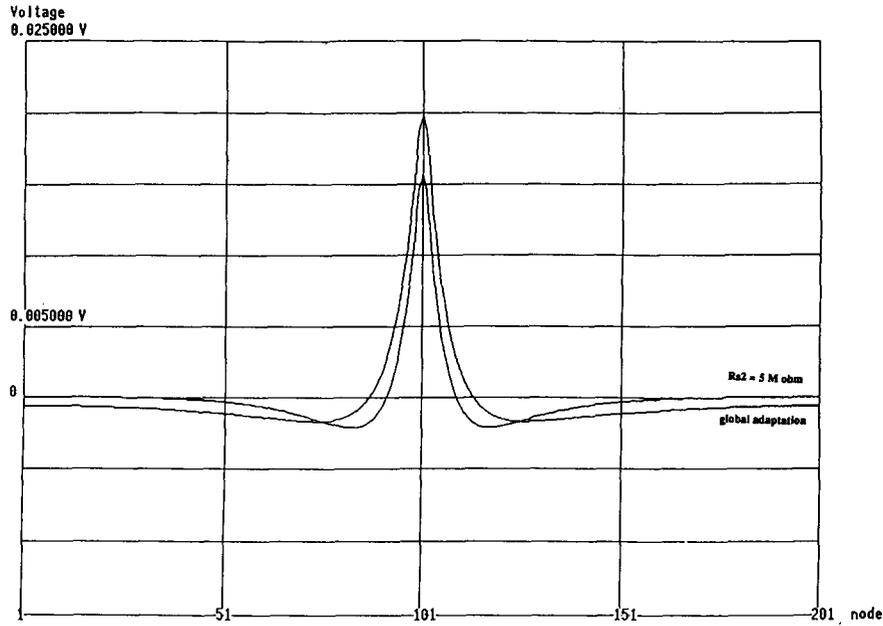


FIGURE 6. Responses of the networks in Figure 2 with $1/g_{s2} = 5M\Omega$ (no adaptation) and $G = 2.0 \times 10^{15}$ (adaptation) for a slit input in Figure 3a.

et al. (1990, 1991) implements a Gaussian-like convolver. In these chips, some of the network parameters are changed manually, and it would be interesting to implement adaptation networks that modify the parameters automatically.

5. ANALOG CMOS CIRCUITS FOR LIGHT ADAPTATION

An analog CMOS implementation of the network given in Figure 2 without adaptation has been reported (Matsumoto et al., 1992; Shimmi et al., 1992). Figure 7 shows one of the 53×52 cells arranged on a hexagonal array. The input photosensor is realized as a photo-transistor in CMOS process (Mead, 1989) and the photo current is converted by a logarithmic law to a voltage using a diode-connected MOS FET in its sub-threshold region (Vittoz, 1985) to obtain wide input dynamic range. The input voltage is fed into the first-layer network through a buffer and the node voltage of the first-layer network is then applied via a buffer to the corresponding node of the second-layer network. This implements the Thévenin equivalent of the current sources in Figure 2. The buffer is realized by a transconductance amplifier with unity gain feedback. The first- and second-layer networks consist of horizontal conductances g_{s1} and g_{s2} , respectively, and vertical conductances g_{m1} and g_{m2} , respectively. g_{s1} , g_{m1} , and g_{m2} can be implemented with MOS conductance, polysilicon or diffusion conductance. The node voltage of the first-layer network is subtracted from the corresponding node voltage of the second-layer network using a differential pair subtracter, and it is read out through analog switches.

Analog CMOS circuits are suitable to implement the adaptation mechanism. A conductance can be realized with MOS FETs in triode region, and there are several ways to realize the MOS conductance (Mead, 1989; Banu & Tsividis, 1982). Because its value is changed by adjusting the bias voltage, the adaptation mechanism can be realized by incorporating circuits that vary the bias voltage of the MOS conductance according to input images.

Now let us describe a circuit that realizes eqn (11). Figure 8 shows a possible configuration, and note that the input circuit in Figure 7 is the Thévenin equivalent of the current source in Figure 2. Let us denote this equivalent voltage by

$$v_k^0 := g_{m1} u_k.$$

In Figure 8, this voltage v_k^0 is first converted into current I_k by the $V-I$ converter so that I_k is proportional to v_k^0 . The summation of all these currents can be obtained *for free* by simply connecting the wires together because of the Kirchhoff Current Law, and the summed current I is given by

$$I := \sum_k I_k \propto \sum_k v_k^0.$$

The current I is fed into the bias voltage generator that produces a bias voltage v_c so that the g_{s2} value is inversely proportional to I . Figure 9 shows circuit design example of the $V-I$ converter, g_{s2} , and the bias generator. The $V-I$ converter is designed with a differential pair and g_{s2} is implemented with two parallel MOS FETs (Banu & Tsividis, 1982) for which the value becomes larger as v_c increases. In the bias generator, the summed current I is subtracted from a bias current I_b ,

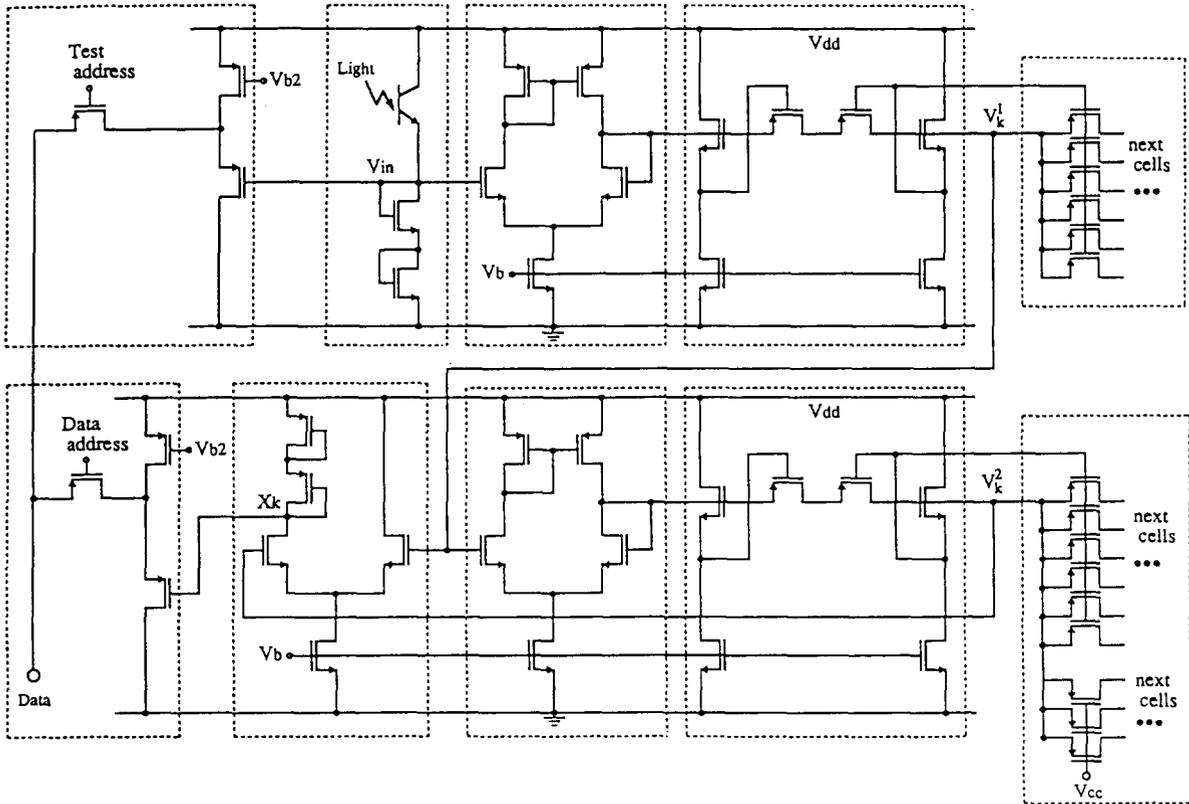


FIGURE 7. Unit cell circuit diagram of a double-layered network (Matsumoto et al., 1992; Shimmi et al., 1992).

and the resultant current $I_b - I$ flows into a resistor R and a diode-connected NMOS, which generate a bias voltage v_c . Thus, as I becomes smaller, v_c (and then g_{s2}) increases. Figure 10 shows SPICE simulation results of g_{s2} characteristics at several different values of $\sum_k v_k^0$ and we see that as $\sum_k v_k^0$ becomes larger, g_{s2} decreases. It should be noted that perfect linearity is not necessary at all.

6. OTHER ADAPTATIONS

6.1. Local Adaptation

The adaptation (11) is global in that the g_{s2} value changes according to the global information $\sum_k u_k$. If

$$g_{s2(k,k+1)} := \frac{1}{L(v_k^1 + v_{k+1}^1)}, \quad L > 0, \quad (14)$$

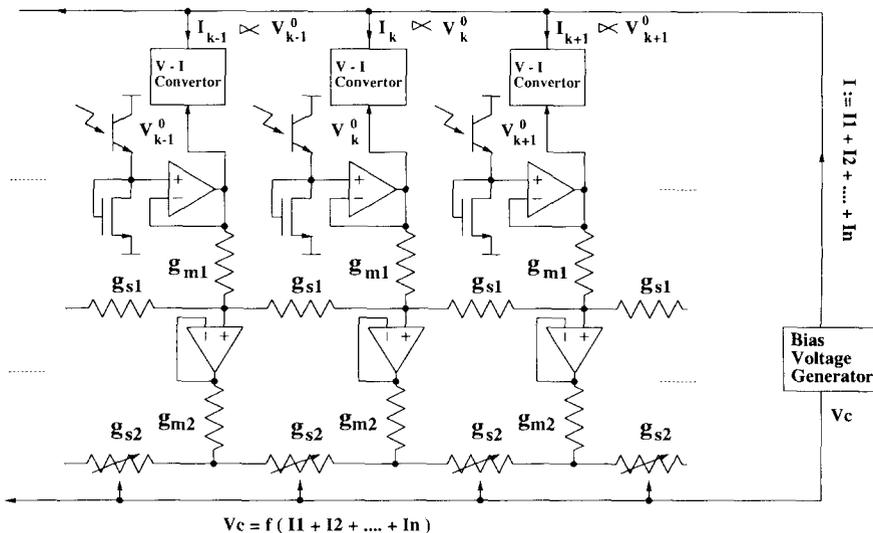


FIGURE 8. Block diagram of the light-adaptive network.

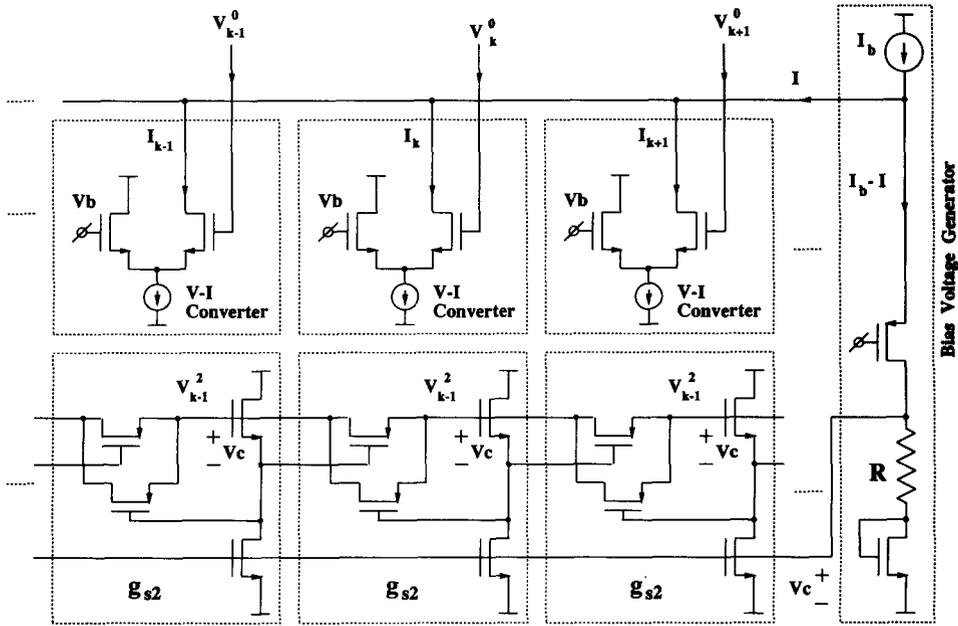


FIGURE 9. Circuit design of V-I converters, bias generators, and variable conductances g_{s2} in Figure 8.

where L is a constant, then the second-layer horizontal conductance value $g_{s2(k,k+1)}$ between node k and node $k + 1$ is inversely proportional to the sum of the first-layer voltages v_k^1 and v_{k+1}^1 . Figure 11a is a simple rectangular input whereas Figure 11b compares the response incorporating the local adaptation (14), where $L = 2 \times 10^7$ with those responses without adaptations

and where $1/g_{s2} = 5 M\Omega$ and $1/g_{s2} = 500k\Omega$, respectively. Even though the effect of the local adaptation is not as dramatic as in Figure 5, where the global adaptation is incorporated, one can see that where the input intensity is high, the response with eqn (14) is closer to that with $1/g_{s2} = 5 M\Omega$. On the other hand, where the intensity is low, the adapted response behaves

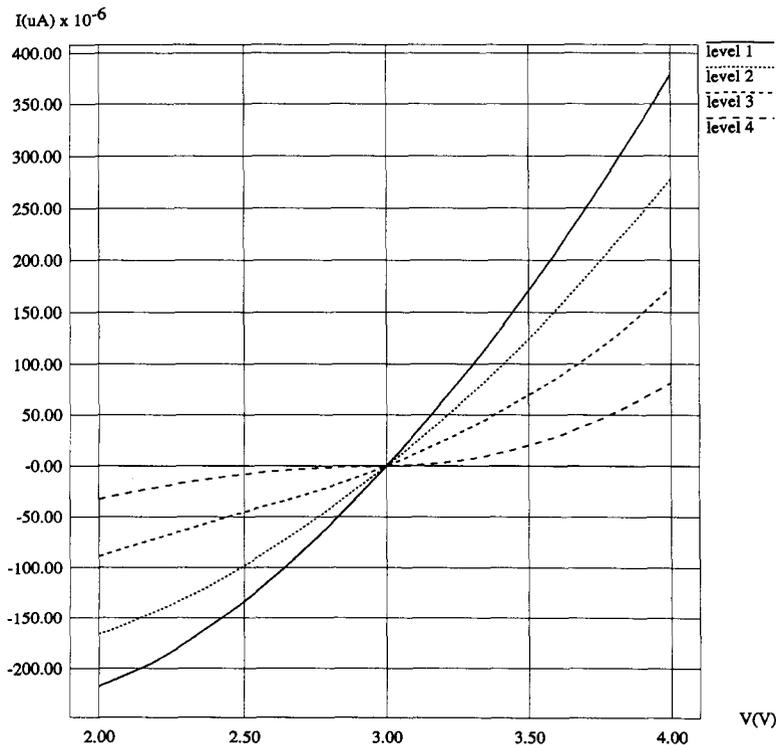


FIGURE 10. Simulation results of Figures 8 and 9. V-I characteristics of g_{s2} are shown at several different values of $\sum_k v_k^0$. The higher the level, the greater the value of $\sum_k v_k^0$.

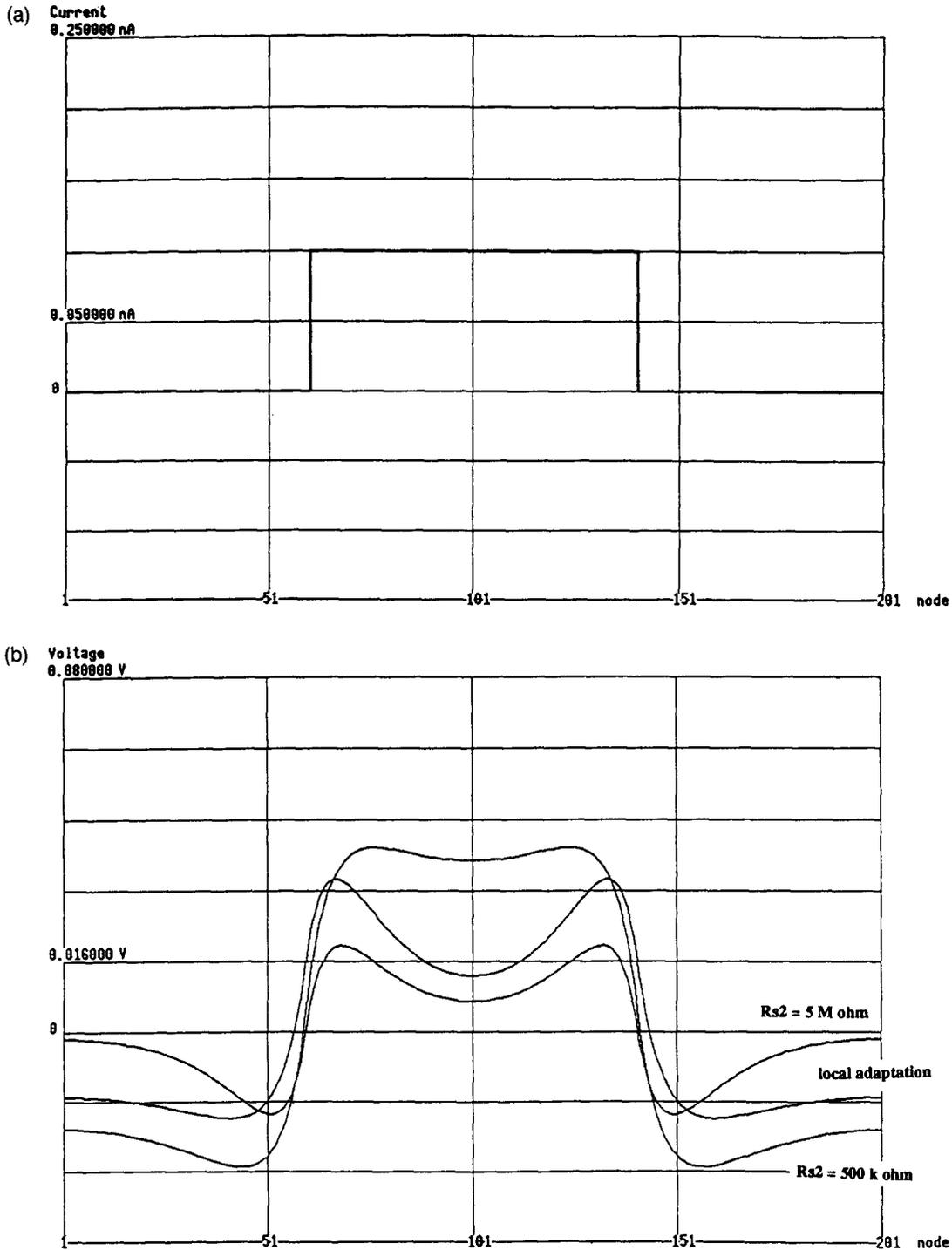


FIGURE 11. Response of the locally adaptive network. (a) A rectangular input image with 81-pixel width. (b) Responses of the networks with $1/g_{s2} = 5M\Omega$ (no adaptation), $1/g_{s2} = 500k\Omega$ (no adaptation), and $1/g_{s2(k,k+1)} = 2 \times 10^7(v_k^1 + v_{k+1}^1)$ (local adaptation).

in a manner similar to the one with $1/g_{s2} = 500k\Omega$. Therefore, with eqn (14), contrast is even more enhanced where interesting difference exists.

Figure 12 shows a possible circuit block diagram for the local adaptation and Figure 13 shows a circuit design of locally adaptive conductances g_{s2} and bias gen-

erators in Figure 12. The bias voltage generator at node k outputs v_k^c inversely proportional to the first-layer node voltage v_k^1 , and $g_{s2(k,k+1)}$ is implemented with two parallel MOS FETs for which the value is roughly proportional to $v_k^c + v_{k+1}^c$, and then this approximates eqn (14). Figure 14 shows SPICE simulation results of

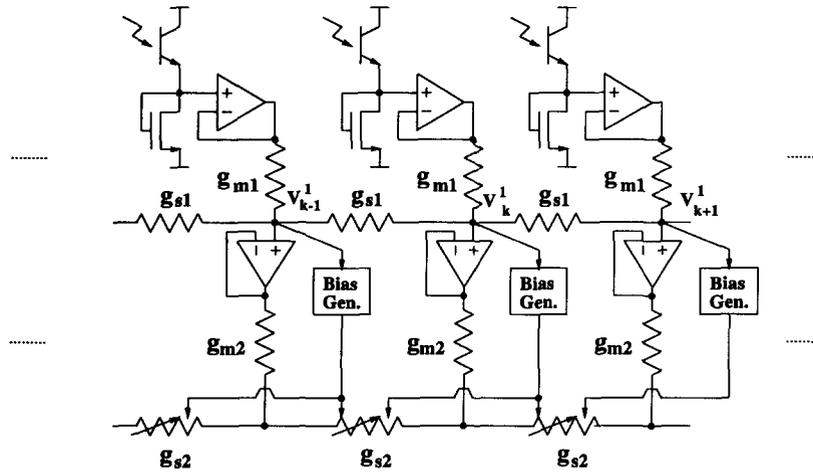


FIGURE 12. Block diagram of the locally adaptive network.

$g_{s2(k,k+1)}$ characteristics at several different values of $v_k^1 + v_{k+1}^1$. One sees that as $v_k^1 + v_{k+1}^1$ becomes larger, $g_{s2(k,k+1)}$ decreases.

6.2. Maximum Value Adaptation

Consider

$$u_k^* := \frac{u_k}{M \max_i(u_i)}, \quad M > 0, \quad (15)$$

which is implemented by the network in Figure 15, where it senses the maximum input voltage and changes the gain of PGAs (programmable gain amplifiers) uniformly to as high a value as possible without overloading the network. Because there are all kinds of noises in a chip, one obtains a better signal-to-noise ratio if the

input signal is amplified as much as possible without overloading the network. A similar method is widely used in A/D converters, where one can obtain a good signal-to-noise ratio if the converter is preceded by a PGA that amplifies small input signals so that the input signal stays within the full input range of the A/D converter.

REMARKS.

- (i) When looked as a regularization filter, the local adaptation mechanism (14) changes λ_1 and λ_2 according to v_k^1 and its local values so that they are described as $\lambda_1(v^1, k)$ and $\lambda_2(v^1, k)$, which are nonlinear.
- (ii) Equation (15) corresponds to a different, still linear though, regularization problem. Namely, the function minimized is of the form

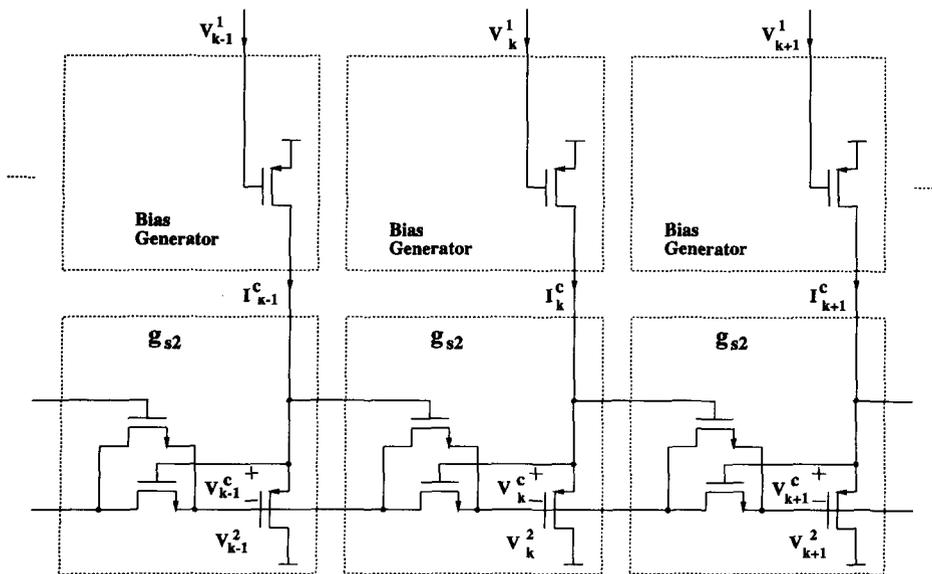


FIGURE 13. Circuit design of locally adaptive conductances g_{s2} and bias generators in Figure 12.

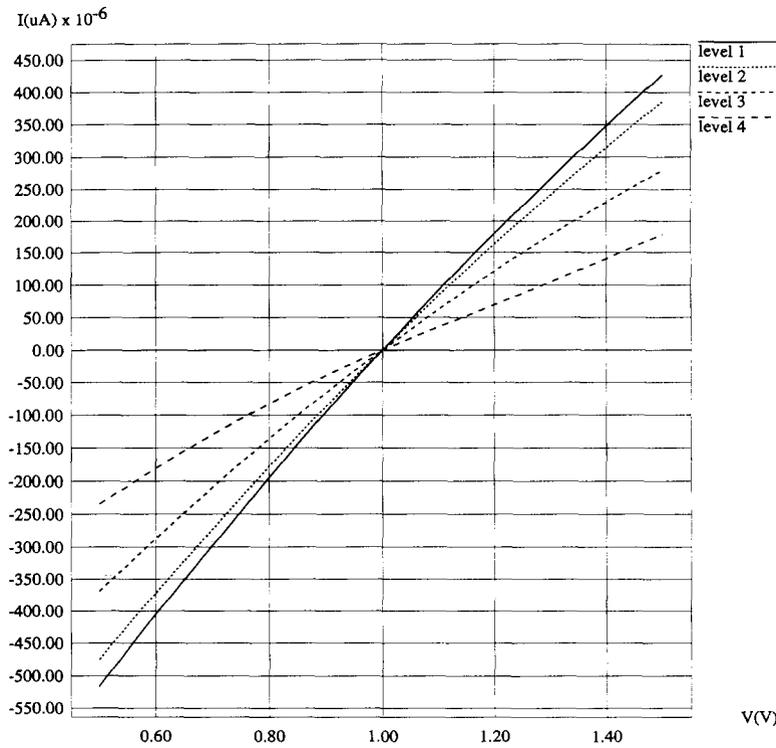


FIGURE 14. Simulation results of Figures 12 and 13. $V-I$ characteristics of $g_{s2(k,k+1)}$ are shown at several different values of $v_k^1 + v_{k+1}^1$. The higher the level the greater the value of $v_k^1 + v_{k+1}^1$.

$$G(v, \mathbf{d}^*(\mathbf{d})) = \|v - \mathbf{d}^*(\mathbf{d})\|^2 + \lambda_1 \|Dv\|^2 + \lambda_2 \|Lv\|^2,$$

where $\mathbf{d}^*(\mathbf{d})$ indicates eqn (15).

7. CONCLUSION

Light-adaptive architectures are proposed for regularization vision chips. Specifically, the architectures enable

controlling the filter width of a ∇^2G -like filter in an adaptive manner in accordance with the input image. Explanations are given of how this can be viewed as a regularization problem where the regularization parameter adapts to the intensity of input images. CMOS circuits are proposed to implement the adaptation mechanisms. The adaptation algorithm has been in-

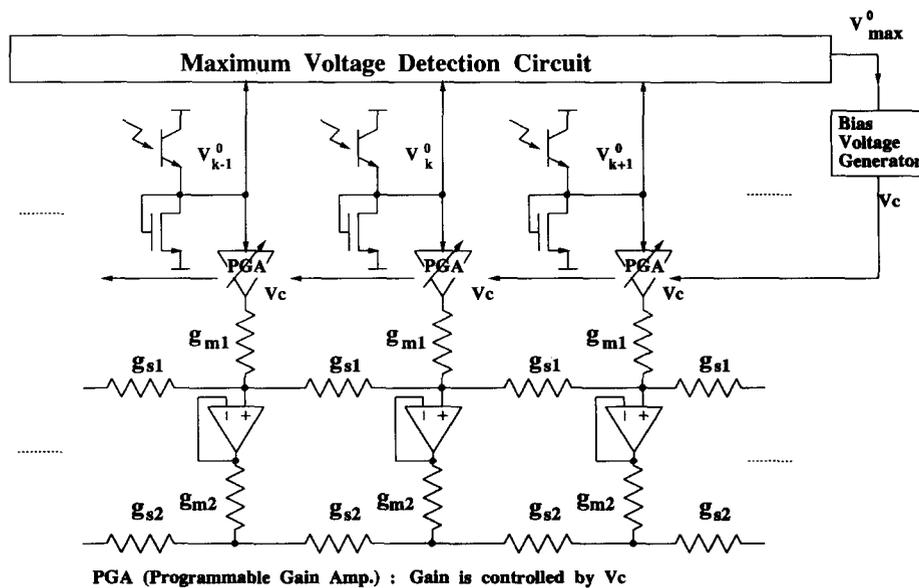


FIGURE 15. Maximum value adaptation network.

spired by the adaptation mechanism of the horizontal cells in the lower vertebrate retina.

REFERENCES

- Banu, M., & Tsividis, Y. (1982). Floating voltage-controlled resistors in CMOS technology. *Electronics Letters*, **18**, 678–679.
- Baylor, D. A., Fuortes, M. G. F., & O'Bryan, P. M. (1971). Receptive fields of cones in the retina of the turtle. *Journal of Physiology*, **214**, 256–294.
- Boahen, K. A., & Andreou, A. G. (1992). A contrast sensitive silicon retina with reciprocal synapses. *Advances in Neural Information Processing Systems*, **4**, 764–772.
- Clark, J. (1989). Authenticating edges produced by zero-crossing algorithms. *IEEE Transactions of PAMI*, **11**, 43–57.
- Courant, R., & Hilbert, D. (1953). *Methods of mathematical physics*. New York: Interscience Publishers.
- Dowling, J. E., & Ehinger, B. (1978). Synaptic organization of the dopaminergic neurons in the rabbit retina. *Journal of Comparative Neurology*, **180**, 203–220.
- Harris, J. (1986). The coupled depth/slope approach to surface reconstruction. Masters thesis, MIT, Cambridge, MA.
- Harris, J. (1988). An analog VLSI chip for thin-plate surface interpolation. *IEEE Conference on Neural Information Processing Systems—Natural and Synthetic*.
- Harris, J., Koch, C., Luo, J., & Wyatt, J. Jr. (1989). Resistive fuses: Analog hardware for detecting discontinuities in early vision. In Mahowald, M. & C. Mead (Eds.), *Analog VLSI implementation of neural systems*. New York: Kluwer Academic.
- Heatin, R., Blevins, D., & Davis, E. (1990). A bit-serial VLSI array processing chip for image processing. *IEEE Journal of Solid-State Circuits*, **25**(2), 364–368.
- Hutchinson, J., Koch, C., Luo, J., & Mead, C. (1988). Computing motion using analog and binary resistive network. *IEEE Computer*, **21**, 52–63.
- Kobayashi, H., Matsumoto, T., Yagi, Y., & Shimmi, T. (1993). Image processing regularization filters on layered architecture. *Neural Networks*, **6**, 327–350.
- Kobayashi, H., White, J. L., & Abidi, A. A. (1990). An analog CMOS network for Gaussian convolution with embedded image sensing. *ISSCC Digest of Technical Papers*, 216–217, San Francisco.
- Kobayashi, H., White, J. L., & Abidi, A. A. (1991). An active resistor network for Gaussian filtering of images. *IEEE Journal of Solid-State Circuits*, **26**(5), 738–748.
- Lamb, T. D., & Simon, E. J. (1976). The relation between intracellular coupling and electrical noise in turtle photoreceptors. *Journal of Physiology*, **263**, 256–286.
- Liu, S. C., & Harris, J. (1989). Generalized smoothing networks in early vision. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, 184–191.
- MacKay, D. J. C. (1991). Bayesian methods for adaptive models. Doctoral dissertation, Caltech, Pasadena, CA.
- Marr, D., and Hildreth, E. (1980). Theory of edge detection. *Proceedings of Royal Society of London, B*, **207**, 187–217.
- Maruyama, M., Nakahira, H., Araki, T., Sakiyama, S., Kitao, Y., Aono, K., & Yamada, H. (1990). An image signal multiprocessor on a single chip. *IEEE Journal of Solid-State Circuits*, **25**(6), 1476–1483.
- Mathur, B., Lin, S. C., & Wang, H. T. (1985). Analog neural networks for focal-plane image processing. *SPIE*, **1242**, 141–151.
- Matsumoto, T., Kobayashi, H., & Togawa, Y. (1992). Spatial versus temporal stability issues in image processing neuro chips. *IEEE Transactions on Neural Networks*, **3**(4), 540–569.
- Matsumoto, T., Shimmi, T., Kobayashi, H., Abidi, A. A., Yagi, T., & Sawaji, T. (1992). A second order regularization vision chip for smoothing-contrast enhancement. *IJCNN 92*, Beijing.
- Mead, C. (1989). *Analog VLSI and neural systems*. Reading, MA, Addison-Wesley.
- Mead, C., & Mahowald, M. (1988). A silicon model of early visual processing. *Neural Networks*, **1**(1), 91–97.
- Naka, K-I., & Rushton, W. A. H. (1967). S-potential from luminosity units in the retina of the fish (*Cyprinidae*). *Journal of Physiology*, **192**, 437–461.
- Ohshima, S., Yagi, T., and Funahashi, Y. (in press). Computational studies on the interaction between red cone and H1 horizontal cell.
- Poggio, T., Torre, V., & Koch, C. (1985). Computational vision and regularization theory. *Nature*, **317**, 314–319.
- Poggio, T., Voorhees, H., & Yuille, A. (1985). A regularized solution to edge detection. *MIT AI Lab., Memo 833*.
- Ruetz, P. A., & Brodersen, R. W. (1987). Architectures and design techniques for real time image processing ICs. *IEEE Journal of Solid-State Circuits*, **22**(2), 233–250.
- Schwartz, E. A. (1973). Responses of single rods in the retina of the turtle. *Journal of Physiology*, **232**, 503–514.
- Shigematsu, T., & Yamada, M. (1988). Effects of dopamine on spatial properties of horizontal cells in the retina of the goldfish. *Neuroscience Research Supplement*, **8**, s69–s80.
- Shimmi, T., Kobayashi, H., Yagi, T., Sawaji, T., Matsumoto, T., & Abidi, A. A. (1992). A parallel analog CMOS signal processor for image contrast enhancement. *Proceedings of European Solid-State Circuits Conference*.
- Teranishi, T., Negishi, K., & Kato, S. (1983). Dopamine modulates S-potential amplitude and dye-coupling between external horizontal cells in carp retina. *Nature*, **301**, 234–246.
- Tessier-Lavigne, M., & Attwell, D. (1988). The effect of photoreceptor coupling and synaptic nonlinearity on signal: Noise ratio in early visual processing. *Proceedings of the Royal Society of London, B*, **234**, 171–197.
- Tikhonov, A. N. (1963a). Solution of incorrectly formulated problems and the regularization method. *Soviet Mathematics Doklady*, **4**, 1035–1038.
- Tikhonov, A. N. (1963b). Regularization of incorrectly posed problems. *Soviet Mathematics Doklady*, **4**, 1624–1627.
- Tikhonov, A. N. (1965). Nonlinear equations of the first kind. *Soviet Mathematics Doklady*, **6**, 559–562.
- Vittoz, E. A. (1985). Micropower techniques, design of MOS VLSI circuits for telecommunications. Y. Tsividis & P. Antognetti, (Eds.), Englewood Cliffs, NJ: Prentice-Hall.
- Whaba, G. (1987). Three topics in ill-posed problems. In H. W. Engl & C. W. Groetsch (Eds.), *Inverse and ill-posed problems*. New York: Academic Press.
- Yagi, T. (1986). Interaction between the soma and the axon terminal of retinal horizontal cells in *Cyprinus Carpio*. *Journal of Physiology*, **375**, 121–135.
- Yagi, T., & Kaneko, A. (1987). Membrane properties and signal conduction of the horizontal cell syncytium of teleost retina. *Neuroscience Research Supplement*, **6**, s119–s132.
- Yagi, T., Ariki, F., & Funahashi, Y. (1989). Dynamic model of dual layer neural network for vertebrate retina. *Proceedings of IJCNN*, **1**, 787–789.