VLSI Memo No. 89-551
June 1989

Resistive Fuses: Analog Hardware for Detecting Discontinuities in Early Vision

John Harris, Christof Koch, Jin Luo, and John Wyatt

Abstract

The detection of discontinuities in motion, intensity, color, and depth is a well studied but difficult problem in computer vision. We discuss our "resistive fuse" circuit--the first hardware circuit that explicitly implements either analog or binary line processes in a controlled fashion. We have successfully designed and tested an analog CMOS VLSI circuit that contains a 1-D resistive network of fuses implementing piece-wise smooth surface interpolation. The segmentation ability of this network is demonstrated for a noisy step-edge input.

We derive the specific current-voltage relationship of the resistive fuse from a number of computational considerations, closely related to the early vision algorithms of Koch, Marroquin and Yuille (1986) and Blake and Zisserman (1987). We discuss the circuit implementation and the performance of the chip. In the last section, we show that a model of our resistive network--in which the resistive fuses have no internal dynamics--has an associated Lyapunov function, the co-content. The network will thus converge, without oscillations, to a stable solution, even in the presence of arbitrary parasitic capacitances throughout the network.

Research Center
Massachusetts Institute of Technology
Cambridge Massachusetts 02139
Telephone (617) 253-8138

Acknowledgements

To appear in *Analog VLSI Implementations of Neural Systems*, C. Mead, and M. Ismail, eds., Kluwer, Norwell, MA, 1989. This research was supported in part by a NSF under grant number IST-8700064, the Office of Naval Research Young Investigator, the National Science Foundation Presidential Young Investigator Award, the James S. McDonnell Foundation, Rockwell International Science Center, Hughes Aircraft Artificial Intelligence Center, the Defense Advanced Research Projects Agency under contract number N00014-87-K-0825, the National Science Foundation under contract number MIP-8814612, and the DuPont Corporation.

Author Information

Harris, Koch, and Luo: Computation and Neural Systems Program, California Institute of Technology, Pasadena, CA 91125.
Wyatt: Department of Electrical Engineering and Computer Science, Room 36-864, MIT, Cambridge, MA 02139. (617) 253-6718.

To appear in
Analog VLSI Implementations of Neural Systems, Mead, C. and Ismail, M. eds., Kluwer, Norwell, MA. (1989).

RESISTIVE FUSES: ANALOG HARDWARE FOR DETECTING DISCONTINUITIES IN EARLY VISION

John Harris, Christof Koch, Jin Luo
Computation and Neural Systems Program
California Institute of Technology
Pasadena, California 91125

and

John Wyatt
Department of Electrical Engineering and Computer Science
and Research Laboratory of Electronics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

Abstract: The detection of discontinuities in motion, intensity, color, and depth is a well studied but difficult problem in computer vision. We discuss our "resistive fuse" circuit—the first hardware circuit that explicitly implements either analog or binary line processes in a controlled fashion. We have successfully designed and tested an analog CMOS VLSI circuit that contains a 1-D resistive network of fuses implementing piece-wise smooth surface interpolation. The segmentation ability of this network is demonstrated for a noisy step-edge input.

We derive the specific current-voltage relationship of the resistive fuse from a number of computational considerations, closely related to the early vision algorithms of Koch, Marroquin and Yuille (1986) and Blake and Zisserman (1987). We discuss the circuit implementation and the performance of the chip. In the last section, we show that a model of our resistive network—in which the resistive fuses have no internal dynamics—has an associated Lyapunov function, the co-content. The network will thus converge, without oscillations, to a stable solution, even in the presence of arbitrary parasitic capacitances throughout the network.

INTRODUCTION

Most early vision algorithms incorporate the generic constraint that variables such as surface orientation and reflectance, depth or optical flow vary slowly in space (Marr and Poggio, 1976; Grimson, 1981; Ikeuchi and Horn, 1981; Horn and Schunck, 1981; Terzopoulos, 1983; Hildreth, 1984; Poggio,

Voorhees and Yuille, 1985; Nagel, 1987). Within the standard regularization approach, this is reflected in the use of stabilizing operators corresponding to various measures of smoothness (Poggio, Torre and Koch, 1985). Thus, in the problem of interpolating a 2-D surface through sparse and noisy depth measurement, the final surface should be as close as possible to the initial data as well as being as smooth as possible (Grimson, 1981); or, in the problem of computing optical flow from the time-varying intensity, the final flow field should be compatible with the locally measured velocity data as well as being smooth (Horn and Schunck, 1981; Hildreth, 1984; Nagel, 1987). However, surfaces display discontinuities where the smoothness constraint is violated. Thus, the to-be-reconstructed surface may have been generated by an underlying piece-wise smooth or even piece-wise constant depth distribution. Or, the 2-D velocity field induced by a rigid object moving/rotating in an otherwise stationary environment varies smoothly across the surface of the object but is zero beyond the contours of the object (since the background is stationary).

In the last years, a number of researchers have introduced powerful algorithms to deal with the representation of such discontinuities. Geman and Geman (1984) first proposed binary line processes to model discontinuities in intensity within the stochastic framework of Markov Random Fields. Discontinuities are subject to various constraints, such that they should form along continuous contours, should not intersect nor form parallel lines. Their approach was extended and modified to account for discontinuities in depth, texture and color by Poggio and his collaborators (Marroquin, Mitter and Poggio, 1984; Poggio, Gamble and Little, 1988) as well as to discontinuities in the optical flow (Hutchinson, Koch, Luo and Mead, 1988). The principal drawback of the Geman and Geman-type method is the computational expense involved in minimizing the associated non-convex cost functionals using stochastic optimization methods, in particular when numerous constraints (e.g. continuity of discontinuities) are incorporated. A number of authors have used deterministic methods to find the (local) minimum of the associated convex or non-convex variational functionals, with next-to-optimal results (Terzopoulos, 1986; Koch, Marroquin and Yuille, 1986). A rigorous deterministic approach has been championed by Blake and Zisserman (1987). Their "graduated non-convexity" (GNC) algorithm bears many similarities to the above methods, and leads to excellent results in the case of piece-wise continuous reconstruction of surfaces (Blake, 1989).

Poggio and Koch (1985) show how standard regularization algorithms can map onto simple resistive networks. Finding the minimum of the standard regularized and quadratic cost functional is equivalent to finding the state of least power dissipation in an appropriate electrical network, where the data are given by injecting current into certain nodes and the solution by the stationary voltage distribution. Figure 1 shows the appropriate network for membrane-type surface interpolation, where the "strength" of smoothing is given by the

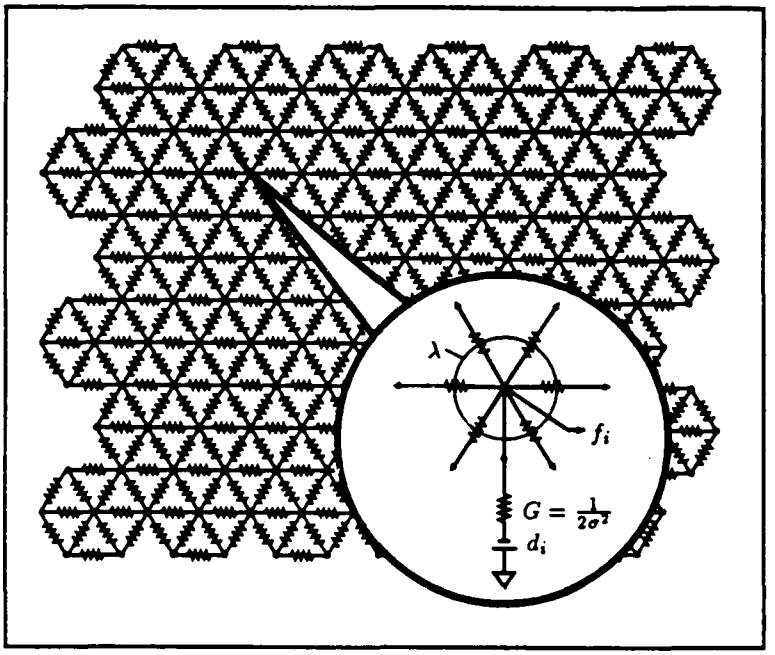

Figure 1 Resistive network for fitting the smoothest surface f through sparse and noisy data d. The circuit minimizes the variational functional of the two-dimensional extension of eq. (1) in the absence of line discontinuities. In the continuum limit, minimization of this functional corresponds to the Euler-Lagrange equation $\lambda\nabla^2 f + Gf = Gd$. The battery supplies the measured depth data d_i, while the vertical conductance G corresponds to $1/(2\sigma^2)$ and the horizontal conductance of the grid to λ. If no data are present at a particular location i, G is set to zero. The stationary voltage distribution then corresponds to the interpolated surface f_i. The amplitude of the horizontal grid conductance, λ, controls the amount of smoothing. A 48 by 48 pixel hexagonal network has been built and tested successfully (Luo, Koch and Mead, 1988).

value of the horizontal grid conductance. For an overview of analog circuits for implementing early vision algorithms see Koch (1989) and Horn (1989).

The recent development of subthreshold, analog CMOS VLSI circuits for various sensory tasks by Carver Mead (see in particular his recent textbook, Mead, 1989) has enabled us to implement these resistive networks—together with the photo-transduction stage—using this real-time, low power and robust

technology. Two circuits are particularly attractive for our purposes: a phototransistor with a logarithmic voltage output over five orders of intensity brightness (Mead, 1985, 1989) and a transistor circuit with a linear current-voltage relationship for small voltage gradients (Sivilotti, Mahowald and Mead, 1987; Mead, 1989). The value of the slope, i.e. the resistance, can be varied over five orders of magnitude. Using this as our basic construction element, we built and tested a 48 by 48 pixel resistive network for smoothing and interpolating noisy and sparse data (Luo, Koch and Mead, 1988; see Fig. 1).

We introduce in this paper an analog, purely deterministic approach to locating discontinuities in the case of interpolating noisy and sparsely sampled depth data. It leads to a very simple and elegant circuit implementation in terms of a two-terminal, nonlinear, voltage-controlled resistor termed "resistive fuse" (Harris and Koch, 1989). We have implemented this device in analog CMOS and demonstrate its performance here.

THEORY

Let us begin by justifying "resistive fuses" as specialized circuit elements for implementing discontinuities. Since our methodology does not distinguish between a 1-D and a 2-D implementation of smoothing in the presence of discontinuities, we will first consider the 1-D case. The simplest possible variational functional for interpolating noisy and sparsely sampled data d_i in the presence of binary line discontinuities ℓ_i is a membrane type of surface interpolation:

$$J(f,\ell) = \lambda \sum_i (f_i - f_{i+1})^2 (1 - \ell_i) + \frac{1}{2\sigma^2} \sum_i (d_i - f_i)^2 + \alpha \sum_i \ell_i, \qquad (1)$$

where f_i is the value of the final surface f at location i, σ^2 the variance of the additive Gaussian noise process assumed to corrupt the data d_i and λ and α are free parameters. The first term in this functional implements the constraint that surfaces should, in general, vary smoothly. If all variables, with the exception of f_i, f_{i+1} and ℓ_i, in eq. (1) were held fixed and $\lambda(f_i - f_{i+1})^2 < \alpha$, it would be "cheaper" to pay the price $\lambda(f_i - f_{i+1})^2$ and set $\ell_i = 0$ than to pay the larger price α. However, if the gradient becomes too steep, the line process is switched on, i.e. $\ell_i = 1$, and the "price" α is paid. The second term in eq. (1), where the sum only includes those locations i where data exist, forces the final solution f to be close to the measured data d. How close depends on the estimated magnitude of the noise, in this case on σ^2. Thus, the surface f, with its associated set of discontinuities ℓ, minimizing eq. (1) will be the one that best satisfies the conflicting demands of piece-wise smoothness and fidelity to the measured data. The functional of eq. (1) is non-convex and a large number of both stochastic and deterministic methods have been designed to find optimal or nearly optimal solutions for this and similar functionals (Geman and Geman,

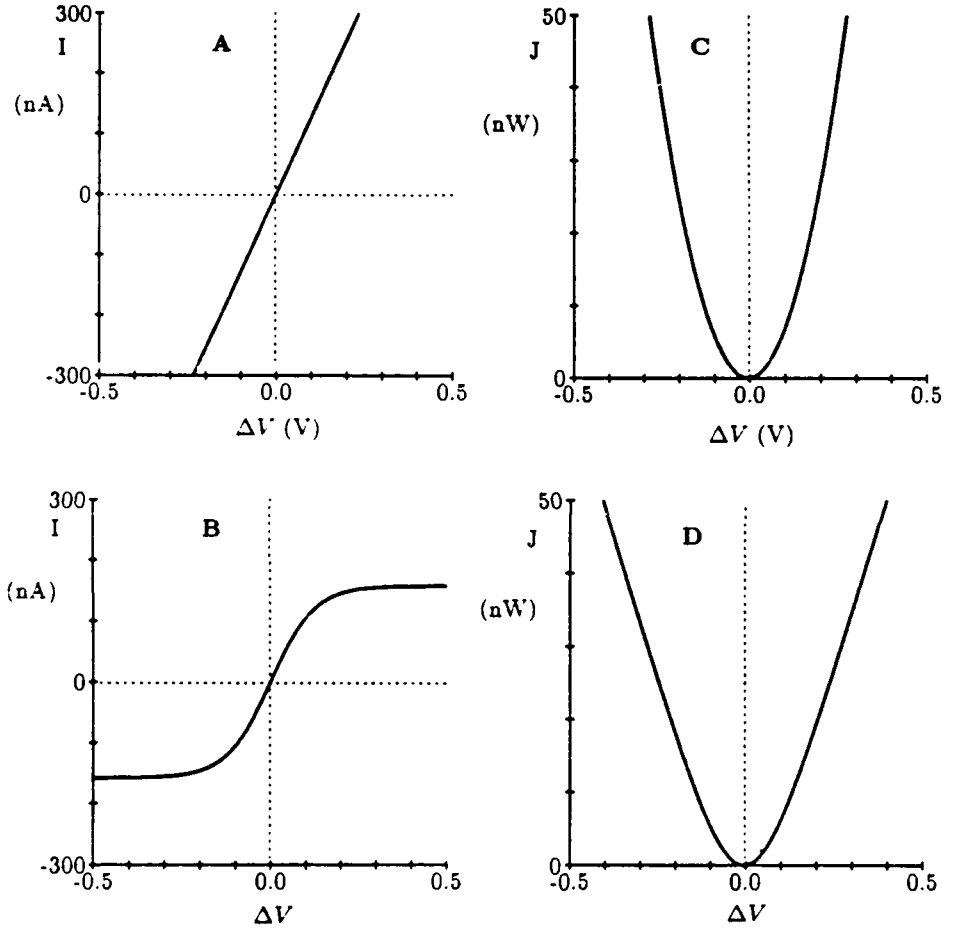

Figure 2 Theoretical I-V curves for a linear resistor (A) and a measured I-V curve for Mead's saturating resistor (B). Integrating numerically over these curves gives the co-content of the linear resistor (C) and the saturating resistor (D). Co-content is defined by eq. (2) and represents generalized power for non-linear systems. The co-content for the linear resistor is equivalent to half the dissipated power, and thus a quadratic function in ΔV, while the co-content for the saturating resistor becomes a linear function of ΔV as $|\Delta V| \rightarrow \infty$.

1984; Marroquin, Mitter and Poggio, 1987; Koch, Marroquin and Yuille, 1986; Blake and Zisserman, 1987; Terzopoulos, 1983, 1986).

Figure 3C shows a plot of $J(f, \ell)$ as a function of the depth at locations f_i and f_{i+1} and as a function of the discontinuity ℓ_i. The values of the surface and of the line discontinuities are assumed to be fixed at all other locations. As long as $\lambda(f_i - f_{i+1})^2 \leq \alpha$, the function $\mathcal{J}$ is quadratic in the gradient. However, once $|f_i - f_{i+1}|$ exceeds the gradient limit $\sqrt{\alpha/\lambda}$, $\mathcal{J}$ remains flat at $\mathcal{J} = \alpha$, independent of the magnitude of $f_i - f_{i+1}$ (Blake and Zisserman, 1987).

The appropriate circuit implementation is a straightforward modification of the network shown in Fig. 1. The surface f_i represents the final reconstructed points. The voltage on the battery is d_i, and the conductance G equals $1/(2\sigma^2)$. If no measured surface value d is present at a particular location, $G = 0$ at that location. The value of the grid conductance λ controls the amount of smoothing. Binary switches, breaking the resistive connections among neighboring nodes, would implement discontinuities in the surface. As long as the switch is closed, the current is linear in the voltage drop across the device. Since the electrical power in a linear network is proportional to the square of the voltage gradient across all resistances, the power is quadratic in the gradient and can thus be identified with the first term in eq. (1). Once the threshold has been exceeded, the binary switch opens and no more current flows through the device. The digital processors controlling the switches need access to the state of the neighboring switches as well as to the neighboring depth values. We will now demonstrate, however, how this mixed analog-digital circuit can be replaced by a single analog non-linear resistor, the "resistive fuse."

The circuit implementation of binary discontinuities will require nonlinear circuit components. As pointed out by Poggio and Koch (1985), the notion of minimizing power in linear networks implementing quadratic "regularized" algorithms must be replaced by the more general notion of minimizing the total resistor co-content (Millar, 1951). For a two-terminal voltage-controlled resistor characterized by $I = f(V)$, the co-content is defined as

$$J(V) = \int_0^V f(V')dV'. \tag{2}$$

For a linear resistor, $I = GV$, the co-content is given by $\frac{1}{2}GV^2$, which is just half the dissipated power $P = GV^2$ (Fig. 2). For a network consisting of a collection of resistors, voltage sources and other elements, the total network co-content is defined as the sum of all the (linear or nonlinear) resistor co-contents, that is,

$$J_{total}(t) = \sum_{\text{all resistors}} J_k(V_k(t)). \tag{3}$$

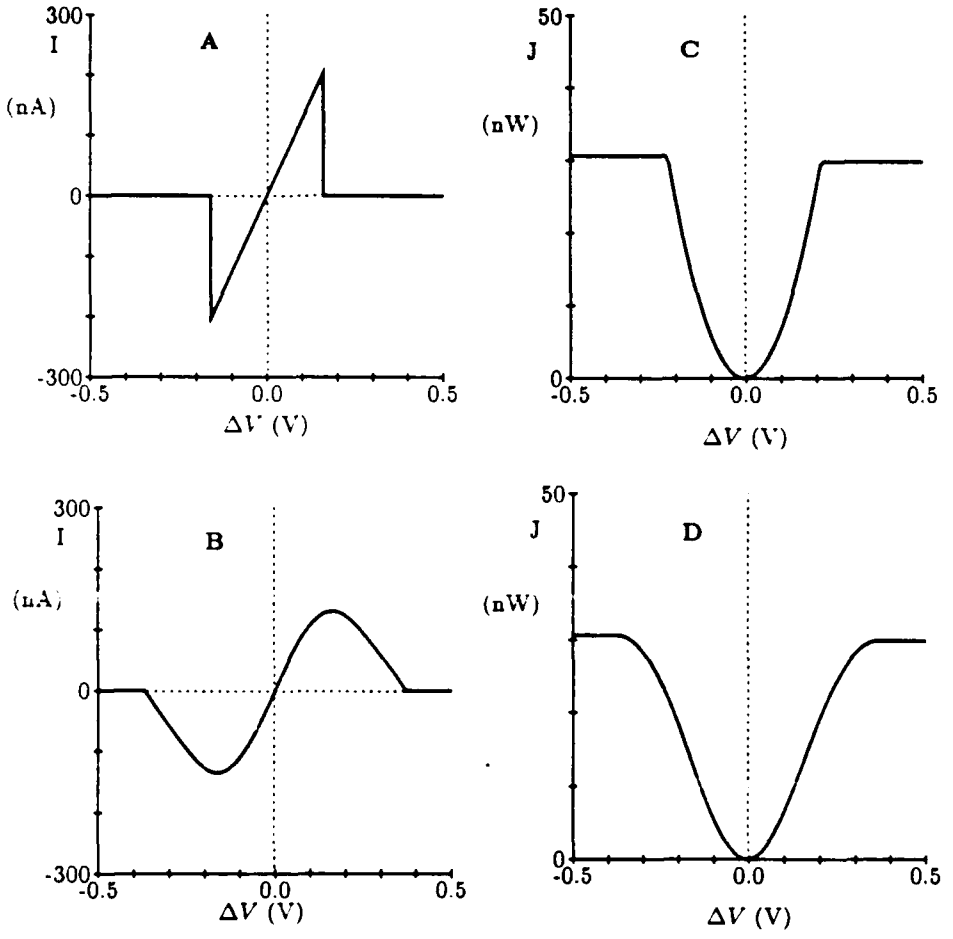

Figure 3 Theoretical I-V curve for an infinite-gain fuse (A) and a measured I-V curve for a finite-gain resistive fuse (B). Integrating numerically over these curves gives the co-content J for the infinite-gain (C) and the finite-gain fuse (D).

The co-content for various resistors is plotted in Figs. 2 and 3. Differentiating eq. (2), we have:

$$f(V) = \frac{dJ}{dV}. \tag{4}$$

The appropriate current-voltage relationship of an *infinite-gain resistive fuse* is illustrated in Fig. 3A. As long as the voltage drop across this device is below the threshold, the current through the nonlinear resistor is linearly related to the voltage across it. Once past the threshold, the circuit breaks (hence the name "fuse"), and the current is zero for all values of the voltage gradient. This two-terminal device then implements the high-level constraint that surfaces should be smooth unless their neighboring values differ by more than $\pm\sqrt{\alpha/\lambda}$, at which point the surfaces will break.

The I-V relationship of the device we have built is shown in Fig. 3B. The most salient difference from the infinite-gain fuse are the smooth flanks, where the current decreases smoothly to zero for increasing values of the voltage gradient [1], in contrast with the discontinuity in the I-V relationship for the infinite-gain fuse. In this region the slope conductance dI/dV will be negative (Fig. 13C). Our measured I-V curve can be related directly to the concept of analog line discontinuities of Koch *et al.* (1986). The key idea is that, following Hopfield and Tank (1985) in their neural network implementation of the Traveling Salesman Problem, binary discontinuities are mapped onto continuous "neurons," whose output is constrained to lie between 0 and 1. The input-output relationship of these "discontinuity neurons" is governed by the sigmoidal function $V = g(U)$, where $g(U)$ is a strictly monotonic function, usually taken to be

$$g(U) = \frac{1}{1 + e^{-2\eta U}}, \tag{5}$$

with the "gain" $\eta > 0$. The network converges to a stationary solution using a steepest descent rule. The solutions obtained were qualitatively very similar to the solutions obtained with binary line processes. It is rather straightforward to derive an "analog" version of resistive fuses (Harris, Koch, Staats, Luo and Wyatt, 1989), with the following I-V relationship

$$I = f(V) = \left[1 - g(\frac{V^2 - \alpha}{\beta})\right] V, \tag{6}$$

where $\beta > 0$ is a parameter related to the analog line process implementation (identical to c_G of eq. (7c) of Koch *et al.*, 1986). Our measured I-V curve for the fuse (Fig. 3B) implements this function. As $\beta \to 0$ the function g

[1] The I-V characteristic of our experimental fuse relates somewhat to the theoretical work of Perona and Malik (1988) who simulated a network of elements with similar I-V characteristics to perform image segmentation.

becomes binary and $f(V)$ of eq. (6) approaches the form of the infinite-gain fuse (Fig. 3A).

So far we have only discussed the implementation of binary or analog discontinuities in 1-D. For 2-D image problems, horizontal as well as vertical line processes need to be incorporated into the variational functional. Furthermore, it has been standard practice to constrain the geometry of line processes by adding appropriate terms to the 2-D extension of eq. (1). Some of the more common constraints are that discontinuities should occur along continuous contours, should not intersect nor form along parallel lines (Geman and Geman, 1984). Furthermore, Poggio *et al.* (1988) introduced the notion that discontinuities in depth should in general coincide with discontinuities in intensity, that is intensity edges.

We previously demonstrated how a piece-wise smooth optical flow field, induced by moving objects, can be successfully recovered in the presence of binary motion discontinuities with the above set of constraints (Koch *et al.*, 1986; Hutchinson *et al.*, 1988). We repeated these simulations using only the finite-gain resistive fuses of eq. (6) together with the constraint that motion discontinuities should only occur together with intensity discontinuities, in our case zero-crossings of the $\nabla^2 G$ operator. The performance of both algorithms—for 128 by 128 video image sequences of several moving and partially occluding people—is very similar (for more details see Harris *et al.*, 1989). Since the co-localization of all or most motion discontinuities with intensity discontinuities (but not necessarily the reverse) is relatively simple to implement at the circuit level, we feel that we can now design VLSI circuits to compute intensity, motion and depth discontinuities for real, two-dimensional images. The following section discusses the detailed circuit implementation of the resistive fuse.

CIRCUIT DETAILS

The circuit schematic for the fuse is shown in Fig. 4. The circuitry above the dotted line in the figure is Mead's saturating resistor (Mead, 1989) with a p-type pullup transistor that sets the nominal resistance of the fuse. In subthreshold operation, the current through a transistor varies exponentially with the gate-to-source voltage. Thus, the voltage V_B produces a current I_B equal to:

$$I_B = I_0 e^{\kappa(V_{DD} - V_B)} \tag{7}$$

Following Mead (1989), all voltages are assumed to be normalized by kT/q. The variable κ is a process-dependent parameter that reflects the inability of the gate to be 100% effective in reducing the barrier potential. I_0 is a constant that includes the width and length of the transistor as well as process-dependent

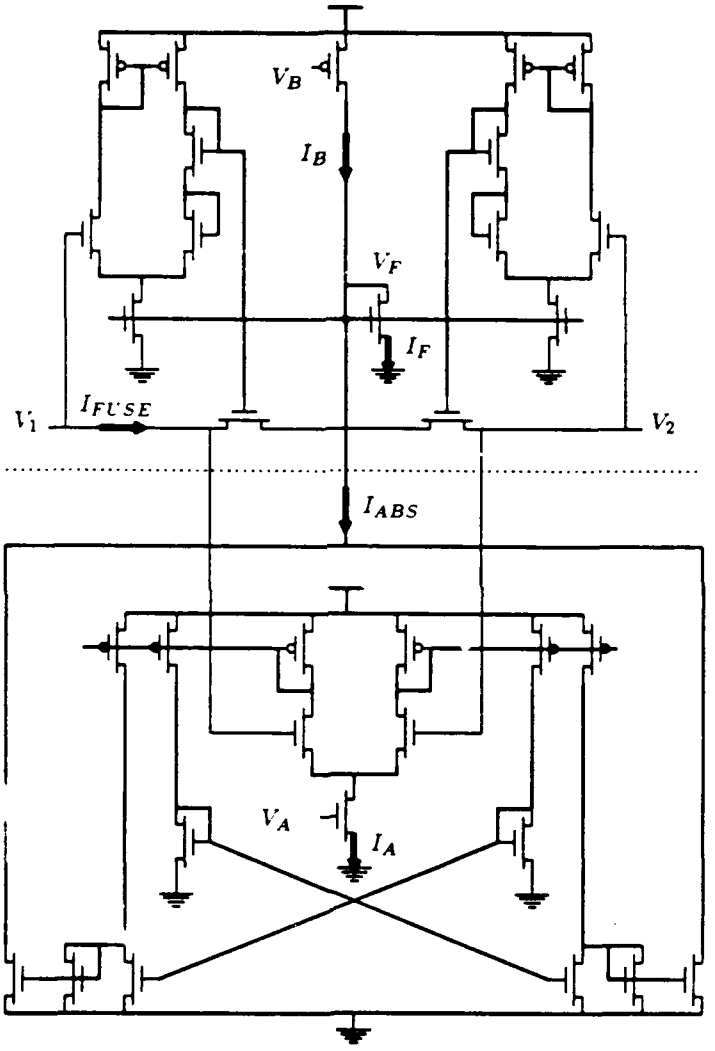

Figure 4 Schematic of the fuse circuit. The nonlinear, voltage-controlled resistance is seen across the V_1 and V_2 terminals. The circuitry above the dotted line is a saturating resistor (Mead, 1989) with V_B controlling the nominal amount of resistance. The circuit below the dotted line is a saturating absolute-value circuit that turns off the resistor for large $|V_1 - V_2|$. V_A determines the magnitude of the current pulled away by the absolute-value circuit.

fabrication parameters. Letting $I_F = I_B$, the I-V relation of the resistor can be derived as:

$$I_{FUSE} = \frac{I_F}{2} \tanh\left(\frac{\Delta V}{2}\right) \tag{8}$$

where $\Delta V = V_1 - V_2$. For small ΔV this portion of the circuit operates as a linear resistor with a resistance of

$$R = \frac{4kT/q}{I_F} \tag{9}$$

Because we are working in the subthreshold region, I_F and thus the resistance can be varied over five orders of magnitude. For large ΔV the resistor saturates and provides a constant current of $I_F/2$. A measured I-V curve for this circuit is shown in Fig. 2B.

The circuit below the dotted line in the figure performs a saturating absolute-value operation. This portion of the circuit is enabled by the voltage V_A, which creates a current I_A equal to:

$$I_A = I_0 e^{\kappa V_A} \tag{10}$$

The positive parts of the outputs of a dual-output wide-range transconductance amplifier are combined to create a current of:

$$I_{ABS} = I_A \tanh\left(\frac{\kappa|\Delta V|}{2}\right) \tag{11}$$

By Kirchhoff's current law, the current I_F is:

$$I_F = \lfloor I_B - I_{ABS} \rfloor \tag{12}$$

where the symbols $\lfloor\ \rfloor$ are defined as

$$\begin{aligned} \lfloor x \rfloor &= x \quad \text{if} \quad x \geq 0 \\ &= 0 \quad \text{if} \quad x < 0 \end{aligned}$$

Substituting (11) and (12) into eq. (8), gives

$$I_{FUSE} = \frac{1}{2}\left\lfloor I_B - I_A \tanh\left(\frac{\kappa|\Delta V|}{2}\right)\right\rfloor \tanh\left(\frac{\Delta V}{2}\right) \tag{13}$$

When $|\Delta V|$ is small, the fuse acts as a linear resistor whose nominal resistance is set by I_B. When $|\Delta V|$ is large, I_A increases above the current supplied by the p-type pull-up, and V_F is pulled to ground, shutting off the resistor. In between these extremes, the fuse exhibits a gradual transition.

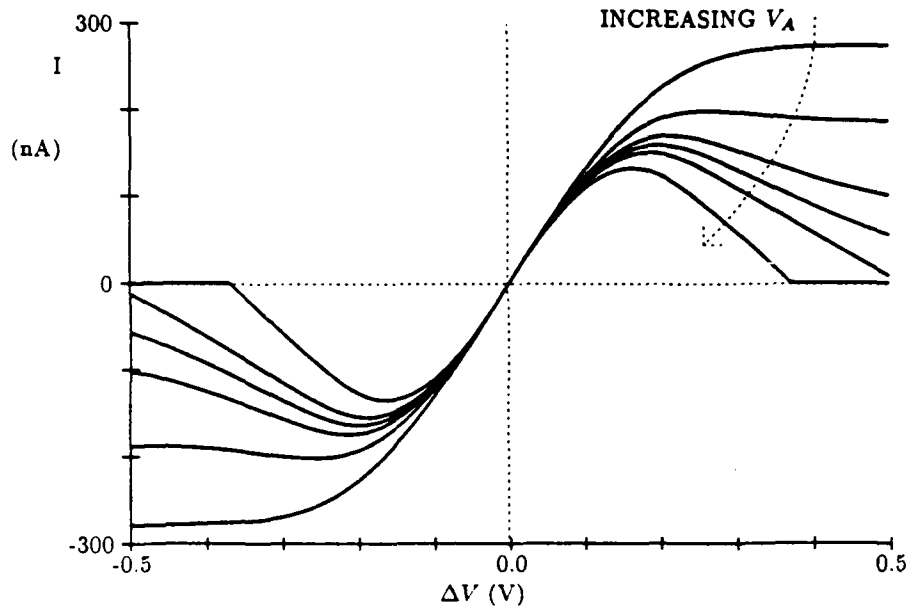

Figure 5 Measured I-V curves that show the effect of continuously varying from the saturating characteristic to that of the fuse curve. V_B was set to 4V and V_A was varied from 0V to 2V. When $V_A = 0$, the resulting I-V curve is identical to that of Mead's saturating resistor.

Figure 5 shows a family of curves measured by varying V_A while keeping V_B constant. By varying V_A in this way, the circuit's I-V characteristic can be continuously and smoothly changed from that of a saturating resistor to the fuse I-V curve. Setting $V_A = 0$ gives $I_A = 0$ disabling the absolute-value circuit, and giving the fuse a saturating I-V relationship (Fig. 2B).

Integration of the I-V curves in Fig. 5 gives the family of co-content curves shown in Fig. 6. For small ΔV the co-content is quadratic and for large ΔV the co-content saturates at a constant value. Instead of saturating for large voltage differences, the co-content of the saturating resistor increases linearly with voltage. As will be seen in the following section, networks of resistors with positively sloped I-V curves are guaranteed to converge to a single unique minimum value of the co-content. By turning the voltage control, we are changing the energy landscape in a continuous fashion ("continuation method") from containing one unique global minimum to a landscape containing many local minima.

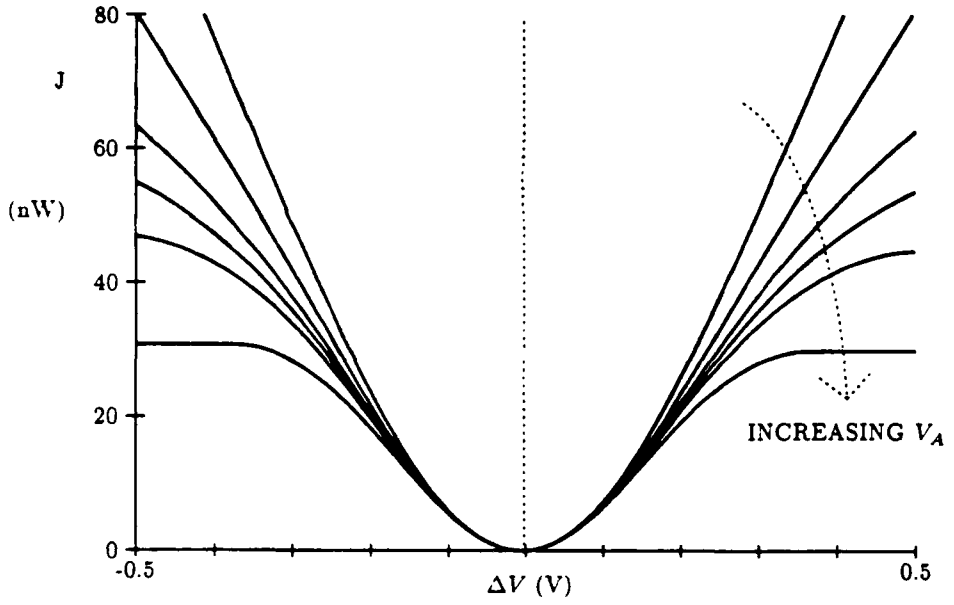

Figure 6 Co-content functions: each curve was numerically integrated from the family of curves in Fig. 5. Continuously varying the co-content curves in this way performs a useful computation that is explored more in Fig. 10 and Fig. 11.

The fuse provides a mechanism for changing the threshold value. If we assume that the circuit is operating in the linear region of the two hyperbolic tangents, I_{FUSE} becomes twin parabolas of the form:

$$I_{FUSE} = \left\lfloor \frac{I_B}{4} - \kappa \frac{I_A}{8} |\Delta V| \right\rfloor \Delta V \tag{14}$$

This linear analysis indicates that the measured curve in Fig. 3B consists of a parabola in each of the first and third quadrants. This current in eq. (14) is cut to zero for:

$$|\Delta V| \geq 2 \frac{I_B}{I_A} \frac{kT}{q\kappa} \tag{15}$$

I_{FUSE} reaches extremum points at:

$$|\Delta V| = \frac{I_B}{I_A} \frac{kT}{q\kappa} \tag{16}$$

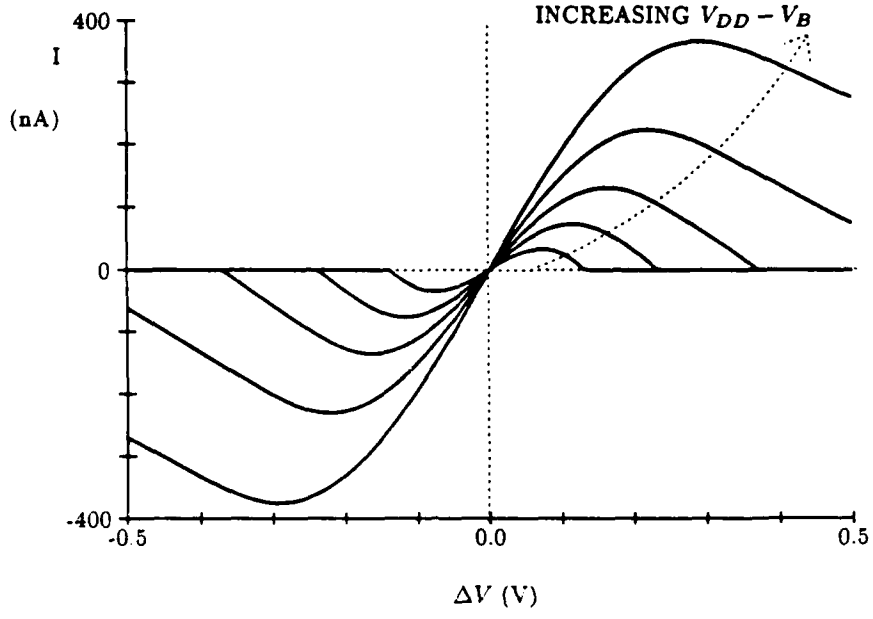

Figure 7 Measured I-V curves illustrating different line process penalties. V_A was kept constant at 2V and V_B was varied from 3.9V to 4.1V.

The extremum points can be set by the ratio of I_B to I_A. In subthreshold operation, the width of the saturating tanh curves is about 100mV. The extremum points can then only be be varied from 0 to about ±100mV. For gate voltages above the threshold of the bias transistors, the width of the linear region of the hyperbolic tangent function increases by $V_{GS} - V_T$, where V_{GS} is the gate-to-source voltage and V_T is the threshold voltage of the bias transistors. Thus, by going slightly above threshold the extremum point can be varied from 0 to about ±500mV. Figure 7 shows a family of I-V curves measured by varying V_B and holding V_A constant.

We are studying the use of a high-gain fuse, a circuit that does not have a large incrementally active region in its I-V curve (Fig. 8). Circuit simulations of the high-gain fuse show I-V curves that look like those of the infinite-gain fuse in Fig. 3A. Instead of feeding the absolute-value current back to the resistor bias circuits, current is fed back to a pass gate that acts as a binary switch in the current path. When $I_B > I_{ABS}$ the voltage on the gate of the binary switch (V_F) is charged to V_{DD}. On the other hand, when $I_B < I_{ABS}$, V_F is pulled to ground, effectively open-circuiting the resistor. The resistance of the resistor is controlled by V_R, which sets the bias current I_R. Notice that the current that controls the line process penalty is decoupled from the current that

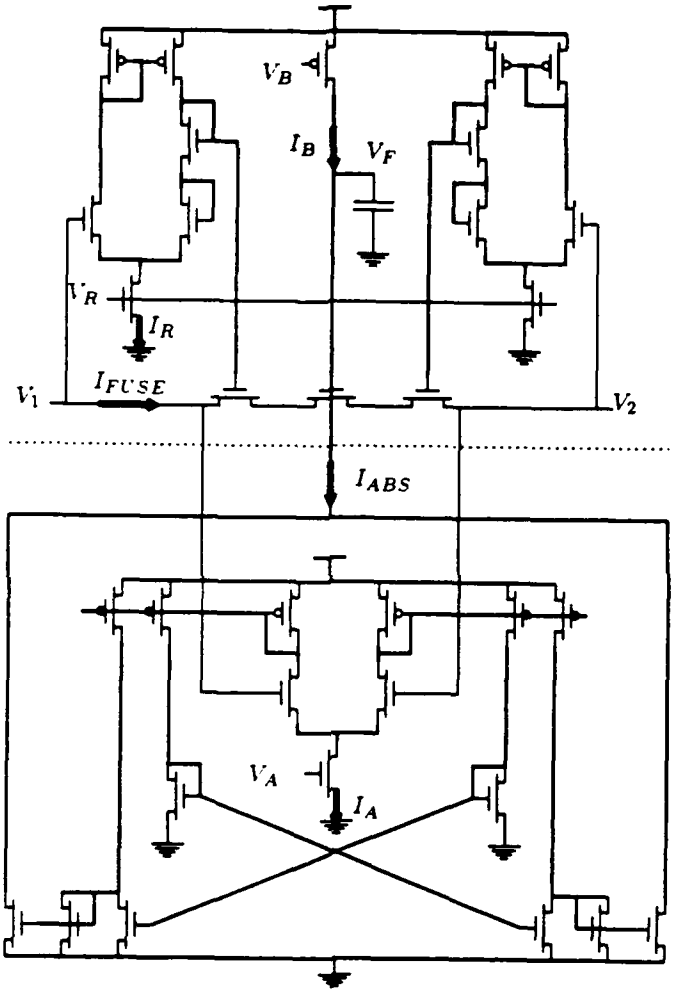

Figure 8 Modification of the fuse to obtain a high-gain characteristic. As before, a saturating resistor and an absolute-value circuit are combined to create a fuse. However, different from the circuit of Fig. 4, the absolute-value circuit discharges the gate of a pass transistor that has been added in the resistance path. This pass gate acts as a binary switch that is opened or closed dependent on whether or not the absolute-value current is greater than the threshold current provided by V_B. V_R provides independent control of the resistance of the fuse when the binary switch is closed.

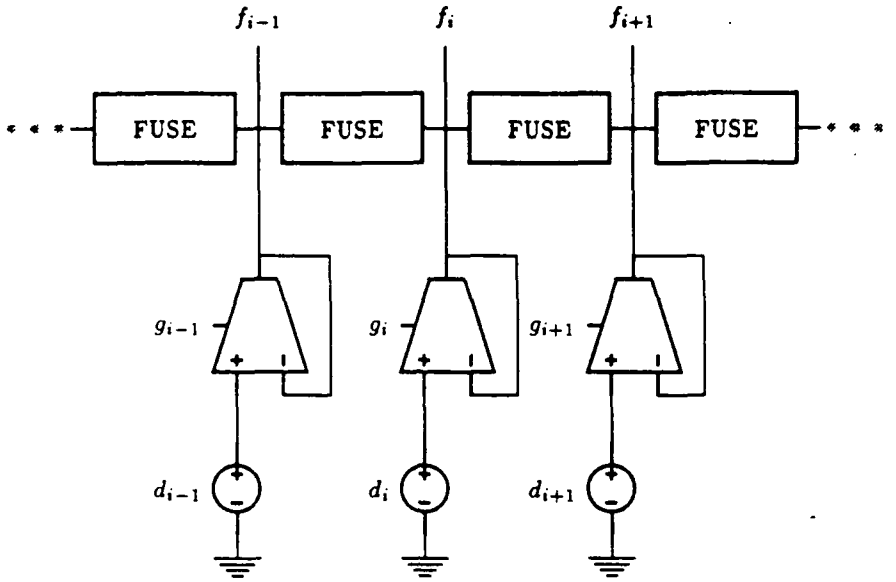

Figure 9 Layout of the 1-D fuse network. Voltage sources d_i provide input to the network through wide-range transconductance amplifiers. The bias voltages on these amplifiers g_i controls their conductance. The smoothed and segmented outputs are given as voltages at f_i. This network was designed to implement eq. (1).

sets the resistance of the fuse. Assuming high-gain elements, the I-V equation for the high-gain fuse is given by:

$$\begin{aligned} &\text{if } I_A \tanh\left(\frac{\kappa|\Delta V|}{2}\right) < I_B \text{ then } I_{FUSE} = \frac{I_R}{2}\tanh\left(\frac{\Delta V}{2}\right) \\ &\text{if } I_A \tanh\left(\frac{\kappa|\Delta V|}{2}\right) > I_B \text{ then } I_{FUSE} = 0 \end{aligned} \tag{17}$$

This implementation of the fuse shares an advantage with Mead's saturating resistor layout, because only one biasing circuit is needed for each node. This saves many transistors, especially in 2-D layouts. The low-gain fuse requires 33 transistors per connection, while the high-gain fuse requires only 21 transistors per connection plus 6 transistors per node. For a hexagonal mesh, each basic cell needs to contain one node plus half of the six neighboring connections, requiring a total of 69 transistors per cell for the high-gain fuse and 99 transistors per cell for the low-gain version.

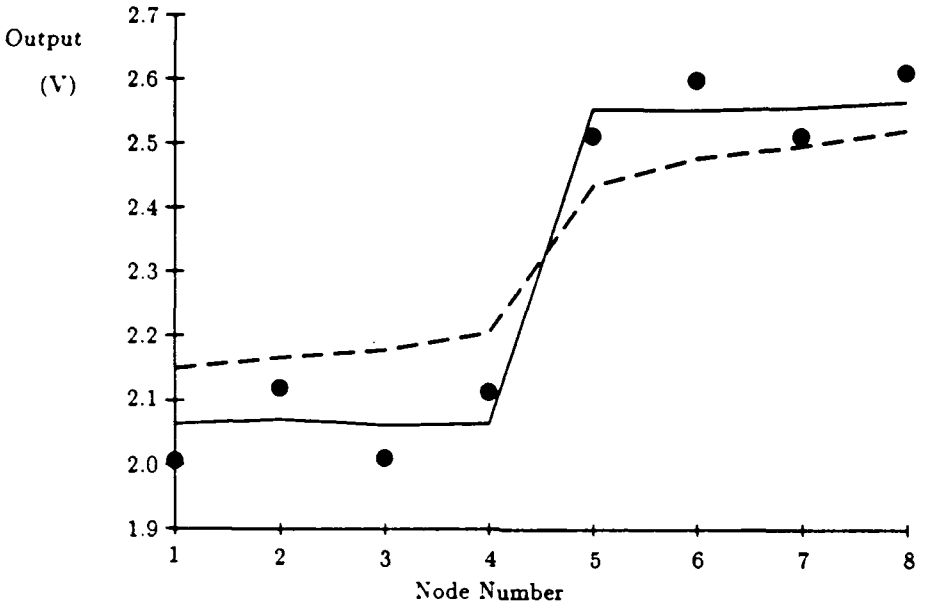

Figure 10 Measured segmentation from an experimental resistive fuse network. The circles denote "noisy" step data that was used as the input to the network. The solid-line curve indicates measured voltages from the chip. The dotted-line curve shows the measured voltage output given by a network of Mead's saturating resistors.

A network of eight fuses (of the type shown in Fig. 4) was fabricated and successfully demonstrated. The schematic is shown in Fig. 9. Eight voltage values are input as the d_i values. The smoothed and segmented f_i voltages are the resulting outputs. Figure 10 shows a segmentation result for a "noisy" 1-D step edge. The network effectively smooths out small steps without degrading large step edges. The I-V curves of the fuses in this example have been set to the form shown in Fig. 3B. In this configuration, the network exhibits a hysteresis property in which two stable final states are possible. The two stable states correspond to segmenting or smoothing the step edge. The segmented stable state is shown as the solid line in Fig. 10. The smoothed stable state becomes essentially a flat horizontal line. The final state depends on the temporal history of the network. To ensure that the proper stable state is reached in a deterministic fashion, V_A is initially set to 0V and then gradually moved to its final value.

The hysteresis properties of the network can be better understood through a load-line analysis of a much simplified circuit (Fig. 11). The current through

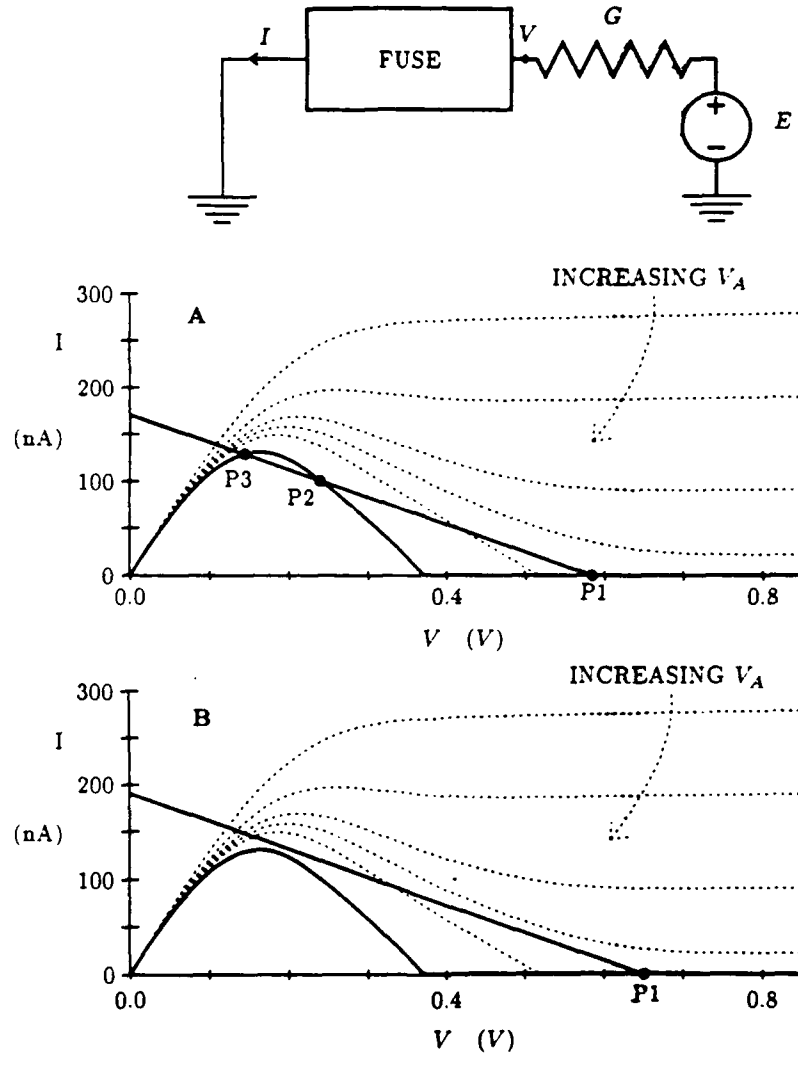

Figure 11 Simple load-line analysis shows that there can be up to three equilibrium points for the fuse/resistor circuit given above. The I-V curves for the measured fuse and the simulated voltage source/resistor are shown as solid lines. For plot A, points P1 and P3 are stable, and P2 is unstable. Voltages in the neighborhood of P2 will be driven to either P1 or P3. By increasing the value of the voltage source E, a single stable equilibrium point P1 remains (plot B). The dotted-line curves show the effect of changing V_A.

the fuse is plotted as a function of the voltage across the fuse. The simulated voltage source/resistor is also illustrated as a solid line, with the negative slope of this line given by the conductance G and the x-intercept given by the value of the voltage source E. A stability analysis reveals that the system possesses up to three equilibria. In the case illustrated in Fig. 11A, the middle equilibrium is unstable and the voltage will tend toward the two stable solutions P1 and P2. Point P1 corresponds to segmentation, and P3 corresponds to smoothing. By increasing the value of the voltage source E (Fig. 11B), only a single stable equilibrium point remains, corresponding to segmentation. Of course, stability cannot be guaranteed for negative values of G. The dotted-line curves show the effect of changing V_A.

Figure 12 shows the computed total co-content from the I-V curves shown in Fig. 11. For Fig. 12A, P1 is the global and P3 is only a local minimum, while P2 corresponds to an unstable local maximum. In contrast, Fig. 12B contains a single equilibrium point, P1, which corresponds to a discontinuity. The dotted lines show the effect of increasing V_A, deforming the energy surface from one with a single equilibrium point to one with two local minima. By using a continuation method in this fashion, discontinuities are deterministically located. Reasonable performance may be obtained by using a single setting of the fuse control voltages and keeping the voltages constant over time. This static approximation of the continuation method will still smooth small step edges while preserving large steps. However, medium steps, such as those simulated in Fig. 11, can be either smoothed or segmented depending upon the temporal history of the network. This load-line analysis is a simplified version of the true dynamics of networks of fuse elements, but serves to illustrate the complexity of even a single fuse element circuit.

STABILITY

Though the chord resistance of the fuse circuit is always positive, its incrementally negative resistance regions (see Fig. 13) raise doubts about the stability of networks of resistive fuse elements. One question that has already been alluded to above is the issue of whether the network will converge at all and whether a unique stationary solution exists. The reasoning presented later in this section supports the following conclusions.

1. Monotonic Resistors

Suppose all the nonlinear resistors are *incrementally strictly passive*, i.e., have I-V curves with positive slope, $dI/dV > 0$, everywhere. One instance of such a device is Mead's saturating resistor (Fig. 2B). Then the stationary network solution for a given input image will be unique. If we further suppose that the nonlinear resistors are ideal memoryless elements (i.e., that we can neglect the fast parasitic dynamics internal to each resistor circuit), then the network

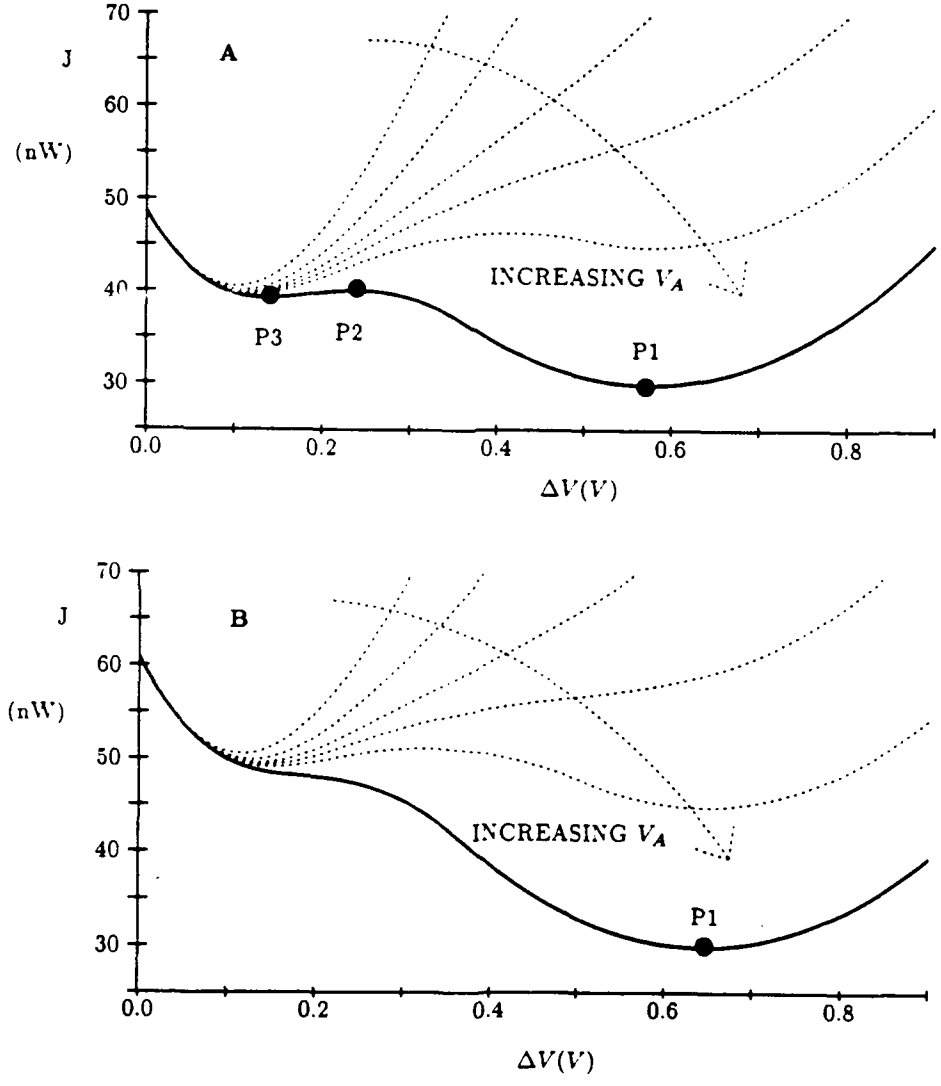

Figure 12 Computed total co-content from the I-V curves shown in Fig. 11. In plot A, P1 and P3 correspond to stable minima while P2 is an unstable maximum. In contrast, Plot B contains a single equilibrium point P1 that corresponds to a discontinuity. The dotted lines show the effect of increasing V_A.

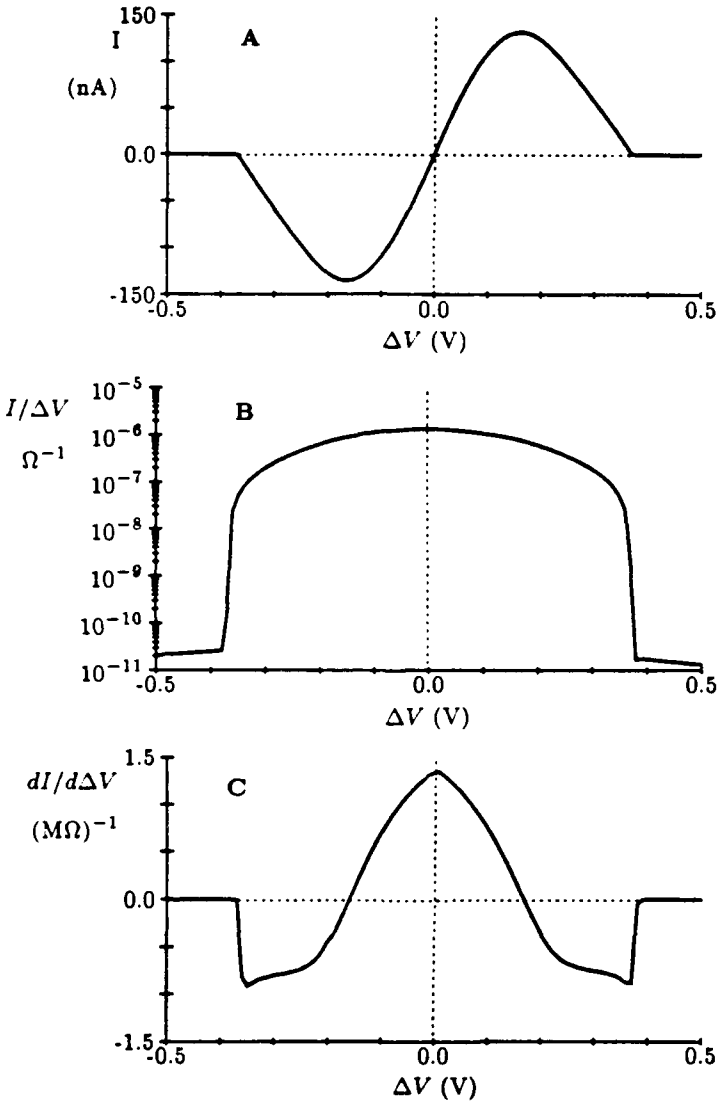

Figure 13 The I-V curve of the fuse measured in 10mV increments is shown in (A). (B) shows the numerically computed chord conductance, which is defined as $I/\Delta V$. Incremental conductance is defined to be $dI/d\Delta V$, which is the derivative of the I-V curve. (C) shows the incremental conductance computed using a two-point derivative approximation. Note the two regions of negative incremental conductance in (C).

will be *globally asymptotically stable*, i.e., for *any* voltage input and *any* initial condition it will converge to the unique stationary solution mentioned above. This conclusion holds even if positive, parasitic capacitances are distributed arbitrarily throughout the network, provided there are no inductors. This result assures us then that implementing the ideal, linear resistances dictated by standard regularization theory with Mead-type saturating resistances will not cause additional stationary solutions to appear.

2. Nonmonotonic Resistors

Now suppose the nonlinear resistors are externally *passive* (i.e., their I-V curves lie in the 1st and 3rd quadrants of the I-V plane) but are *incrementally active*, i.e., have regions of negative slope, as the resistive fuse in Fig. 3. Then there will in general exist a *number* of stationary network solutions for a given input image. If we further suppose that we can neglect the *internal* dynamics of the incrementally active resistor circuit, then for any voltage input and any initial condition the network will *not* oscillate indefinitely but must eventually settle to *some* stationary state. This conclusion also holds even if parasitic (positive) capacitances are distributed arbitrarily throughout the network, provided there are no inductors. This is a rather surprising result in view of the well-known instability problems with negative incremental resistance circuits.

3. Resistors with Internal Dynamics

The nonlinear resistors are of course multiple transistor circuits themselves and will inevitably have internal transient dynamics due to charge storage in transistors and parasitic wiring capacitance. Although each of the resistor circuits reported here is known to be stable in isolation, networks of such elements may, in principle, be unstable. This is an active research area, and many questions remain. Recent theoretical work (Wyatt and Standley, 1989; Standley and Wyatt, 1989; Standley, 1989) gives sufficient conditions for stability of such networks when the complex high-frequency dynamics are confined to the *linear* elements in any circuit consisting only of such linear elements, nonlinear memoryless resistors, and positive nonlinear capacitors. These results can be applied to yield *local* stability criteria for networks in which the resistor circuits are incrementally passive (such as Mead's saturating resistor) but have complex internal dynamics. But in their present form they are not applicable to networks in which the resistors are incrementally active (such as the resistive fuse) with internal dynamics.

The conclusions given in 1 and 2 above follow from well-established nonlinear network principles outlined below. Since the derivations follow with remarkable ease in these two cases, complete proofs are given.

We have sometimes found that experienced circuit designers can be deeply skeptical about the dynamic stability (non-oscillation) claim made above, and

tunnel diode oscillator circuits are sometimes mentioned as counterexamples. It may be helpful to clarify what the precise result, Theorem II below, actually assumes. In the first place, it assumes an inductorless circuit, i.e., the only circuit elements allowed are positive (but possibly nonlinear) capacitors, ideal constant voltage sources, and nonlinear (possibly incrementally active) resistors. Thus oscillators that rely on inductors, even the distributed inductance in connecting wires, are not ruled out by the theorem. Note also that nonreciprocal building blocks, such as amplifiers, are not allowed under the assumptions, and that the individual resistors are assumed to have *no* internal dynamics of their own. Finally, the theorem does *not* assert that *every* stationary network solution is stable. Some will be unstable and some will be stable, but the network will eventually always settle to one of the latter.

The "no-inductors" assumption and the "no resistor dynamics" assumption are modelling approximations. Their appropriateness in a particular context is always open to question, and the issue can be settled for any given circuit only by experimentation. We note here that neglecting on-chip inductance has proven to be an excellent approximation in the analysis of many practical circuits, and that the nonlinear resistor circuits reported here are *intended* by the designer to operate as essentially memoryless resistors.

All the conclusions in 1 and 2 above follow easily from Tellegen's theorem, restated below for convenience (Tellegen, 1952; Penfield, Spence and Duinker, 1970; Chua, Desoer and Kuh, 1987).

4. Tellegen's Theorem

Assume we are given a network with sign conventions for branch voltages V_k and branch currents I_k such that the product $V_k \cdot I_k$ represents the power flowing *into* branch k. Then

$$\sum_{\text{all network branches}} V_k \cdot I_k = 0 \tag{18}$$

Furthermore, suppose x_k represents either V_k or any quantity derived from V_k such that at each instant the set of all x_k satisfies Kirchhoff's Voltage Law (KVL), i.e., the x_k sum to zero around any loop in the network. And suppose y_k represents either I_k or any quantity derived from I_k such that at each instant the set of all y_k's satisfies Kirchhoff's Current Law (KCL) i.e., the sum of the y_k's entering any node is zero (examples include $x_k = dV_k/dt, x_k(t) = V_k(t+3), y_k = \int I_k$, etc.). Then

$$\sum_{\text{all network branches}} x_k(t_1)y_k(t_2) = 0, \text{ for all } t_1, t_2. \tag{19}$$

Tellegen's theorem makes it very easy to show why the stationary solution to any network with incrementally passive resistors must be unique, as claimed in section 1.

5. Theorem I (Uniqueness)

There exists at most one solution for the resistor voltages and currents in any network of arbitrary topology consisting of strictly incrementally passive resistors and ideal voltage and current sources.

Proof: Suppose on the contrary there exist two such solutions, solution a and solution b (if more exist, pick any two). Let V_k^a and V_k^b denote the voltage across branch k in the two solutions, ΔV_k denote $V_k^b - V_k^a$, and let ΔI_k be defined similarly. Then the set of ΔV_k's satisfies KVL and the ΔI_k's satisfy KCL, so from eq. (19)

$$\sum_{\text{all resistors and sources}} \Delta V_k \cdot \Delta I_k = 0. \tag{20}$$

Since $V_k^a = V_k^b$ for all voltage sources and $I_k^a = I_k^b$ for all current sources, the product $\Delta V_k \cdot \Delta I_k$ vanishes for all source branches and eq. (20) reduces to

$$\sum_{\text{all resistors}} \Delta V_k \cdot \Delta I_k = 0. \tag{21}$$

But each resistor curve has positive slope by assumption, so $\Delta V_k \cdot \Delta I_k \geq 0$. Thus eq. (21) guarantees that $\Delta V_k = 0$ or $\Delta I_k = 0$ for each resistor. Therefore ΔV_k *and* ΔI_k *both* vanish since each resistor curve is assumed to be single-valued and invertible.
Q.E.D.

This theorem first appeared in Duffin (1947); see also Birkhoff and Diaz (1956). A more recent treatment can be found in Hasler (1986).

The non-oscillation claims in sections 1 and 2 follow with similar ease from Tellegen's theorem. The key quantity of interest is the *resistor co-content* of eq. (2) (see also Poggio and Koch, 1985). Thus, the reason nonlinear RC networks cannot exhibit unforced sustained oscillations, even if the resistors are incrementally active, is because $J_{total}(t)$ is always "running down," i.e. J_{total} acts (roughly speaking) as a Lyapunov function.

6. Theorem II (Stability)

Consider a network of arbitrary topology consisting of nonlinear voltage-controlled resistors, ideal time-invariant voltage sources, and nonlinear but positive capacitors described by $I_k = C_k(V_k)\frac{dV_k}{dt}$, with $C_k(V_k) > 0$ everywhere. Then J_{total} is strictly decreasing at each instant during any transient, i.e.,

$$\frac{dJ_{total}(t)}{dt} \leq 0, \tag{22}$$

and the inequality is strict except at equilibrium.

Proof: From Tellegen's theorem, eq. (19), we have

$$\sum_{\text{all network branches}} I_k(t)\frac{dV_k(t)}{dt} = 0. \tag{23}$$

For the voltage sources $\frac{dV_k(t)}{dt} = 0$, so these drop out of the sum in eq. (23), which now reads

$$\sum_{\text{all resistors}} I_k(t)\frac{dV_k(t)}{dt} + \sum_{\text{all capacitors}} I_k(t)\frac{dV_k(t)}{dt} = 0. \tag{24}$$

For each resistor,

$$I_k(t)\frac{dV_k(t)}{dt} = \frac{dJ_k(t)}{dt}, \tag{25}$$

which follows from eq. (2), using the chain rule for derivatives. Thus the first sum in eq. (24) is just $dJ_{total}(t)/dt$. And for each capacitor,

$$I_k(t)\frac{dV_k(t)}{dt} = C_k(V_k(t))\left(\frac{dV_k(t)}{dt}\right)^2 \geq 0. \tag{26}$$

The inequality (22) follows upon substituting eqs. (25) and (26) into (24). **Q.E.D.**

This theorem is a special case of results in (Brayton and Moser, 1964), but the proof given here is much more elementary.

If J_{total} is bounded from below and slopes upward for large values of the voltages, then Theorem II implies that the network will settle into a steady-state. A sufficient condition for this is that the I-V curve of all resistors in the network should lie somewhere in the interior of the 1st and 3rd quadrants for large values of ΔV.

Note that Theorem II rules out sustained oscillation because $J_{total}(t)$ would have to be periodic if the network state were periodic, and this is impossible since $dJ_{total}/dt \leq 0$, with equality only at equilibrium. However, J_{total} does not necessarily meet all the standard criteria for a Lyapunov function since its *shape* is essentially arbitrary. It is easy to show that J is convex if and only if the resistors are all incrementally passive. With incrementally active resistors such as resistive fuses, J can have many *local minima*, which are then the (locally) *stable* equilibria of the network. In the case of positive linear resistors, Theorem 2 has the special interpretation that the total *dissipated power* decreases monotonically during transients in any RC circuit with voltage sources, even if the capacitors are nonlinear. In this linear case the co-content (and the total power) are convex functions of those voltages that are not constrained by the sources, so the local minimum to which the network converges is in fact the global minimum of the dissipated power, subject to the source constraint. Stripped of all dynamics, the static version of this statement is known as *Maxwell's Minimum Heat Theorem* (Maxwell, 1891).

CONCLUSION

We have successfully demonstrated in this manuscript for the first time a simple and elegant analog circuit implementation of the line discontinuities of Geman and Geman (1984) and of the graduated non-convexity algorithm of Blake and Zisserman (1987). We only report on the experimental data for an 8 pixel 1-D circuit. We have sent out a 20 by 20 pixel 2-D version of this network to MOSIS for fabrication. We previously demonstrated a 48 by 48 pixel circuit implementing smooth surface interpolation (Luo, Koch and Mead, 1988). This work can be extended to include 2nd order or thin-plate surface interpolation (Harris, 1989), where the energy functional embodies the discretized square of the ∇^2 operator. Computer simulations have shown that detection of discontinuities in surface orientation, such as occurring along creases, is feasible in problems such as edge detection and surface interpolation (Blake and Zisserman, 1987; Liu and Harris, 1989) and can be incorporated into our thin-plate interpolation circuits (Harris, 1989).

We thus have all the elementary circuit elements in hand—phototransistors for on-chip image acquisition (Mead, 1989), resistive networks for smoothing, and resistive fuses for detecting discontinuities—to design analog, resistive network chips to compute the 2-D optical flow field in the presence of motion discontinuities, the depth and depth discontinuities in 2-D images as well as intensity discontinuities.

Acknowledgements

We thank Carver Mead for laying the framework upon which we have built our research. Our theoretical ideas would never have been ported into silicon without him. Hewlett-Packard provided computing support. All chips were fabricated through MOSIS with DARPA's support. Research in the laboratory of C.K. is supported by a National Science Foundation grant IST-8700064, Office of Naval Research Young Investigator and National Science Foundation Presidential Young Investigator Awards, the James S. McDonnell Foundation, Rockwell International Science Center and the Hughes Aircraft Artificial Intelligence Center. J.W. acknowledges DARPA Contract No. N00014-87-K-0825, National Science Foundation Contract No. MIP- 8814612, and the DuPont Corporation. We thank David Standley for his careful reading of this document and Bimal Mathur for inspiring the high-gain fuse circuit.

References

Birkhoff, G. and Diaz, J. B. (1956). Nonlinear network problems. *Quart. Appl. Math.* **13**:431–443.

Blake, A. (1989). Comparison of the efficiency of deterministic and stochastic algorithms for visual reconstruction. *IEEE Trans. Pattern Anal. Mach. Intell.* **11**:2–12.

Blake, A. and Zisserman, A. (1987). *Visual Reconstruction.* Cambridge, MA: MIT Press.

Brayton, R. K. and Moser, J. K. (1964). A theory of nonlinear networks—I, II. *Quart. Appl. Math.* **22(1)**:1–33 (April) and **22(2)**:81–104 (July).

Chua, L. O., Desoer, C. A., and Kuh, E. S. (1987). *Linear and Nonlinear Circuits.* New York: McGraw-Hill, pp. 23–34.

Duffin, R. J. (1947). Nonlinear networks IIa. *Bull. Amer. Math. Soc.* **53**:963–971.

Geman, S. and Geman, D. (1984). Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.* **6**:721–741.

Grimson, W. E. L. (1981). *From Images to Surfaces.* Cambridge, MA: MIT Press.

Harris, J. G. (1989). An analog VLSI chip for thin plate surface interpolation. In *Neural Information Processing Systems*, ed. D. Touretzky. Palo Alto: Morgan Kaufmann.

Harris, J. G. and Koch, C. (1989). Resistive fuses: circuit implementations of line discontinuities in vision. *Snowbird Neural Network Workshop*, April 4–7.

Harris, J. G., Koch, C., Staats, E., Luo, J. and Wyatt, J. (1989). Analog hardware for detecting discontinuities in early vision: computational justification and VLSI circuits, in preparation.

Hasler, M. and Neirynck, J., (1986). *Nonlinear Circuits*. Norwood, MA: Artech House Inc., pp. 172–173.

Hildreth, E. C. (1984). *The Measurement of Visual Motion*. Cambridge, MA: MIT Press.

Hopfield, J. J. and Tank, D. W. (1985). Neural computation in optimization problems. *Biol. Cybern.* **52**:141–152.

Horn, B. K. P. (1989). Parallel networks for machine vision. *Artif. Intell. Lab. Memo No.* **1071** (MIT, Cambridge).

Horn, B. K. P. and Schunck, B. G. (1981). Determining optical flow. *Artif. Intell.* **17**:185–203.

Ikeuchi, K. and Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artif. Intell.* **17**:141–184.

Hutchinson, J., Koch, C., Luo, J., and Mead, C. (1988). Computing motion using analog and binary resistive networks. *IEEE Computer* **21**:52–63.

Ikeuchi, K. and Horn, B. K. P. (1981). Numerical shape from shading and occluding boundaries. *Artif. Intell.* **17**:141–184.

Koch, C., Marroquin, J., and Yuille, A. (1986). Analog "neuronal" networks in early vision. *Proc. Natl. Acad. Sci. USA* **83**:4263–4267.

Koch, C. (1989). Seeing chips: analog VLSI circuits for computer vision. *Neural Computation* **1**:184–200.

Liu, S. C. and Harris, J. G. (1989). Generalized smoothing networks in solving early vision problems. *Computer Vision and Pattern Recognition Conference.*

Luo, J., Koch, C., and Mead, C. (1988). An experimental subthreshold, analog CMOS two-dimensional surface interpolation circuit. *Neural Information Processing Systems Conference*, Denver, November.

Marr, D. and Poggio, T. (1976). Cooperative computation of stereo disparity. *Science* **194**:283–287.

Marroquin, J., Mitter, S., and Poggio, T. (1987). Probabilistic solution of ill-posed problems in computational vision. *J. Am. Statistic Assoc.* **82**:76–89.

Maxwell, J. C. (1891). *A Treatise on Electricity and Magnetism*, 3rd ed., Vol. I, pp. 407–408. Republished by New York: Dover Publications, 1954.

Mead, C. A. (1985). A sensitive electronic photoreceptor. In *1985 Chapel Hill Conference on Very Large Scale Integration*, pp. 463-471.

Mead, C. A. (1989). *Analog VLSI and Neural Systems*. Reading: Addison-Wesley.

Millar, W. (1951). Some general theorems for non-linear systems possessing resistance. *Phil. Mag.* **42**:1150–1160.

Nagel, H. H. (1987). On the estimation of optical flow: relations between different approaches and some new results. *Artif. Intell.* **33**:299–324.

Penfield, P., Jr., Spence, R., and Duinker, S. (1970). *Tellegen's Theorem and Electrical Networks*, Cambridge, MA: MIT Press.

Perona, P. and Malik, J. (1988). A network for multiscale image segmentation. *Proc. 1988 IEEE Int. Symp. on Circuits and Systems*, Espoo, Finland, June, pp. 2565–2568.

Poggio, T., Gamble, E. B., and Little, J. J. (1988). Parallel integration of vision modules. *Science* **242**:436–440.

Poggio, T. and Koch, C. (1985). Ill-posed problems in early vision: from computational theory to analogue networks. *Proc. R. Soc. Lond. B* **226**:303–323.

Poggio, T., Torre, V., and Koch, C. (1985). Computational vision and regularization theory. *Nature* **317**:314–319.

Poggio, T., Voorhees, H., and Yuille, A. (1986). A regularized solution to edge detection. *Artif. Intell. Lab Memo* **No. 833** (MIT, Cambridge).

Sivilotti, M.A., Mahowald, M.A. and Mead, C.A., Real-time visual computation using analog CMOS processing arrays. In: *1987 Stanford Conf. VLSI*, pp. 295-312 (MIT Press, Cambridge, 1987).

Standley, D. L., and Wyatt, J. L., Jr. (1989). Stability criterion for lateral inhibition and related networks that is robust in the presence of integrated circuit parasitics. In *IEEE Trans. Circuits and Systems* **36**, May., pp. 675–681

Standley, D. L. (1989). Design criteria extensions for stable lateral inhibition networks in the presence of circuit parasitics. *Proc. 1989 IEEE Int. Symp. on Circuits and Systems*, Portland, Oregon, May, pp. 837–840.

Tellegen, B. D. H. (1952). A general network theorem, with applications. *Phillips Research Reports* **7**:259–269.

Terzopoulos, D. (1983). Multilevel computational processes for visual surface reconstruction. *Comp. Vision Graph. Image Proc.* **24**: 52–96.

Terzopoulos, D. (1986). Regularization of inverse problems involving discontinuities. *IEEE Trans. Pattern Anal. Machine Intell.* **8**:413–424 (1986).

Wyatt, J. L., Jr. and Standley, D. L. (1989). Criteria for robust stability in a class of lateral inhibition networks coupled through resistive grids. *Neural Computation* **1**:58–67.