Analog CMOS IC for Object Position and Orientation

David Standley

Rockwell International Science Center, Imaging Devices Dept.
Thousand Oaks, CA 91360

Berthold Horn

Massachusetts Institute of Technology
Dept. of Electrical Engineering and Computer Science
Cambridge, MA 02139

ABSTRACT

A CMOS VLSI chip that determines the position and orientation of an object is described. The chip operates in a continuous-time analog fashion, with a response time as short as 200 μs and power consumption under 50 mW. A self-contained phototransistor array acquires the image directly, and the output is a set of eight currents from which the position and orientation can be found. Orientation is determined to within $\pm 2\%$ or better for moderately sized and sufficiently elongated objects. Chip dimensions are 7900 μm by 9200 μm.

1. INTRODUCTION

An analog VLSI chip that determines the position and orientation of an object against a dark background is described. The implemented algorithm is based on finding the first and second moments of the object's spatial intensity distribution[1]. These moments allow the centroid (an indicator of position) and the axis of least inertia (an indicator of orientation) to be computed (Fig. 1). The chip has a self-contained array of phototransistors, which are available in CMOS,[2] so that the input is acquired by focusing the scene directly onto the chip surface. Using a novel scheme,[3] the moments are computed using a uniform grid of linear resistors. Only eight measurements are required. Resistive sheets have been used in earlier systems to determine the position of objects,[3] e.g. a small, bright spot. Also, the analog VLSI chip by DeWeerth[4] et al. finds the centroid of an object, though using a different method. Yet neither performs the orientation task. A dim scene background can be completely removed by an adjustable brightness threshold, so that it does not interfere with the calculation. Earlier reports on this work have appeared elsewhere.[5,6,7]

 0-8194-0582-5/91/$4.00

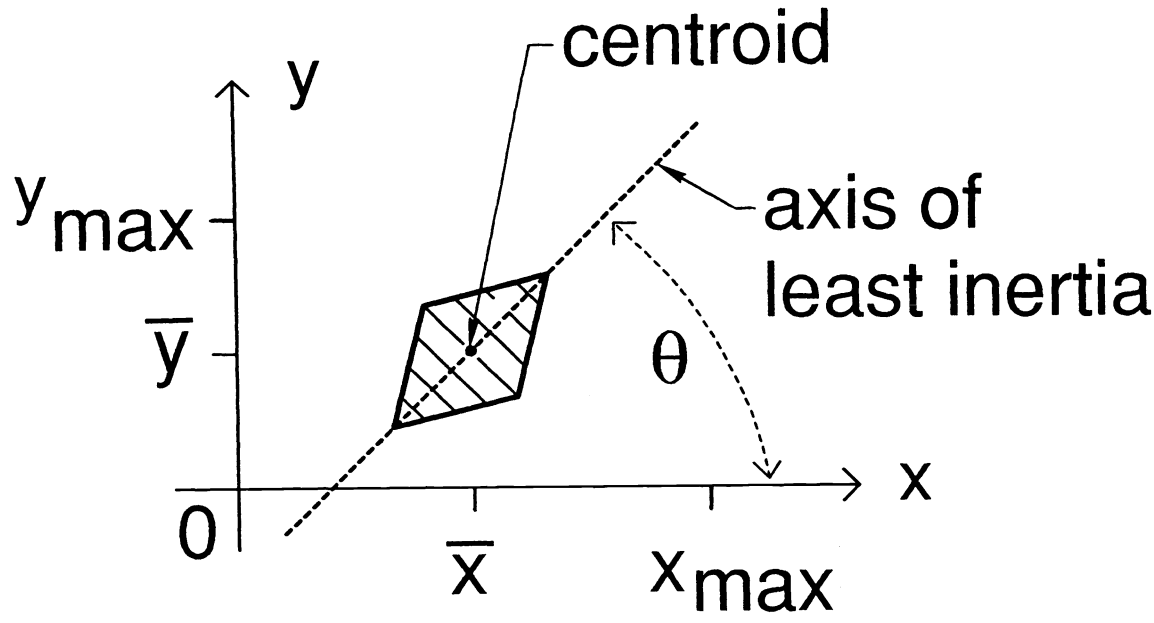

Figure 1: Example of object centroid and axis of least inertia.

2. ALGORITHM

Consider an image (in continuous space) having brightness $B(x,y)$ in a field where $0 \leq x \leq x_{\max}$ and $0 \leq y \leq y_{\max}$. At each location, the brightness is assigned a weighting

$$m(x,y) \triangleq f(B(x,y)), \tag{1}$$

where

$$f(B) \triangleq \begin{cases} 0, & 0 \leq B \leq B_{\text{thr}} \\ B - B_{\text{thr}}, & B > B_{\text{thr}} \end{cases} \tag{2}$$

and B_{thr} is a threshold value. This piecewise-continuous *brightness conditioning function* $f(B)$ removes a dim scene background while essentially retaining the gray-level character of the image (Fig. 2).

The first moments of the weighting $m(x,y)$ are

$$M_x \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} x\, m(x,y)\, dx\, dy \tag{3}$$

$$M_y \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} y\, m(x,y)\, dx\, dy. \tag{4}$$

Similarly, the second moments are

$$M_{xy} \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} xy\, m(x,y)\, dx\, dy \tag{5}$$

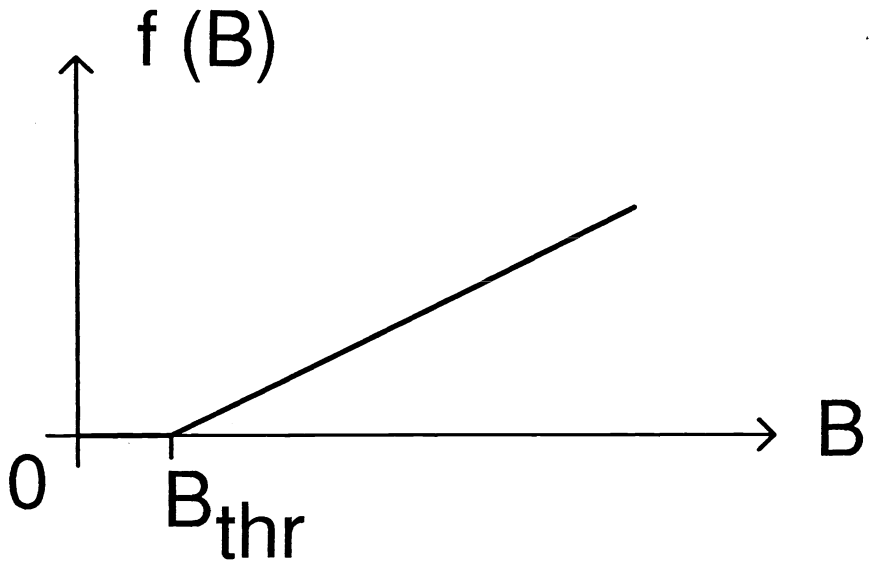

Figure 2: This brightness conditioning function shows the image data weighting vs. the brightness. Thresholding is included.

$$M_{x^2} \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} x^2\, m(x,y)\, dx\, dy \tag{6}$$

$$M_{y^2} \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} y^2\, m(x,y)\, dx\, dy. \tag{7}$$

The zeroth moment M_0 is simply

$$M_0 \triangleq \int_0^{x_{\max}} \int_0^{y_{\max}} m(x,y)\, dx\, dy. \tag{8}$$

The centroid $(\bar{x}, \bar{y})$, which indicates the position, is given by the normalized first moments

$$\bar{x} = M_x / M_0 \tag{9}$$

and

$$\bar{y} = M_y / M_0. \tag{10}$$

The angle of the axis of least inertia, which indicates the orientation, can be found by first calculating the quantities

$$a' \triangleq M_{x^2} - M_0(\bar{x})^2, \tag{11}$$

$$b' \triangleq M_{xy} - M_0(\bar{x}\bar{y}), \tag{12}$$

and

$$c' \triangleq M_{y^2} - M_0(\bar{y})^2. \tag{13}$$

Then the angle of the axis of least inertia, as shown in Fig. 1, is given by

$$\theta = \frac{1}{2}\arctan(2b', a' - c') \tag{14}$$

where $\arctan(v,u)$ for any $(v,u) \neq (0,0)$ is the unique angle $\phi \in [0^o, 360^o)$ such that

$$u = \sqrt{u^2 + v^2}\cos\phi$$

and

$$v = \sqrt{u^2 + v^2}\sin\phi.$$

Note that certain objects, such as a square, will not have a unique axis of least inertia, and hence orientation is not defined by this method. Yet many obejcts, e.g., uniform elongated ones, will have an orientation angle. Also note that the three second moments are not needed separately; only M_{xy} and the difference $M_{x^2} - M_{y^2}$ are required.

3. RESISTIVE GRID SCHEME

The resistive grid scheme proposed by Horn[3] is essentially processing by dimensional reduction of data. In the continuous image field case, a uniform resistive sheet can be used to reduce the double integrals in the above moment formulas to a single integral; this can be thought of as a 2-D to 1-D reduction. For simplicity we assume a square image field; i.e., $x_{\max} = y_{\max}$. In the discrete image case (i.e. with pixels), which is implemented, each corresponding weighted sum (which replaces an integral) over an $N \times N$ image array can be reduced to an equvalent sum over $4N$ quantities with a resistor grid. These, in turn, can be reduced to eight quantites with resistor lines, independent of the array size N. Figure 3 shows a uniform resistor grid, which is grounded around the perimeter. Each internal node corresponds to a pixel (j,k) in the 2-D array, and a current $i(j,k)$ (proportional to the weighting $m(j,k)$ at the pixel) is injected into the grid. The $4N$ currents flowing out of the perimeter contain enough information to extract the moments. In particular,

$$\sum_{j=1}^{N}\sum_{k=1}^{N} h(j,k)i(j,k) = \sum_{(j,k)\in P} h(j,k)i_{\mathrm{p}}(j,k), \tag{15}$$

where P is the set of all nodes on the perimeter, $i_{\mathrm{p}}(j,k)$ is the current flowing out of the grid at each perimeter node, and $h(j,k)$ is any *harmonic* function (i.e. the Laplacian of $h(j,k)$ vanishes identically)[5]. In the continuous case the required spatial weightings are $h(x,y) = 1$, x, y, xy, and $x^2 - y^2$; to within a scale factor, these correspond to discrete weightings of $h(j,k) = 1$, j, k, jk, and $j^2 - k^2$, respectively, which are all harmonic. Thus, (15) shows that the 2-D weighted sum over the image array can be converted to an equivalent 1-D weighted sum around the perimeter of the grid.

Resistor lines are used to further reduce the perimeter currents to a set of eight currents, which are available at the corners; they operate as current dividers. The following section describes the setup more specifically; algebra is omitted for brevity. Complete details can be found in the thesis[5].

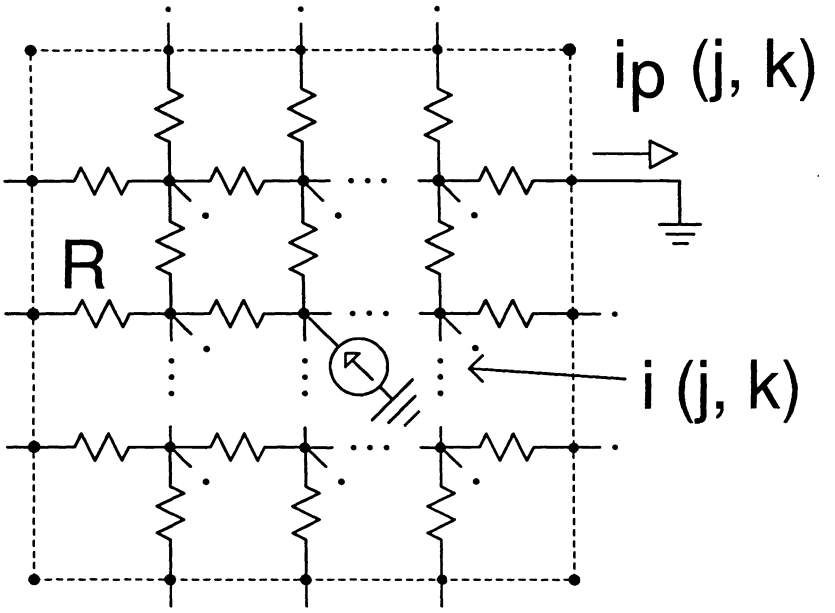

Figure 3: Uniform resistor grid is grounded around the perimeter. A 2-D array of current (corresponding to the image data) is injected into the grid, and the currents leaving around the perimeter contain the information needed to compute the required first and second moments.

4. ARCHITECTURE AND CIRCUITRY

Figures 4 and 5 show the chip architecture. The resistor grid and *photoreceptor cell array* (Fig. 4) occupy most of the active chip area. The grid is a 30 by 30 array of 3 kΩ polysilicon resistors, which is driven by a 29 by 29 array of photoreceptor cells. Each photoreceptor cell contains a phototransistor together with other circuitry to implement the brightness conditioning function of Fig. 2; i.e., the incident brightness is converted to a current proportional to $m(j,k)$, which is injected into the grid. (Note that effective subpixel resolution is available if the image is blurred slightly to avoid aliasing resulting from the finite sampling grid.) Around the perimeter of the grid are the *current buffers*, which hold each perimeter node at a common d.c. voltage (effectively ground) and convey the currents flowing out of the grid into uniform and quadratic resistor lines at the periphery, as shown in Fig. 5. The buffer outputs can be simultaneously steered to either the uniform or quadratic lines, and the corresponding sets of output currents, $i_1 - i_4$ and $i_5 - i_8$, respectively, are measured. Each uniform line simply gives a linear weighting as a function of position (or array index) on a side of the grid. Each quadratic line weights the currents according to the square of the normalized distance along the line, where the origins are at the left and the bottom for the horizontally and vertically oriented lines, respectively. The resistors in the lines are polysilicon, and each resistor in each quadratic line is made by connecting certain other resistors (selected from a set of "primitives") in series. The ends of the lines are held at a virtual ground by *external* op-amp circuits (not shown) that convert the currents flowing out of the chip into voltages. The eight output quantities are enough to find the object position and orientation with simple formulas.[5]

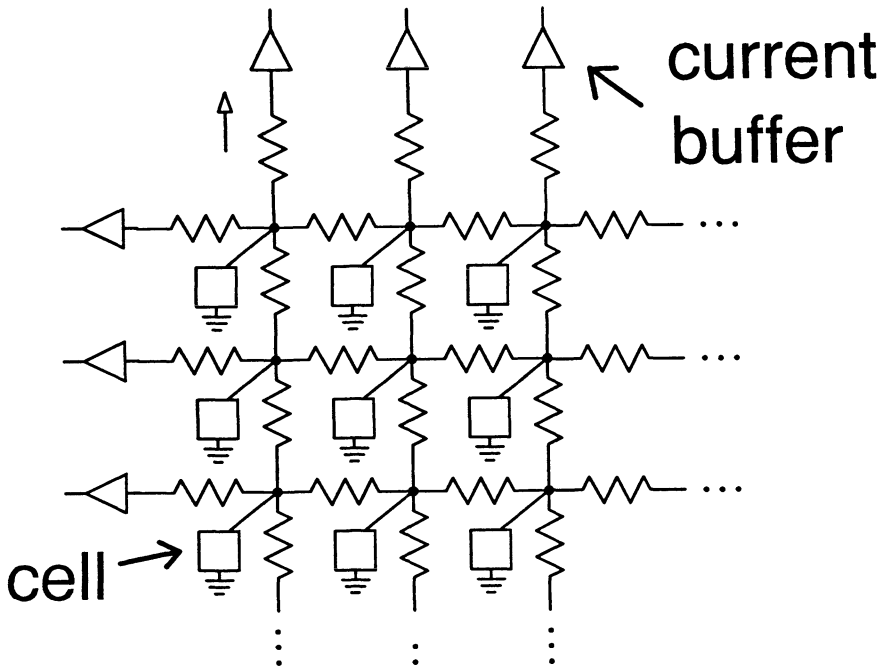

Figure 4: Resistor grid and photoreceptor cell array.

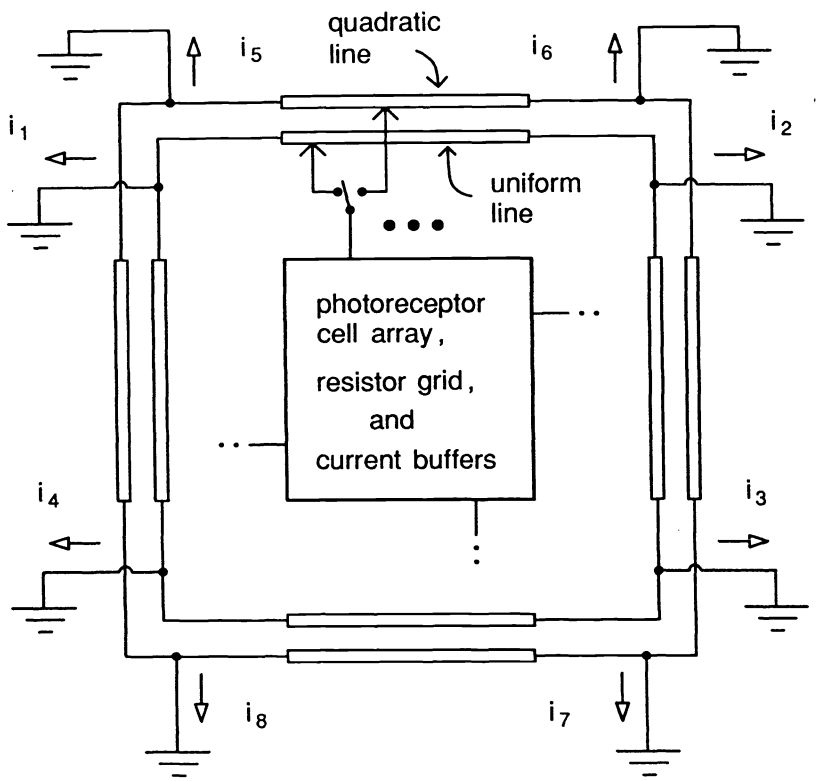

Figure 5: Main chip architecture.

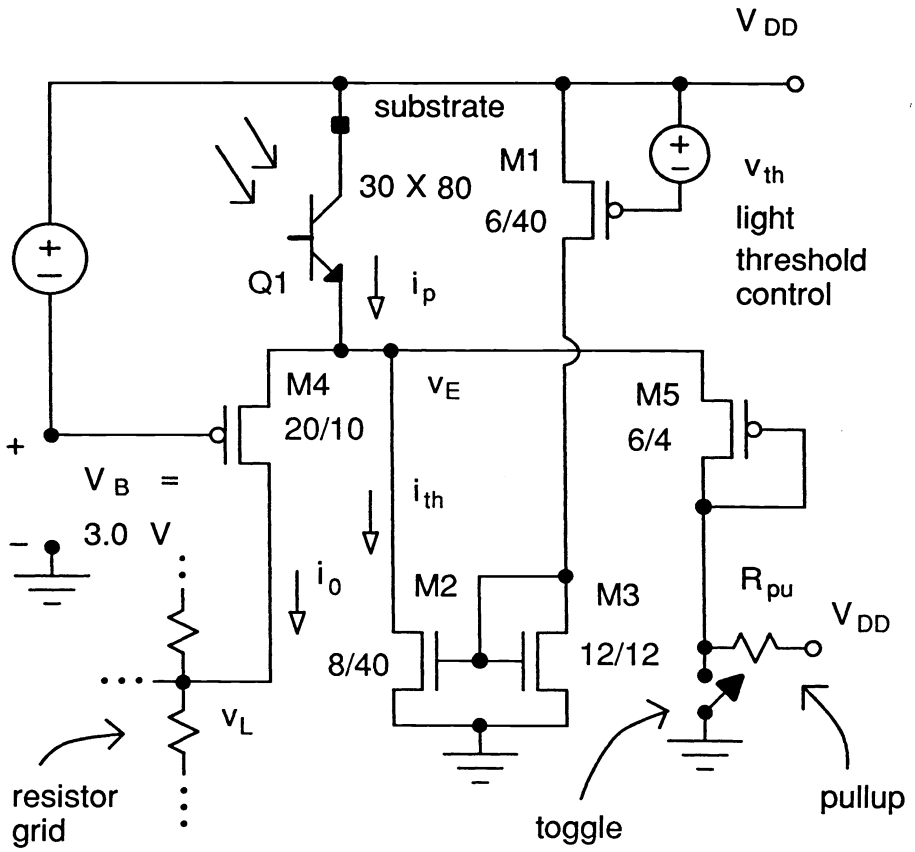

Figure 6: Circuit for photoreceptor cell. Voltage sources and the toggle are on global busses.

Figure 6 shows the schematic of a photoreceptor cell, together with bias sources on global busses, which are realized by diode-connected PFETs driven by external current sources. All sizes are in microns. Transistors M1 – M3 form the thresholding current source i_{th}, which subtracts from the photocurrent i_p of Q1. The gate of output transistor M4 is held at about 3.0 V above ground. The drain of M4 is the cell output, which is connected to the resistor grid. In normal operation, the grid voltage is at most 3.0 V; and thus M4, which is either cut off or in saturation, acts like a diode in addition to a cascode transistor (which gives a higher output resistance than if M4 were simply diode-connected). If the incident light level is below the threshold, i.e., if $i_p < i_{th}$, the output current $i_o = 0$. If the light level is above the threshold, i.e., $i_p > i_{th}$, then $i_o = i_p - i_{th}$. The result is a continuous, piecewise-linear d.c. response curve, ideally like $f(B)$ (Fig. 2). Diode-connected PFET M5 is connected to a bus normally held high, so it is cut off. Grounding this toggle bus forces i_o to zero for all the cells. By measuring the final outputs with all cells "shut off", the net effects of offset errors in the system are measured and thus cancellable; this was done in all experiments. Typical operating currents are $i_p = .1$ to 1.0 μA and $i_{th} = .2$ μA. (See the thesis[5] for a discussion of the current buffers.)

5. EXPERIMENTAL RESULTS

Working chips have been fabricated in a 2 μm p-well process. Total chip dimensions are 7200 μm by 9200 μm, and the light-sensitive cell array occupies a 5500 μm by 5500 μm area. Power is nominally 30 mW, which is shared between the 841 photoreceptor cells and the 116 current buffers. Performance is dependent[5] on object size and shape, but changes in orientation can typically be measured to within $\pm 2^o$ for elongated and moderately sized objects, e.g., a rectangle of dimensions 30 by 60 on a 100 by 100 image field. The position of a 25 by 25 square is determined to within $\pm$ .3 % of the usable range (i.e., the range for which the object remains completely in the image field) when referenced to a linear least squares fit line. Measurements of the response to light pulses show that the chip operates at about 5000 images per second. (Though the system operates in continuous time, speed can be characterized for images that remain constant over a certain period; this is analogous to the settling performance of operational amplifiers.) The speed is limited by the photoreceptor cells.

6. ACKNOWLEDGMENTS

The authors thank Professors Hae-Seung Lee, Charles Sodini, and John Wyatt at the Massachusetts Institute of Technology for helpful comments. This work was performed at the Research Laboratory of Electronics at M.I.T., and further work is continuing at Rockwell. This work was supported by the National Science Foundation and the Defense Advanced Research Projects Agency under Contract MIP-8814612, and DuPont Corp..

7. REFERENCES

1. Horn, B.K.P., **Robot Vision**, M.I.T. Press, Cambridge, MA, and McGraw-Hill, New York, NY, 1986, pp. 48–57.

2. Mead, C., "A Sensitive Electronic Photoreceptor," *Proc. 1985 Chapel Hill Conf. on VLSI*, 1985, pp. 463–471.

3. Horn, B.K.P., *A.I. Memo No. 1071*, M.I.T. Artificial Intelligence Lab., M.I.T., Cambridge, MA, Dec. 1988, pp. 31–34.

4. DeWeerth, S.P. and C.A. Mead, "A Two-Dimensional Visual Tracking Array," *Proc. 1988 MIT Conf. on VLSI*, M.I.T. Press, Cambridge, MA, pp. 259–275.

5. Standley, D.L., *Analog VLSI Implementation of Smart Vision Sensors: Stability Theory and an Experimental Design*, Ph.D. thesis, Dept. of Elec. Eng. and Comp. Sci., M.I.T., Cambridge, MA, Jan. 1991.

6. Standley, D.L. and B.K.P. Horn, "An Object Position and Orientation IC with Embedded Imager," *Proc. IEEE International Solid-State Circuits Conference*, San Francisco, CA, Feb. 13-15, 1991, pp. 38-39.

7. Wyatt, J.L. Jr., D.L. Standley, and W. Yang, "The MIT Vision Chip Project: Analog VLSI Systems for Fast Image Acquisition and Early Vision Processing," to appear in *Proc. 1991 IEEE International Conference on Robotics and Automation*, April 1991.