CALIFORNIA INSTITUTE OF TECHNOLOGY
COMPUTATION AND NEURAL SYSTEMS PROGRAM

JULY 5, 1991 CNS MEMO 11

SCANNERS FOR VISUALIZING ACTIVITY OF ANALOG VLSI CIRCUITRY

Carver A. Mead
Tobias Delbrück

PASADENA, CALIFORNIA 91125

Scanners for Visualizing Activity of Analog VLSI Circuitry

Carver A. Mead and Tobias Delbrück
Computation and Neural Systems Program
139-74
California Institute of Technology
Pasadena, CA, 91125
Internet E-mail: tdelbruck@caltech.edu

July 5, 1991

Abstract

This paper tutorially describes mixed digital–analog serial multiplexers (scanners) that we use to visualize the activity of one- and two-dimensional arrays of analog VLSI elements. These scanners range from simple one-dimensional devices designed to scan a one-dimensional array onto an oscilloscope, to complete video scanners with integrated sync and blank computation and on-chip video amplifiers. We discuss practical details of design and performance, and we give a source for example scanner layout.

1 Using Scanners in Analog VLSI Design

We use scanners as a diagnostic tool, to observe the behavior of large analog VLSI chips, when the chips have far more nodes than the pins available to us, or when the chip has a topography that maps well onto a display device.

This paper is a tutorial description of the distilled knowledge gained from over 100 chip designs using scanners. We describe how our scanners work and how we use them to multiplex outputs from one- and two-dimensional arrays onto output devices such as oscilloscopes and monitors. In Section 2, we discuss one-dimensional scanners and their component parts as prototypical of two-dimensional scanners. In Section 3, we discuss two-dimensional scanners and generation of video control signals. In Section 4, we discuss the analog parts of the scanners: output representations, current-sense amplifiers, and video drivers. In Section 5, we discuss performance of the digital and analog parts of the scanners. Finally, in Section 6, we discuss practical design details. In the appendix, we give an analysis of a logarithmic current-sense amplifier.

2 A One-Dimensional Scanner

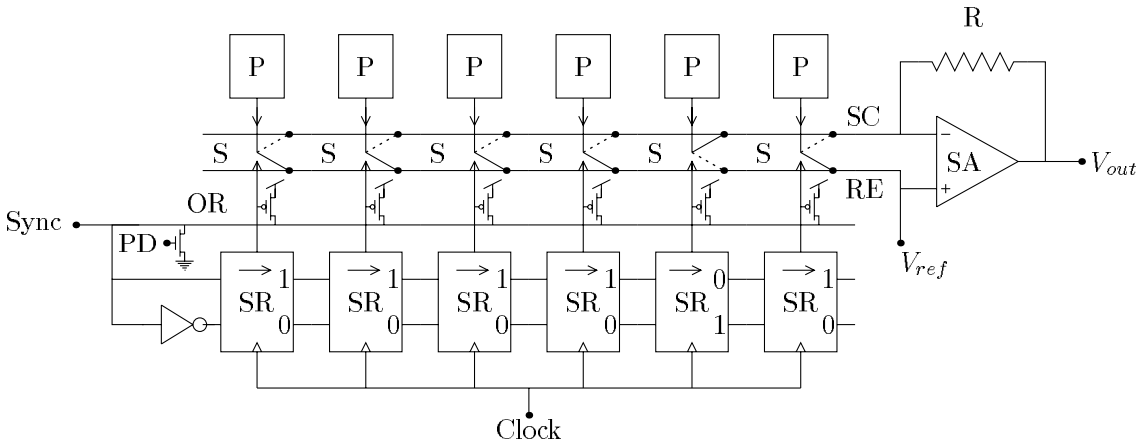

Figure 1: Schematic of a one-dimensional scanner. Pixels (P) send their output current through switches (S) to either a scan (SC) or a reference (RE) wire. The pixel output current is routed through the SC wire and is sensed by the current-sense amplifier (SA). The current-sense amplifier holds the SC wire at approximately $V_{\rm ref}$ and outputs a voltage proportional to the pixel output current. The switches are controlled by the shift-registers (SR). The shift-registers shift in the direction of the arrow. A new bit is loaded into the start of the shift-register at the end of a scan by means of the wired NAND, which computes the OR of low bits in the shift-register (OR). The OR output also acts as a sync input to the oscilloscope. The pulldown transistor on the OR line is biased with the bias PD. The output from the sense-amplifier goes to an oscilloscope.

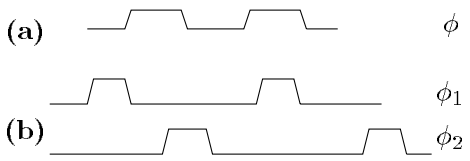

Figure 2: Two clocking schemes. (a) Single-phase. (b) Two-phase, nonoverlapping. Clock signals go from ground to V_{dd}.

Figure 1 shows a one-dimensional scanner. This scanner multiplexes the output from a one-dimensional array of pixels serially onto a single output line. The pixels generate analog currents that are to be scanned out and displayed. The pixel outputs are multiplexed by MOS switches onto the output line or onto the reference line. At a particular point in the scan, one pixel sends its output into the output line, and all the other pixels feed into the reference line. The reason for this dual-output scheme is that, generally, when we want to scan quickly, we must hold the wires at a fixed potential (virtual ground) while we sense an output current. The dual-output scheme allows us to hold all the wires at the same virtual ground. If we scanned out a voltage, each pixel would need to charge the output wire to the output voltage, and this circuit would run much more slowly than does the current-sense arrangement. The current-sense amplifier in Figure 1 shows one method for sensing the current coming out of the scanned pixel.

The blocks at the bottom of Figure 1 represent a digital shift-register that controls the scanning process. With each clock cycle, whatever bit is in a given stage is shifted to the next stage. Whatever bit is in the last stage in the line gets shifted out and disappears forever. Typically, there will be a single register that holds a low bit, and all the other registers will hold a high bit. That one register with the low bit determines the one pixel that will be connected to the output amplifier in that clock cycle. In Figure 1, the second pixel from the right is being scanned out. All other pixels are connected to a reference line. The line coming into the bottom of each shift-register stage is the global single-phase clock input. The wired NAND circuit logically ORs together all the shift-register low values so that a new low bit is generated when only high bits are left in the register. This arrangement is self-initializing and requires no off-chip control.

In the following sections, we discuss the component parts of this one-dimensional scanner.

2.1 Shift-Register Operation

In these scanners, the shift-register is a static digital device. The clock input required is a simple single-phase signal alternating between high and low. Other schemes use a two-phase, nonoverlapping clocking scheme, where there are two clock signals, both alternating between high and low, such that the high periods of the two signals are completely nonoverlapping. The two schemes are shown in Figure 2. Since the shift-registers are static devices, they hold their values without refreshing. Hence, we may view a particular pixel continuously by stopping the scan at that pixel. The shift-registers outputs are fully restored to V_{dd} or ground when the clock is high.

The single-phase CSRL shift-register circuitry is a novel design, and was derived by Massimo Sivilotti[5] from an earlier two-phase version [4]. The circuitry for one stage of

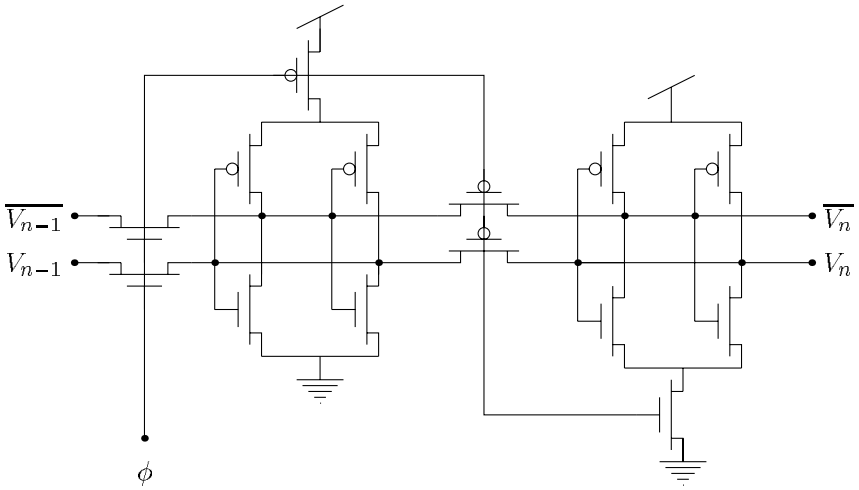

Figure 3: A single-phase static shift-register stage. The bit in the stage is represented by the complementary pair V_n and $\overline{V_n}$; these signals are fully restored to V_{dd} or ground when ϕ is high.

a shift-register is shown in Figure 3. We see that a stage consists of two pairs of cross-coupled inverters, along with n-type pass transistors feeding into the first pair, and p-type pass transistors connecting the first and second pair. [1]

The inverters are conventional CMOS inverters except that they have a transistor of one type or the other in series with their power supply. For the left pair of inverters, the p-type power-supply transistor is between V_{dd} and the inverters; thus the inverters will be powered only when the clock input ϕ is low. For the right set of inverters, the situation is reversed: The inverters will be powered when ϕ is high. The bit in a stage is represented by the values V_n and $\overline{V_n}$; the two signals are fully restored and complementary when the clock is high.

When the cross-coupled inverters are powered, they have two stable states corresponding to V_n high or low and $\overline{V_n}$ complementary to V_n. When they are not powered, we can load bits onto their inputs. That loading is the function of the pass transistors connecting this stage to the previous stage and connecting the first pair of cross-coupled inverters to the second.

It is easy to see how this arrangement works, as long as you already believe that it works. The key requirement to make these single-phase shift-registers function correctly is that the pass transistors must be sufficiently weaker than the power transistors that the driven stage cannot change the state of the driving stage. The correct sizing restriction will ensure unidirectionality of information transfer. Analytically, and also empirically, a sizing restriction that works over the current range of MOSIS parameters is that the pass transistors feeding into the cross-coupled inverters must have a saturation current at most one-fourth that of the power transistors or those in the cross-coupled inverters [5]. If this

[1] In this paper, we use the grounded-substrate convention: Transistors without bubbles on the gate are native devices; transistors with bubbles on the gate are in the well. The potential of the substrate is always ground; the well is at $+5V$ for an n-well technology and $-5V$ for a p-well technology. It is simplest to think of all circuit operation as occurring in an n-well implementation. In that case, a logic-high signal is $+5V$ and a logic-low signal is ground.

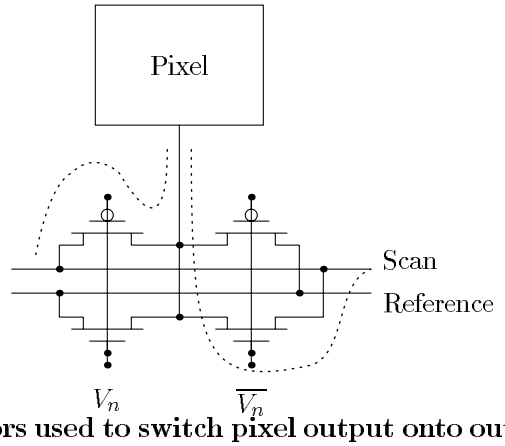

Figure 4: Pass-transistors used to switch pixel output onto output line or onto reference line. The dotted lines show the flow of current when V_n is low.

sizing restriction is obeyed, then the shift-register will function correctly with any clock rise- or fall-time. In fact, a sinusoidal oscillator may be used as the clock input.

This fully-static, single-phase, shift-register stage is a novel design. Other static digital shift registers use two-phase, nonoverlapping clocks that are inconvenient and unnecessarily complicated to generate, and are particularly prone to clock-skew problems. Other single-phase designs are either not fully static, or have a high transistor count [2].

2.2 Generating a New Bit

The scanners discussed in this paper have at most one low bit in the shift-register. Whenever that low bit falls out the end of the shift-register, a new low bit is generated and is loaded in at the beginning. The rest of the time, high bits are loaded into the shift-register. The register is thus self-initializing. This trick is managed by a wire that computes the logical NAND of all the outputs of the shift-register stages. Conceptually, we can think of this circuit as an OR gate for low-going inputs. If there is a low bit in any stage, then the OR line is pulled high, and only highs are loaded into the shift-register. If there is no low bit, then the OR line gets pulled low by a pulldown transistor, generating a new low bit. In addition to generating a new bit, the NAND output acts as a synchronizing signal (sync) to trigger the oscilloscope sweep.

2.3 Multiplexing the Pixel Output

We use the low bit marching along in the shift-register to switch the output from the selected pixel (or column, as we shall see in Section 3) onto the output line. The outputs from all the other pixels are connected to the reference line. The switch, shown in Figure 4, is a conventional analog multiplexer, consisting of four transistors, two of each type. We use this complementary arrangement since a particular type type of transistor will not pass signals well when the signals are far from that type's bulk potential, that is, native transistors will not pass signals near V_{dd}, and transistors in the well will not pass signals near ground.

When the switching transistors are turned off, whatever charge is under the gates of the transistors gets pushed out from the channel and adds to the current we are observing. This

charge injection sometimes causes unwanted transients in the pixel output, particularly at low current levels. The use of complementary pass transistors somewhat ameliorates this charge-injection problem, since one type injects electrons and the other type injects holes [6]. More sophisticated techniques, not described here, may further alleviate this problem[8].

2.4 Sensing the Output

In general, it is difficult to scan rapidly if the pixel generates a voltage that must charge the output line. For a one-dimensional scanner that will be viewed on an oscilloscope, a voltage scan is feasible as long as the scan rate is no higher than a few kHz[7]. A two-dimensional video scanner, however, would require pixels that could drive a long metal wire at several MHz. For this reason, we have chosen schemes that sense a current generated by the pixels. By using a single, fast, feedback amplifier, we can hold the output line at a fixed voltage, *virtual ground*, thus allowing much higher scan rates.

Figure 1 shows the prototypical current-sense amplifier, arranged in a negative feedback configuration. Any deviation from V_{ref} at the scanout line is multiplied by the gain of the amplifier and fed back to the scanout line through a feedback element. Thus the feedback amplifier senses the error signal at the scanout line and uses it to correct the voltage at the scanout line. Under this condition, the current through the feedback element is equal to the pixel output current. If the current were not equal, then the voltage on the scan line would continue to change. Since, in the circuit shown, the output voltage is fed back through a linear resistor, the output voltage is proportional to the pixel output current. The proportionality constant is set by the feedback resistance R. In Section 4, we discuss this topic at greater length.

3 Two-Dimensional Scanning

In this section, we discuss all digital parts of two-dimensional scanning, including selection of rows and generation of video control signals. The horizontal portion of a two-dimensional scanner is identical to a one-dimensional scanner. The added vertical portion performs row selection. This scheme is shown schematically in Figure 5. In the simplest case, the vertical scan will connect a pass-transistor switch inside each pixel in a row of the array. Each pixel in that row will send its current to the horizontal scanner, which will select one of the pixels for output. Section 4 discusses row selection in more detail.

3.1 Making a Raster-Scan Picture

Because of the proliferation of personal computers, many varieties of multiscanning monitors are available at modest cost. These monitors are designed to display output from a variety of different display-adaptor cards; two familiar examples are the EGA and VGA graphics standards for IBM and clone machines. In contrast to the requirements for National Television Standards Committee (NTSC) video (the standard format for broadcast television in the United States), the timing requirements for multiscanning monitors are not nearly as stringent, and there is no need to mix the video brightness, color, and timing information

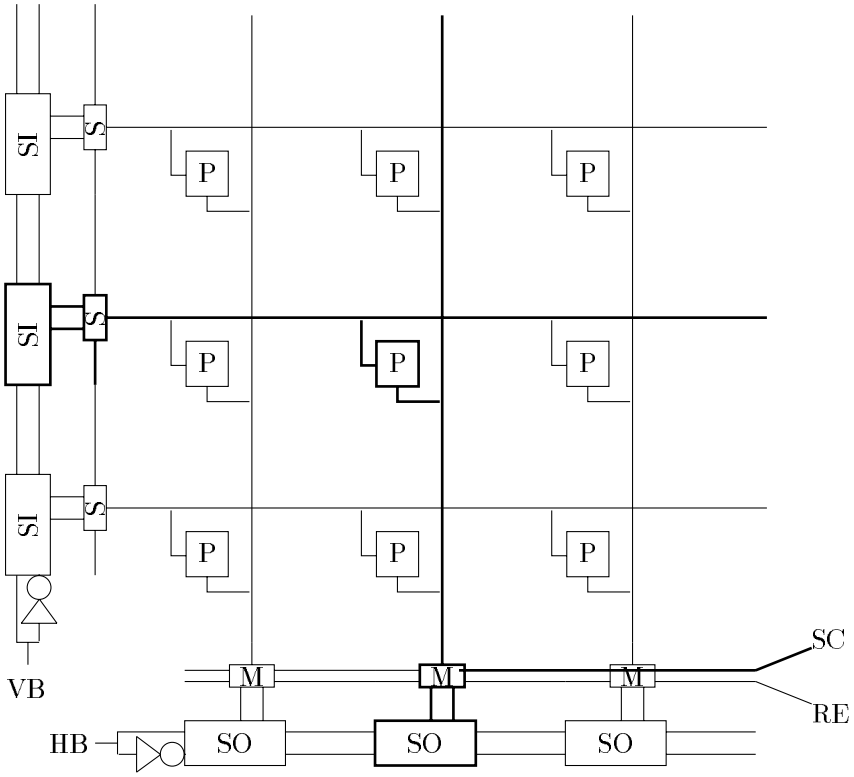

Figure 5: Two-dimensional scanning. Row selection is done by vertical shift-registers (SI). Column selection is done by horizontal shift-registers (SO). Output from row of pixels (P) selected (S) by output from vertical shift-registers is multiplexed, by switches (M) that are controlled by the horizontal shift-registers, onto either scanout wire (SC) or onto reference wire (RE), as for one-dimensional scanner. New bits are loaded into shift-registers as needed (HB and VB). Bold lines show the center pixel in the array selected for scanout.

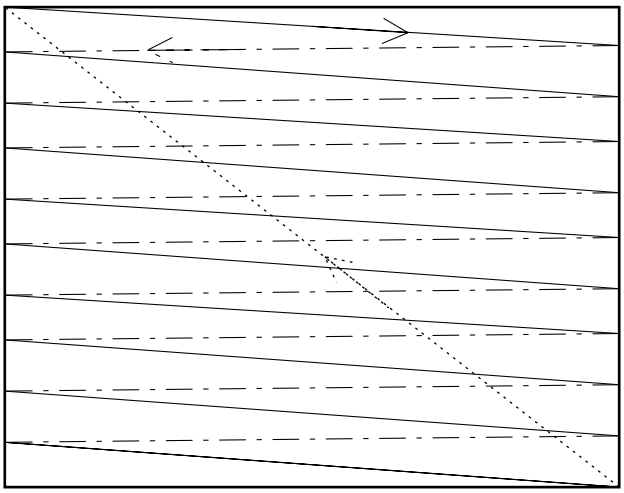

Figure 6: Electron beam movement on monitor display. Horizontal sync occurs after each line; vertical sync occurs at end of frame.

into a single signal. In fact, making a video signal for a multiscanning monitor is relatively simple, so we shall confine our discussion here to the production of such a signal. The resulting signal will not generally be understood by a regular television monitor or video-tape recorder, but is much easier to generate.

In the simplest mode (the only one we discuss here), there are five inputs to the monitor: three analog inputs for the three colors (red, green, and blue (RGB)), and two digital inputs for horizontal and vertical synchronization (sync). The RGB signals will be discussed in Section 4.

Figure 6 shows how the electron beam moves across the monitor face. The beam starts at the upper-left corner of the screen and scans across the screen. After the end of each line, the horizontal-sync signal starts the horizontal retrace. After the last line, the vertical-sync signal starts the vertical retrace. The sync signals tell the monitor when to start the horizontal and vertical retrace.

We generate two additional digital signals that are used to blank the video signal during the horizontal and vertical retrace periods. These are not separate inputs to the monitor; they are used only in our video circuitry. (We could generate a single blank signal, and thus save one pin, but generating separate horizontal and vertical blank is slightly simpler and is a help in debugging.)

The detailed timing requirements for these monitors are given in the owner manuals; what follows is a slightly simplified summary for a particular monitor, the NEC Multisync model II. A video frame must consist of between 200 and 700 lines. The vertical sync must run at 60 $\pm$ 15 Hz. The blank period must be some fraction of the display period, and the sync pulses must occur somewhere near the beginning of the blank period. For a horizontal scan line, the blank period must be 33%$\pm$7% of the display period, and for a vertical scan line, the blank period must be 20%$\pm$10% of the vertical display period. Adherence to these values will almost always produce an acceptable picture, although adjustment of the scan rate will sometimes help produce an a more pleasing aspect ratio or will fill the monitor screen more fully.

3.2 Generation of Sync and Blank Signals

To generate the horizontal and vertical sync and blank signals, we take advantage of the timing already extant in the horizontal and vertical shift-registers by extending the shift-registers that do the row and column selection (the display interval), to include the blank intervals as well. This scheme is shown in Figure 7. The horizontal and vertical scanners act as before, with the addition of extra shift-register stages that encode the sync and blank signals. The sync signal is generated by a wired NAND line that is pulled high when the bit is in the shift-register stages representing the sync signal. Similarly, the blank signal is generated as the NOT of a wired NAND encoding the display period. Thus, there are two additional wired NAND lines for each dimension of scanning. Because the display interval and the sync interval do not overlap, only one additional row of wired NAND transistors are required.

3.3 Counting Rows

Generally, a chip with a complex pixel will have far less than 200 rows of pixels, so we will chose to scan each row several times to fill the rows of the video image with the required number of lines of video. We use a Johnson counter (Figure 8) to count the number of monitor scan lines that display the same row of pixels. The clock for the Johnson counter is the horizontal sync signal; the buffered output of the Johnson counter is the clock for the vertical scanner.

3.4 Crystal Oscillator Circuit for Generating Clock

For video scanners we save ourselves the inconvenience of generating a clock-signal off-chip by using the crystal oscillator arrangement shown in Figure 9. Using this circuit, we generate the main clock signal with an off-chip crystal and a few resistors and capacitors.

3.5 Scanning a Hexagonal Array

Some arrays are best laid out hexagonally. A hexagonal arrangement imposes an additional constraint on scanning, since alternate rows of the array must be delayed by one half of a clock phase to map the chip topography accurately onto the monitor screen. One scheme that we have used is shown in Figure 10. In this scheme, an additional OR line on the vertical scanner encodes whether the row being scanned is even or odd. This signal is used to select either the first half-phase or the second half-phase from the horizontal shift-register stage as the source of the horizontal multiplexer pass-gate signals.

4 Output Representation, Current Sensing, and Video Drivers

In this section, we discuss the analog parts of scanners, including typical transformations between internal pixel representation and output current, the virtual ground scheme, current-sensing amplifiers, and video amplifiers for monitor interfacing.

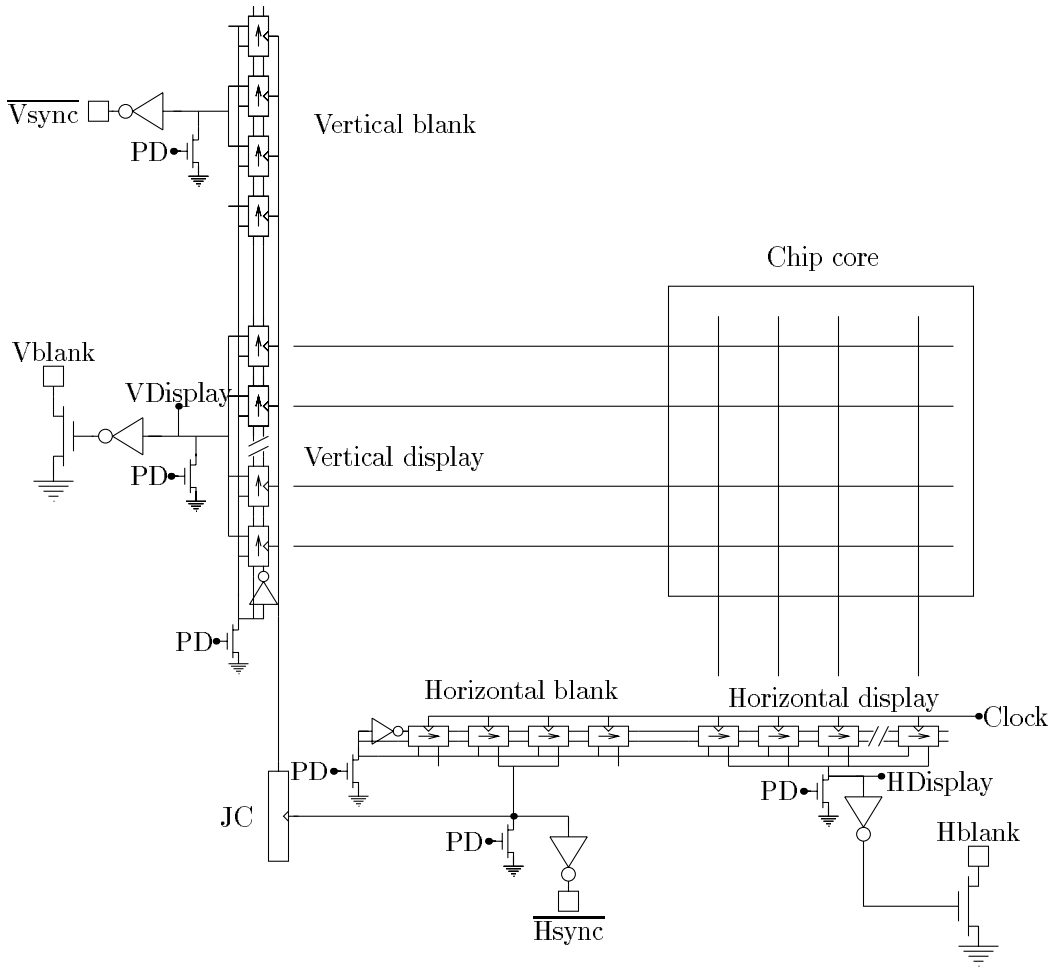

Figure 7: Video scanner with integrated sync and blank signals. Shift-register stages are shown as boxes with arrows showing direction of shift. Clock inputs to shift-registers are shown as carets coming into top of horizontal shift-registers and into right of vertical shift-registers. Wired NAND connections are shown as lines extending outwards from shift-register. Lines extending horizontally into the chip core are row select lines; lines extending vertically into chip core are column output wires. JC is the Johnson row counter. The pulldown transistors (PD) act to pull down the wired NAND lines; these transistors are all biased with a common pulldown bias. Sync and blank output pins are shown as small boxes. The actual scanner is larger; only a few sections are shown for clarity. Actual sizes, for a 50 by 50 pixel chip, are as follows: Horizontal blank section, 11 stages. Horizontal sync stages: stages 2,3 and 4 of horizontal blank section. Vertical blank section: 10 stages. Vertical sync stages: stages 2 and 3 of vertical blank section. Transistor W:L ratios: Sync output inverters: 116:2. Blanking pulldown transistors: 232:2. These large transistors are capable of driving off-chip loads.

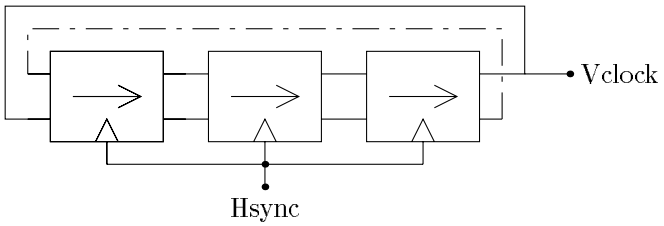

Figure 8: A three-stage Johnson counter. The last stage of the counter feeds back into the first stage with an inversion. Hence, one possible sequence of states is as follows: {0,0,0}, {1,0,0}, {1,1,0}, {1,1,1}, {0,1,1}, {0,0,1}. (There is also a parasitic state consisting of the sequence {1,0,1}, {0,1,0} that is not encountered in practice.) The vertical clock is the buffered output from the last stage; the clock input comes from the horizontal sync signal. An N-stage Johnson counter counts $2N$ lines. Counters with $N \leq 2$ do not have parasitic states.

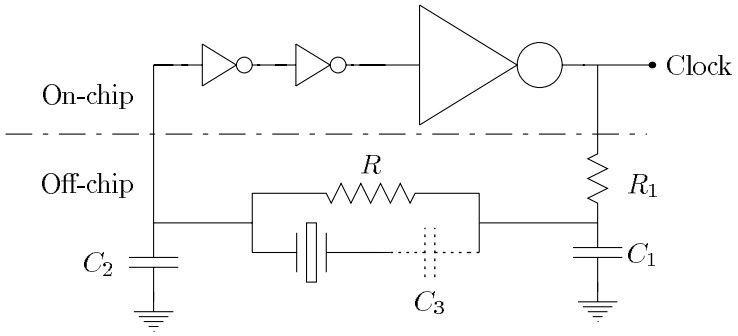

Figure 9: Crystal oscillator circuit. The crystal is driven by and drives the three on-chip inverters. The use of three inverters provides enough gain that the oscillator does not balance. The last inverter is larger than the others to provide the drive for the off-chip components and for a large on-chip fanout. The off-chip resistor R biases the inverters into their high-gain region. At the resonant frequency, the crystal is a short circuit. The inverters provide a 180 degree phase shift, plus some internal delay. The R_1C_1 combination provides additional phase shift, so that the crystal can make up the required 360 degree phase shift. Capacitor C_2 stabilizes the oscillation. Values of components: $R \approx 10M\Omega$, $R_1 \approx 500\Omega$, $1/2\pi R_1C_1 = f$, where f is the crystal frequency, and $C_2 = C_1/4$.
For frequencies above 4 MHz, R_1, C_1, and C_2 may be unnecessary, but $C_3 = 1pF$ may need to be substituted to prevent overdriving the crystal, which results in oscillation at an overtone frequency. If C_3 is not used, the crystal should be driven from C_1 directly, or from the clock output if R_1 and C_1 are not used.
Transistors in first two inverters have W : L ratios of 14:2; last inverter has a W : L ratio of 116:2.

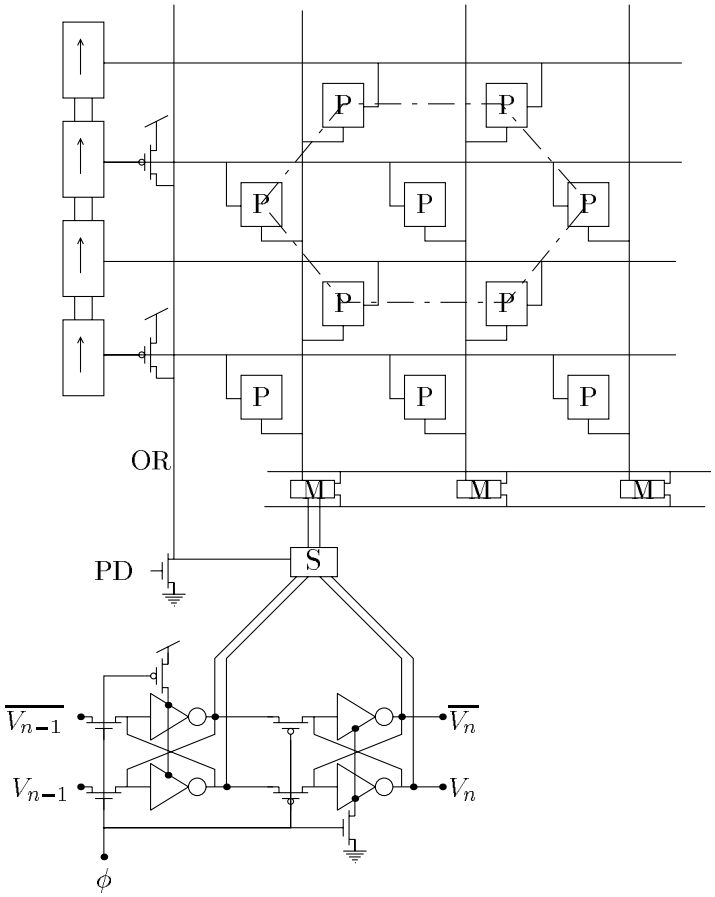

Figure 10: Scanning of a hexagonal array. In the chip layout, alternate rows of the array are shifted by half of a pixel-width. The shift-register is as before. Inverters shown have a power-supply transistor in series with either V_{dd} or ground. Signals from each horizontal shift-register stage pass through MOS switch (S), where either the first half-phase or the second half-phase output from the shift-register is taken to be the switching signal for the output multiplexor switches (M). An odd-row wired NAND (OR) encodes whether an odd or even row is being scanned, and controls which half-phase is taken to be the multiplexer control signal. For the array shown in the figure, the first half-phase is taken for the odd rows of the array, and the second half-phase is taken for the even rows. Only one horizontal shift-register stage is shown in the figure.

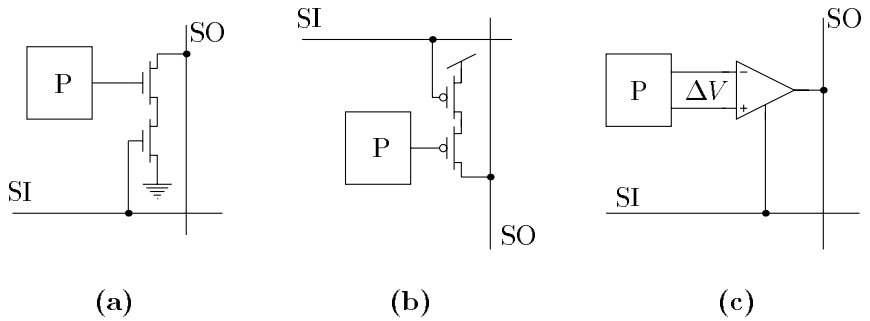

Figure 11: Scanout-selection examples. The scanin (SI) and scanout (SO) lines connect to the vertical and horizontal scanners, respectively. Pixels (P) supply either a single-ended voltage ((a) and (b)) or a differential voltage (c). SI lines select one row of pixels. In (a), SI goes high to select the pixel. In (b), SI goes low to select the pixel. In (c), SI biases the transconductance amplifier to select the pixel.

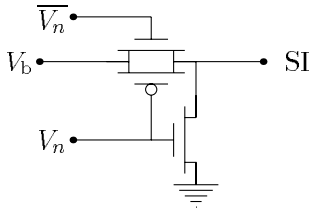

Figure 12: Vertical scanner output circuit used to generate a voltage to bias a row of transconductance amplifiers. V_b is the desired bias voltage, SI is the scanin line, V_n and $\overline{V_n}$ are logic signals coming from the shift-register. V_n goes low to select the row.

4.1 Pixel Output Representation and Row Selection

Figure 11 shows pixel output circuits. In each case, we suppose that the internal pixel representation is a voltage that is converted into an output current that we sense using a current-sense amplifier. In Figure 11 **(a)** and **(b)** , the internal voltage representation is converted into a current using a single transistor. In **(c)**, the internal representation is a differential voltage ΔV that is converted into a current with a transconductance amplifier.

If the scanout requires that a bias voltage be given to each row of the array (as in Figure 11**(c)**), we use a pass-gate arrangement similar to the switches used in the horizontal scanner. The horizontal driver circuit we use for supplying the bias voltage to the row of pixels is shown in Figure 12. The time available during horizontal blanking is sufficient to allow the voltage to settle to the correct value.

4.2 Current-Sense Amplifiers

To sense the current output by the pixels, we use a current-sense amplifier. If the amplifier is off-chip, we use the arrangement shown in Figure 13**(a)**. This amplifier senses a bidirectional current, and the sensitivity is set by the resistance R. The feedback will keep the negative input to the opamp very near V_{ref}; under this condition the current through the feedback resistor must equal the input current, and hence the output voltage will be $V_{\text{ref}} + IR$.

For the off-chip current-sense amplifiers, we successfully use a TL074 opamp; the only

disadvantages are the requirement for a $+12V/-12V$ power supply and the limitation to the pixel rate to less than 1 MHz. Other opamps may be faster, but are more difficult to stabilize. We often use the extra amplifiers (the TL074 comes in a quad configuration), in follower configuration, to supply the reference voltage V_{ref} or other reference voltages used in the core of the chip.

On-chip, we use unidirectional sense amplifiers like those shown in Figure 13 **(b)** and **(c)**. The analysis given in Appendix A shows that when the current is in subthreshold, the output voltage is logarithmic in the current I:

$$V_{\text{out}} = \frac{V_{\text{ref}} + \frac{kT}{q}\ln(I/I_0)}{\kappa}.$$

I_0 is the leakage current, and $\kappa \approx 0.7$ is the back-gate coefficient. Above threshold, the output voltage is related to the square root of the input current:

$$V_{\text{out}} = \frac{V_{\text{ref}} + V_{\text{T}} + \sqrt{I/I_0'}}{\kappa}.$$

V_{T} is the threshold voltage, and I_0' is a constant with units I/V^2.

It is intuitively clear that, the tighter the feedback amplifier clamps the sense line, the more speedup the feedback arrangement will provide. We quantify the speedup by defining a natural time scale that is the open loop time constant of the pixel output line,

$$\tau_{\text{in}} = \frac{C_{\text{in}}}{G_{\text{fb}}},$$

where C_{in} is the capacitance of the line and G_{fb} is the source conductance of the feedback transistor. This time constant represents the speed of a current-sensing arrangement in which we measure the voltage at the source of the feedback transistor, holding the gate and drain of the feedback transistor at a constant voltage. The analysis in Appendix A shows that the speedup provided by using a fast feedback amplifier with voltage gain of A is just κA; in other words, the resulting first-order time constant of the output is

$$\tau_{\text{out}} = \frac{\tau_{\text{in}}}{\kappa A},$$

where $\kappa \approx 0.7$ is the back-gate coefficient. However, if the feedback amplifier is too slow, then the output signal will ring. To prevent ringing, we must ensure that the feedback amplifier is at least $4\kappa A^2$ times faster than the input node. The implication of this result is that, although a high-gain feedback amplifier is desirable for maximum speedup, to prevent ringing, we may have to settle for a gain of 100 or less. In our amplifiers, we use minimum-length transistors, and we run the bias of the amplifiers high enough that the gain is actually reduced by above-threshold effects.

The transconductance amplifiers shown in Figure 13 **(b)** and **(c)**, are simple single-stage transconductance amplifiers [3]. Since the feedback amplifier must run much faster than the input node to prevent output ringing, we build our feedback amplifiers using ring transistors, where the inside of the ring is the drain of the transistor. The ring-like arrangement allows

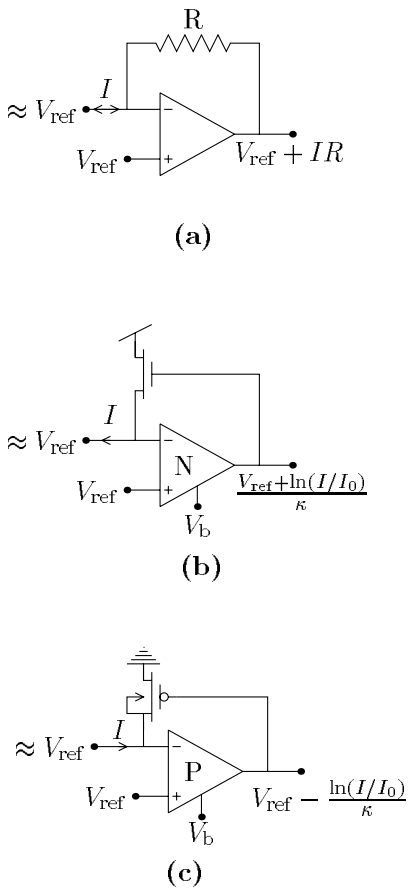

Figure 13: Three current-sense amplifiers: (a) a bidirectional off-chip linear arrangement; (b) and (c) unidirectional on-chip logarithmic amplifiers. For (b) and (c), output voltages are shown for the subthreshold case. Both (b) and (c) are on-chip simple transconductance amplifiers [3] biased with bias $V_{\rm b}$. N and P refer to the type of the amplifier: N means that the differential pair should consist of native devices, P means that the differential pair should be constructed from transistors in the well. In (c), we show the feedback transistor in its own well, with the well tied to the source of the transistor. This modification will prevent the back-gate effect from requiring an output voltage below ground for a large current I and a $V_{\rm ref}$ near ground.

very wide transistors to be placed in a small area, with minimal side-wall capacitance. The resulting amplifier has relatively low gain but high transconductance.

We must take care with the logarithmic sense amplifiers shown in Figure 13 to ensure that the amplifiers are operating in their proper voltage ranges. A simple transconductance amplifier with a native-type bias transistor and differential pair will only operate correctly when the output voltage is above $V_{\min} = \min(V_+, V_-) - V_b$. V_+ and V_- are the amplifier input voltages, and V_b is the bias voltage. This condition is satisfied for the circuit in Figure 13**(b)**. However, the amplifier in Figure 13**(c)** will output a voltage closer to ground than V_{ref}. Hence, the amplifier must be able to operate when the output voltage is substantially below both of the inputs. In this case, a well-type differential pair must be used. In addition, if the pixel requires that V_{ref} be held near ground, the back-gate effect on the exponential feedback transistor would require that the output voltage be below ground. Hence, we usually put this feedback transistor in its own well and tie the source of the transistor to the well. For the logarithmic sense amplifier in Figure 13**(b)**, this arrangement is not possible because the feedback transistor is in the substrate.

If we use these current-sense amplifiers in a video driver, it is essential that we introduce the minimum amount of stray capacitance. Hence, we buffer the output of the current-sense amplifiers, using a voltage follower, before driving a large capacitive load. For the same reason, schemes that place the feedback element off-chip have difficulty at video rates, due to the pad, package, and circuit board capacitance.

In Section 5 we discuss performance of the current-sense amplifiers. In Appendix A, we analyze the logarithmic current-sense amplifier.

4.3 Video Amplifiers

The RGB lines are AC coupled inside the monitor. Hence, their DC levels do not matter. The total brightness range is generally about one volt from black level to full saturation. The brightness of the image is given by the contrast in the video signal between the video level and the blank periods surrounding the sync signals. It is up to the user to provide the video blanking.

The transition from the internal video signal, as computed by the on-chip sense amplifiers, to the off-chip video signal that drives the monitor is often the most painful part of getting a chip running. The monitor RGB inputs are terminated with a standardized 75Ω load — a load that requires very large transistors to drive directly. We use smaller output transistors and off-chip amplifier arrangements like those shown in Figure 14 to do the impedance matching, and to provide maximum flexibility in interfacing to a particular monitor. These video drivers have been evolved to use the minimum number of off-chip components, and function with a minimum amount of trial-and-error component twiddling.

The on-chip current-sense amplifier drives an on-chip voltage follower, which drives the gate of a large on-chip output transistor Q_1. If this transistor is of the same type as the feedback transistor in the logarithmic current-sense amplifier, then the current flowing in the output transistor will be proportional to the input current to the current-sense amplifier. The output transistor has both source and drain coming out to pads, giving us the flexibility to do level and gain adjustment off-chip. The output transistor is incorporated into the off-chip video driver in an inverting mode. The source of the output transistor is tied to

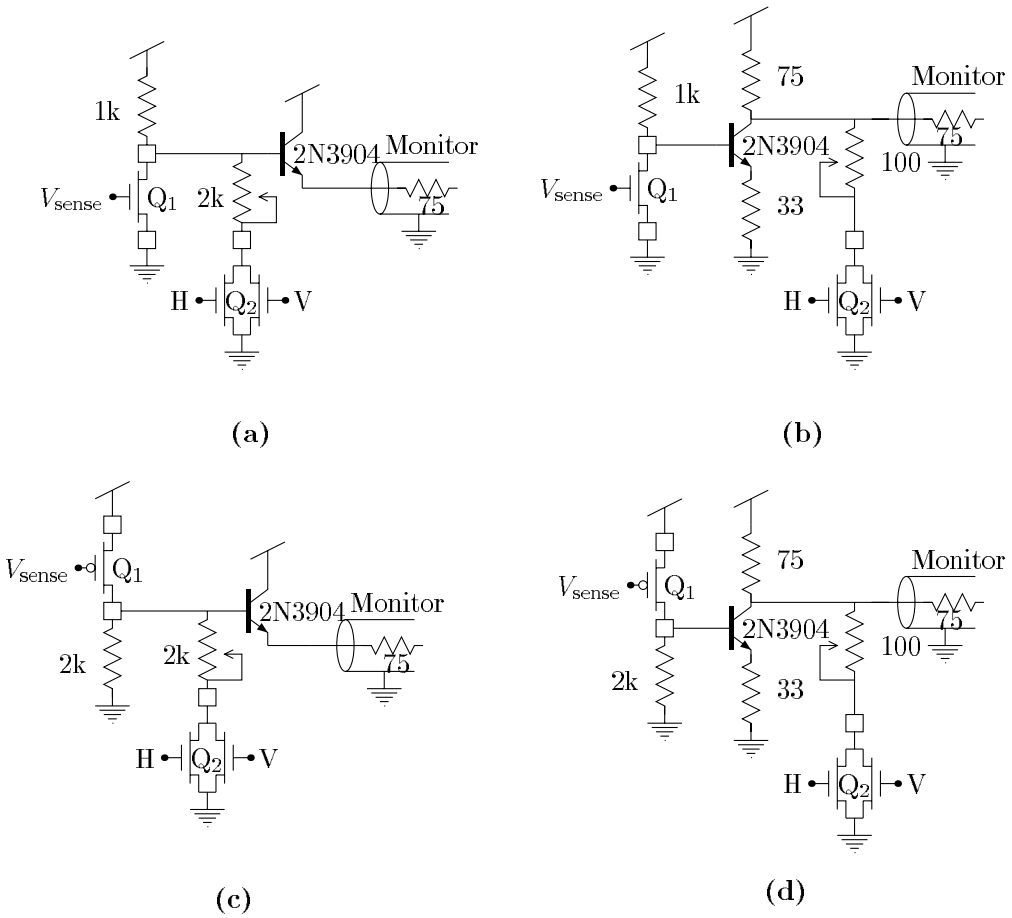

Figure 14: Video-driver amplifiers. (a) Driver used for inverting internal signal with n-type output transistor. (b) Driver used for noninverting internal signal with n-type output transistor. (c) Driver used for inverting internal signal with p-type output transistor. (d) Driver used for noninverting internal signal with p-type output transistor. The Q_1 and Q_2 FETs are on-chip, and all other components are off-chip. V_{sense} comes from voltage-follower driven by output of current-sense amplifier. Bonding pads are shown as small boxes. H and V are the on-chip horizontal and vertical blank signals; they drive the blanking FETs Q_2. The 75 Ω resistor shown in the monitor coax is inside the monitor; we do not supply it. The potentiometers are used to adjust the blanking level, and hence, the brightness and contrast of the video image. FET W : L ratios: Blank transistors, 232:2; Output transistors, 950:2. Blank transistor ohmic resistance at V_g=5 V: n-FET, 140Ω; p-FET, 240Ω.

the appropriate rail, and the drain is pulled to the opposite rail with a resistor. To drive the monitor input, the output of this inverting amplifier is then either amplified with an inverting amplifier or followed with an emitter follower. The on-chip blanking transistors pull down on the appropriate place in the video driver circuit to blank the video signal. For the inverting configuration, the blanking transistor pulls down on the output of the inverting amplifier. For the non-inverting, emitter follower, the blanking transistor pulls down on the input to the emitter follower. The four possible combinations of output transistor type and inverting–noninverting amplifier type shown in Figure 14 cover most possible situations.

The preceding discussion has assumed that we are generating only a single color. If we want to generate a white picture, we connect the output from the video-driver to all three RGB inputs, or duplicate the output driver for each color. If we are scanning out more than one signal from the chip, we may want to generate a separate color with each channel. In this case, the blanking circuitry shown in Figure 14 is slightly more complicated, since we cannot directly connect the separate video signals to a common blank.

In Section 5, we discuss the performance of the video drivers.

5 Performance

In this section, we give measurement results on the power consumption and speed of a typical scanner.

5.1 An Example Scanner

Figure 15 shows a photograph of a silicon retina equipped with the two-dimensional video scanner. The chip has 68 rows and 43 columns of pixels, and each row is counted 6 times. Figure 16 shows the sync and blank outputs from this video scanner. This chip was fabricated in 2μm p-well technology through MOSIS. The behavior of n-well chips is similar. The timing shown in this figure produced excellent framing over a range of ± 10 % clock frequency.

A photograph of the complete video board is shown in Figure 17. The attached cables carry power and video. We use a power supply that provides +5 V and $\pm$ 12 V outputs (Elpac, Model WM113); this particular chip uses an on-chip sense amplifier and so does not require an off-chip opamp powered by the ± 12 V supply. The lens (Fairchild Cinphar, 13 mm, $f1.8$) is from an old movie camera; the lens mount is custom. A Zero Insertion Force (ZIF) socket holds the chip for easy testing. Five 10-turn potentiometers supply bias voltages. Several 10 μF and 100 μF tantalum capacitors reduce supply ripple and help stabilize the V_{ref} reference wire. V_{ref} is supplied from a 100 Ω potentiometer; the other potentiometers are 10 $k\Omega$.

Figure 18 shows the monitor screen when the retina is looking at a hand. The parasitic charge-injection and clock coupling transients are visible as vertical lines on the screen. This silicon retina has only 43 by 68 pixels, yet even with this relatively low resolution, many features are still easily recognizable. In a dynamic image, unlike this static photograph, the image becomes much more salient.

Figure 15: Photomicrograph of a fabricated silicon retina with scanner. The chip measures approximately 4600 by 6800 microns. There are 43 columns and 68 rows of pixels. The horizontal sync and blank shift-registers and the Johnson counter are at the lower left of the chip. At the upper left are the vertical sync and blank shift-registers. At the lower right are the sense amplifiers and output pads. The pads at the upper right of the chip provide bias voltages for the core of the chip.

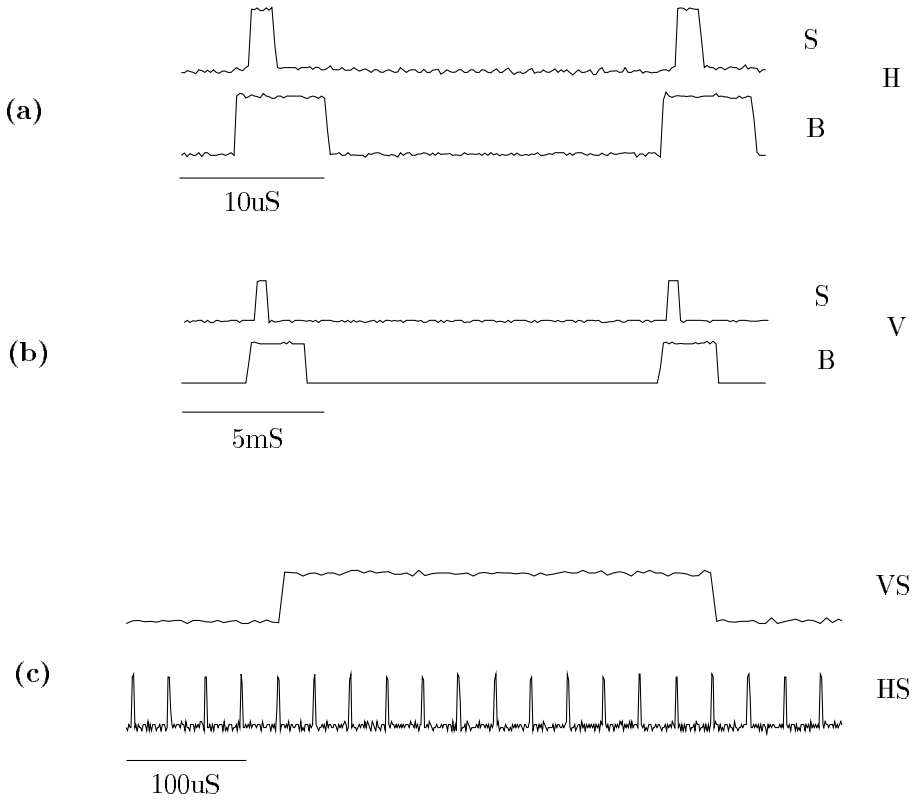

Figure 16: Sync (S) and blank (B) signals from the video scanner shown in Figure 15. (a) Horizontal (H) sync and blank signals. (b) Vertical (V) sync and blank signals. (c) Vertical and horizontal sync signals shown together. This chip has 68 rows and 43 columns. Each row is counted six times, for a total of 408 lines of video.

Figure 17: Complete video scanner board.

Figure 18: Monitor screen showing output of silicon retina.

5.2 Power Consumption and Speed

The digital parts of the scanner in Figure 15 consume 6 mW of power when running at a 1.8 MHz clock rate with a 5 V supply voltage. At a clock frequency of 1.8 MHz, the vertical retrace frequency is 60 Hz. This scanner is capable of running at up to 11 MHz, far in excess of the required speed for chips of this complexity.

5.3 Current-Sense Amplifier and Video-Driver Performance

The on-chip logarithmic current-sense amplifiers shown in Figure 13 have a usable bandwidth that is dependent on the pixel output current and on the voltage gain of the feedback amplifier. We use ring transistors with a W : L ratio of 20:2 in our feedback amplifiers, yielding a gain $A \approx 40$ and an effective speedup κA of about 25. When the pixel output current is very small, the current-sense amplifier will not stabilize the input line fast enough to drive the output at video rates. For moderate pixel output currents, on the order of tens of nanoamperes, the natural time constant of a typical pixel output line of one picofarad capacitance is less than one microsecond, and the bandwidth of the system begins to be limited by the bandwidth of the video-driver circuitry. As an example, the maximum measured bandwidth for the current-sense amplifier in Figure 13(**b**) driving the video driver in Figure 14(**a**)is 8 MHz. We have built complete systems with bandwidth in excess of 5 MHz, more than sufficient for display of arrays of more than 100 by 100 pixels.

6 Practical design details

In this section, we discuss practical design details and give a source for example layout.

6.1 Technology

We have fabricated and tested these scanners using a wide range of fabrication processes. The scanner design itself is generic, and we use the same layout for both n- and p-well technologies. Monitors generally trigger on the low-going edges of the sync signals, but we have found that the same sync generation output circuitry used for n-well technology works for p-well also. The monitors care about the sign of the video-signal voltage, however, so we must pay some attention to the actual technology when designing output circuitry. For example, the video-driver circuits shown in Figure 14 all utilize n-type blanking transistors that pull down to ground. The identical generic CIF layout, fabricated in the complementary technology, will result in p-type blanking transistor that pull up to V_{dd}. In this case, we can modify the video-drivers in Figure 14 **(b)** and **(d)** so that the blanking transistors pull up on the input to the inverting video-driver. The drivers in Figure 14 **(a)** and **(c)** , however, would require additional off-chip active components to function correctly.

6.2 Power-Supply Separation

We have found it very beneficial to separate the power supply for the scanner and core of the chip. In fact, we often have four power-supply pins: one for the digital scanner, one for the clock driver, one for the video circuitry, and one for the core of the chip. This added flexibility is provided at the cost of a few pins and often lets us isolate problem areas having to do with power utilization and latchup. In addition, the sensitive analog circuitry is isolated from the relatively large clock noise generated by the digital scanner.

6.3 Common design errors

In our experience in fabricating more than 100 chip designs with these scanners, we have encountered five common design errors. Some of these errors may be traced to inadequate design tools, such as design rule checkers that do not easily check entire designs or particular parts of large arrays, or circuit netlist extractors that do not easily extract complete designs. Other errors are logical choices of sign or transistor type.
Logical design errors:

Logical design errors in select sign. When designing a pixel that requires a logic high signal for row selection, we have often selected the wrong sign of logic signal. It is easy to tell which of the two possible signals are being sent to the pixel, since one of the two signals must go to the wired OR. This signal must be the logic low signal. Similarly, we have also chosen the wrong sign bit in the horizontal multiplexer, resulting in a situation where all columns, except the one we are scanning out, feed into the current-sense amplifier.

Sense amplifier of wrong type. When using a logarithmic current-sense amplifier such as those in Figure 13, parts **(b)** and **(c)**, we have built the amplifier using a simple transconductance amplifier of the wrong type. This arrangement will not work, since the transconductance amplifier will not operate in its defined output-voltage range.

Reference wire with too much resistance. If the dummy line carrying the pixel-output reference voltage V_{ref} has too much resistance, the dummy line will not be held at V_{ref}, which will greatly slow the output current-sensing bandwidth. Therefore, it is essential that the V_{ref} wire not have poly sections.

Design errors due to inadequate design tools:

Power supply and ground shorts. This problem invariably occurs at the corners of the chip, where the core of the chip abuts the inside of the scanner frame. It only occurs when the design tools are incapable of extracting the netlists for large circuits.

Design-rule violation. It is crucially important to check the design rules, particularly at the abutments and corners of the chip. These are the places where well spacing, well-active spacing, and active-select rules are most likely to be violated. This error only occurs with primative design-rule checkers.

6.4 Layout for Scanners

The value of exploratory design has been greatly enhanced by the availability of fast prototyping through the MOSIS fabrication service. It is no longer necessary that even relatively complicated projects be group efforts, involving months of design, and elaborate negotiations between foundry and customer. Instead, projects now can be designed by individuals in a few days.

Using any one of several readily available and affordable design tools, a designer can lay out the required circuitry for the chip core. This circuitry is integrated with a scanner frame using a simple silicon compiler that places pixels in rows and columns and abuts the scanner cells. To verify the layout of the chip, a layout extractor is used to generate the netlist of a small version of the circuit. Using a netlist comparison program, this netlist is compared with a netlist extracted from a schematic of the chip. When the design verifies correctly, a full-sized version of the chip is compiled and is electronically mailed to MOSIS. Six to eight weeks later, the packaged project comes back, ready to test.

Relatively complicated parts of designs, such as the scanners described here, are non-exploratory and exist only for circuit testability purposes. They still require considerable design and debugging, and a single mistake can kill the entire scanner operation. The aim of this paper is to avoid a time-consuming duplication of the effort that has gone into development of these scanners. Hence, we are making the layout and schematics for these scanners available via anonymous ftp. To access these sources, you can ftp to the internet host **hobiecat.cs.caltech.edu** (preferably during off-hours), log in as **anonymous**, and give your name as password. The scanner directory is **/usr/ftp/pub/scanners**. A README file in that directory will provide further details.

7 Summary and Conclusion

In this paper, we have given a tutorial description of how to design both one- and two-dimensional scanners, with emphasis on practical aspects of design and on the production of useful output images or signals. In our laboratory, we have used scanners on over 100 chip designs over the past three years. In fact, scanners act as our eyes into the microcosm of large arrays of VLSI elements, allowing us to observe the collective nature of the computation that the chip is performing, or the distribution and global effect of circuit offsets. For many circuits, such as retina or cochlea models or visual motion models, scanners provide insights into the spatio-temporal structure of the network operation that would be lost in a static view of a single node. For certain auditory processing chips, the use of a two-dimensional video scanner becomes essential in the process of converting the chip operation into a signal that can be understood and processed by our own visual system into a coherent picture of the computation. In these circuits, a global view is essential in understanding the circuit operation.

We do not regard these scanners as useful for chip-to-chip communication. The well-known pitfalls of sending analog information off-chip, with the attendant corruption by digital noise and the problem of matching references and other values between analog chips, is an obvious concern. Of equal concern, however, is the fundamental problem of temporal aliasing and bandwidth-utilization. Even if we could noiselessly send analog values from one chip to another (or several others), these signals would still be discretely sampled in time — not a desirable quality for an analog, continuous-time, processing system. A serial sampler, such as a scanner, would not be biased in its sampling, but by the same token, it would waste its bandwidth on uninteresting events. That is, a scanner would send information from a location on the sending chip even if nothing were changing there. More sensible, in this regard, is a scheme that "spends" bandwidth on locations with interesting events [1]. These "event-driven" scanners remain to be proven practical, as do many aspects of multi-chip analog design; in the meantime, serial scanners will be indispensable adjuncts of exploratory analog VLSI design.

A Analysis of logarithmic current-sense amplifier

In this appendix we work out in detail the characteristics of the logarithmic current–sense amplifier shown in Figure 13(**b**). In Figure 19 we have redrawn this circuit, along with associated capacitance and output conductance used in the analysis. In the following discussion, we employ the grounded-substrate convention, and we measure all voltages in terms of the thermal voltage $V_t = \frac{kT}{q}$, and all currents in terms of the leakage current I_0. For example, a current $I = e^V$ is really $I = I_0 e^{\frac{V}{V_t}}$, and a conductance $g = \frac{C}{I}$ is really $g = \frac{C}{I/V_t}$.

A.1 DC Characteristics

We can work out the DC transfer characteristics of this circuit as follows. A native transistor operating in subthreshold (weak inversion) obeys the I-V relation

$$I_{ds} = e^{\kappa V_g - V_s}.$$

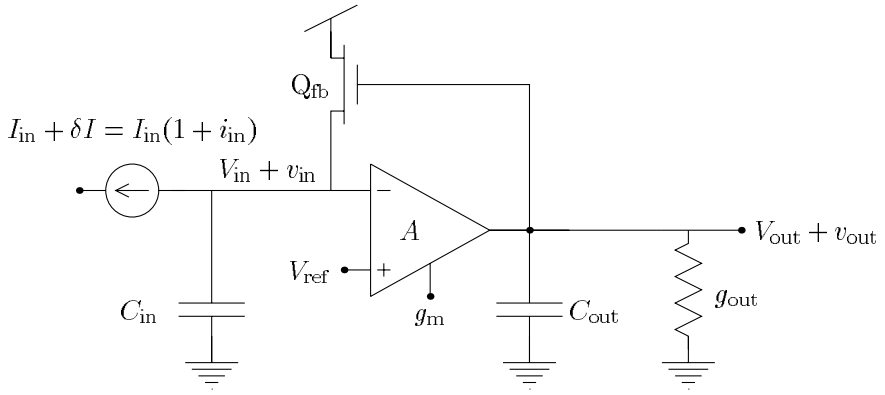

Figure 19: Logarithmic current-sense amplifier. The exponential feedback element is Q_{fb}. $I_{in}+\delta I$ is the input current, $V_{out}+v_{out}$ is the output voltage, $V_{in}+v_{in}$ is the input voltage, g_m is the transconductance of the feedback amplifier, g_{out} is the output conductance of the feedback amplifier, $A = g_m/g_{out}$ is the voltage gain of the feedback amplifier, and C_{in} and C_{out} are associated capacitances. Voltages and currents denoted by upper-case letters are large signal quiescent values; those denoted by lower-case letters are small-signal variations.

Here, I_{ds} is the current between drain and source, $\kappa \approx 0.7$ is the back-gate coefficient, V_g is the gate voltage, and V_s is the source voltage. We assume here that the feedback transistor Q_{fb} is saturated (drain-source voltage is large). Since the feedback will act to keep the negative input to the amplifier very near to V_{ref}, and since the input current I must be supplied by the feedback transistor, we find that the output voltage (i.e., the gate voltage) will be

$$V_{out} = \frac{V_{ref} + \ln I_{in}}{\kappa}.$$

The same analysis applies for the circuit in Figure 13**(c)**, except for the obvious change in sign and the fact that the source of the feedback transistor is at the bulk potential ($V_s = 0$).

The analysis is similar above threshold, except that we use the above-threshold I-V relationship for saturation current:

$$I_{ds} = (\kappa V_g - V_s - V_T)^2,$$

where I_{ds} is the drain–source current, $\kappa \approx 0.7$ is the back–gate coefficient, V_g is the gate voltage, V_s is the source voltage, and V_T is the threshold voltage. This I-V relationship gives the output voltage

$$V_{out} = \frac{V_{ref} + V_T + \sqrt{I_{in}}}{\kappa},$$

with an analogous inversion for the opposite type of feedback element.

A.2 Transient Characteristics

A slightly simplistic analysis of the circuit in Figure 19 treats the feedback amplifier as having finite voltage gain A but infinite transconductance g_m. In this case, the equation governing the input node V_{in} is

$$C_{in}\dot{V}_{in} = -\delta I + e^{\kappa V_{out} - V_{in}}.$$

The equation governing the output node is simple:

$$V_{\text{out}} = -AV_{\text{in}}.$$

We go to the small-signal regime by treating all voltages and currents as small variations around their DC level, to obtain

$$C_{\text{in}}\dot{v_{\text{in}}} = -\delta I + I_{\text{in}}(\kappa v_{\text{out}} - v_{\text{in}}),$$

and

$$v_{\text{out}} = -Av_{\text{in}},$$

where the lower-case symbols refer to small-signal variations. Dividing by I_{in}, defining $i_{\text{in}} = \delta I/I_{\text{in}}$, $\tau_{\text{in}} = C_{\text{in}}/I_{\text{in}}$, and substituting v_{out} for v_{in} in the first equation, we obtain

$$\frac{\tau_{\text{in}}}{A}\dot{v_{\text{out}}} + (\kappa + 1/A)v_{\text{out}} = i_{\text{in}}.$$

We note here that the definition of τ_{in} given here is the same as that given in Section 4.2: $\tau_{\text{in}} = C_{\text{in}}/G_{\text{fb}}$, since in subthreshold, $G_{\text{fb}} = I_{\text{in}}/V_{\text{t}}$. We easily find the transfer function by going to the s-plane (Siebert, 1986). Here, $s = \sigma + j\omega$ is the Laplace transform parameter, and we obtain

$$(\frac{\tau_{\text{in}}}{A}s + \kappa)v_{\text{out}} = i_{\text{in}},$$

where we have also used the approximation that $A >> 1$. The single root of the transfer function is

$$s_0 = \frac{-\kappa A}{\tau_{\text{in}}},$$

corresponding to a first-order time constant $\tau = \tau_{\text{in}}/\kappa A$.

A more sophisticated analysis includes both a finite gain and a nonzero time constant for the feedback amplifier. The analysis of this second-order system is straightforward, however, so we shall just show the resulting transfer function:

$$v_{\text{out}} = \frac{i_{\text{in}}}{\tau_{\text{out}}\tau_{\text{in}}s^2 + (\tau_{\text{out}} + \frac{\tau_{\text{in}}}{A})s + \kappa + 1/A}.$$

$\tau_{\text{out}} = C_{\text{out}}/g_{\text{m}}$ is the time constant of the feedback amplifier. As before, $A = g_{\text{m}}/g_{\text{out}}$ is the gain of the feedback amplifier. We note here that below threshold, A is a constant when g_{m} changes, since the Early effect on the output transistors causes g_{out} to scale with g_{m}.

Figure 20 shows a root-locus plot of the roots of the denominator of this transfer function. This plot shows how the two poles of the resulting second-order system move in the s-plane in response to changes in the time constant τ_{out} of the feedback amplifier. For $\tau_{\text{out}} = 4\tau_{\text{in}}$, the poles are coincident at $s = -1/2\tau_{\text{in}}$, and the response is critically damped but slow. As τ_{out} decreases, the poles split apart from the real axis, forming a conjugate pair, and the response speeds up but also starts to ring. The ringing is maximal when $\tau_{\text{out}} = \tau_{\text{in}}/A$; at this point, the Q of the circuit is $A = \sqrt{\frac{\kappa A}{2}}$. At a value of $\tau_{\text{out}} \approx \tau_{\text{in}}/4\kappa A^2$, the poles reconverge onto the real axis at $s \approx -2\kappa A/\tau_{\text{in}}$, and the response is again critically damped. (These values for τ_{out} assume $A >> 1$.) As τ_{out} approaches zero, the poles split apart along the negative real axis. As we expect, one pole ends up at $s = -\kappa A/\tau_{\text{in}}$; the other shoots off to negative infinity. Figure 21 shows representative step responses of the circuit in Figure 19 in the regimes discussed.

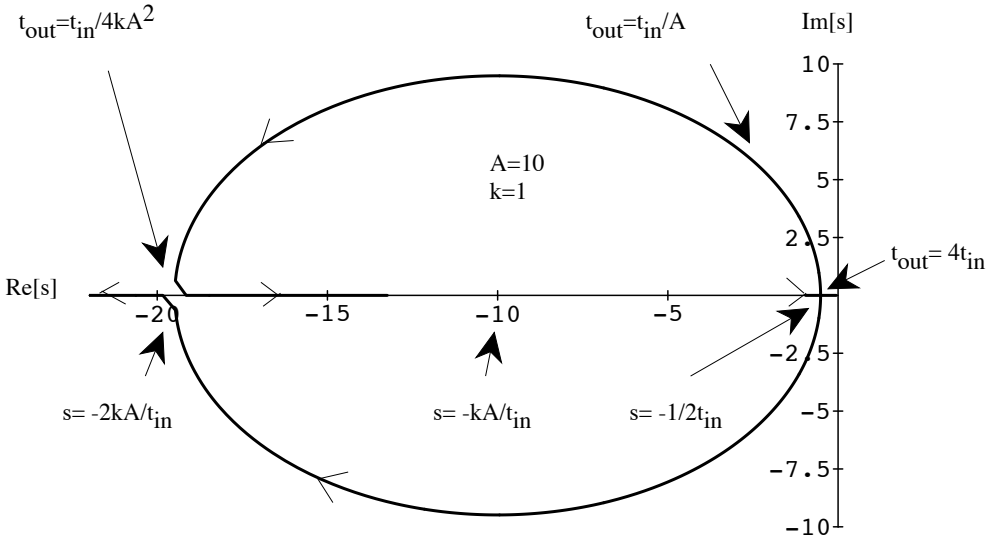

Figure 20: Root-locus plot of roots of transfer function as transconductance τ_{out} of the feedback amplifier is varied. $A = 10$, $\kappa = 1$, $\tau_{\text{in}} = 1$.

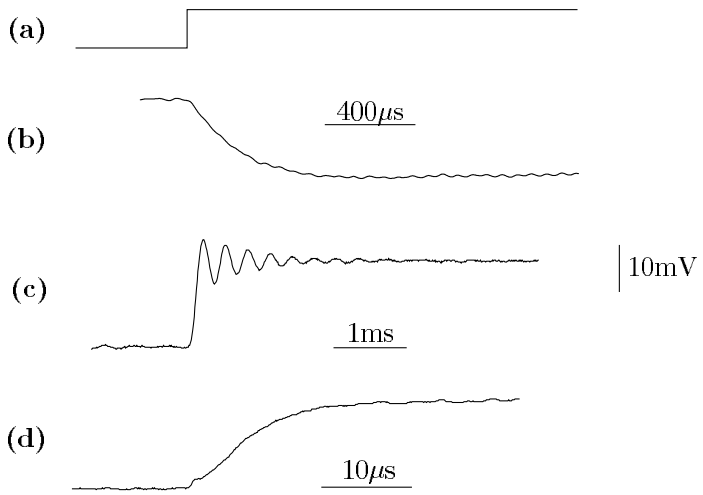

Figure 21: Step responses of logarithmic current-sense amplifier in regimes discussed in Appendix A. (a)Input current step. (b) Open loop response of input node v_{in} when feedback is so slow that v_{out} does not move over time scale shown. (c) Response ringing in output node v_{out} when τ_{out} is decreased. (d) Response of v_{out} when τ_{out} is made small enough that the response is again critically damped. Measurement of the first-order time constants in (b) and (d) shows that the effective speedup κA caused by feedback amplifier is about 25. Measured DC open loop gain of amplifier is about $A = 41$. Measured κ at operating source voltage is $\kappa = .79$. Theoretical value for speedup is $\kappa A = 32.4$. Discrepancy is probably accounted for by Miller capacitance, which dynamically multiplies any parasitic feedback capacitance between output drain and negative input gate by DC gain of feedback amplifier.

B Acknowledgments

Our thanks to the Office of Naval Research for their support under grant NAV N00014-89-J-1675, and, of course, to the MOSIS fabrication service. Thanks also to Mohammed Ismail, Shih-Chii Liu, Lyn Dupré, Xavier Arreguit, Buster Boahen, and anonymous reviewers for helpful comments and discussions.

References

[1] M.A. Mahowald, "Evolving Analog VLSI Neurons," in T. McKenna, J. Davis, S. Zornetzer (eds.), *Single Neuron Computation, Proceedings of ONR Meeting, Woods Hole, MA,* Academic Press, to be published 1991 or 1992, Chap 15.

[2] M.S. McGregor, P.B. Denyer, and A.F. Murray, "A single-phase clocking scheme for CMOS VLSI," in *1987 Stanford Conference on Very Large Scale Integration,* MIT Press: Cambridge, MA, 1987, pp. 257–271.

[3] C.A. Mead, *Analog VLSI and Neural Systems*, Addison–Wesley: Reading, MA, 1989.

[4] C. Mead and J. Wawrzynek, "A new discipline for CMOS design," in H. Fuchs (ed.), *1985 Chapel Hill Conference on Very Large Scale Integration.* Computer Science Press: Chapel Hill, NC, 1985, pp. 87–104.

[5] M. Sivilotti, *Wiring Considerations in Analog VLSI Systems, with Application to Field-Programmable Networks.* Ph.D. thesis, Dept. of Computer Science, California Institute of Technology, Pasadena, CA, June 1991.

[6] M.A. Sivilloti, M.A. Mahowald, and C.A. Mead, "Real-time visual computations using analog CMOS processing arrays." in Losleben, P. (ed.), *Proceedings of the Stanford Conference on Very Large Scale Integration.* MIT Press: Cambridge, MA, 1987, p. 295–311.

[7] L. Watts, "One dimensional resistive networks and scanners," Internal laboratory memorandum, Physics of Computation Laboratory, California Institute of Technology, Pasadena, CA, 1990.

[8] G. Wegmann and E.A. Vittoz, "Analysis and improvements of accurate dynamic current mirrors," *IEEE Journal of Solid-State Circuits* vol. 25(3) pp. 699–706, 1990.