Proceedings of the 1990 IEEE International Conference on Systems, Man and Cybernetics, Los Angeles, California, November 4-7, 3 pages 1990

IMAGE SMOOTHING AT VIDEO RATES WITH ANALOG VLSI

Andrew Moore and Rodney Goodman
Computation and Neural Systems Program
MS 116-81, California Institute of Technology
Pasadena, CA 91125

Abstract

Image smoothing is an important computational primitive in both artificial and biological vision systems. A resistive grid forms a suitable substrate for this operation in both types of systems. Previous artificial systems using this substrate form the image for smoothing either with on-chip photoreceptors in real time or with digitally driven input to an analog sample-and-hold system at rates far below the video frame rate. We have designed, fabricated, and successfully tested a subthreshold CMOS analog VLSI chip which, with a minimum of supporting circuitry, can smooth an image formed from a conventional video signal, at the video frame rate.

Introduction

Image smoothing is an important operation in artificial and natural vision systems. Commonly this is done in man-made systems by digitizing the output of a video camera, performing fast convolutions digitally and, if display of the smoothed result is required, reconverting the image to analog voltages compatible with video monitors. Such systems are compact only if the degree of spatial smoothing is limited to a few pixels.

Analog methods of smoothing images have been previously described by Mead [1]. In the analog retina, the image is formed by on-chip photoreceptors; the smoothing is carried out in real time with a resistive grid. In contrast to conventional digital techniques, smoothing with a resistive grid built from subthreshold analog VLSI is performed with very low power consumption, and with a compactness that does not vary with the degree of smoothing; the equivalent convolution kernal can be varied from in size from a small fraction of the image (one pixel) to the whole image with no loss of compactness and little increase in power consumption. Others [2] have accomplished smoothing with resistive grids, but only "off-line" - the image is input to the chip by an external digital system, at rates much lower than the frame rate.We have designed, fabricated, and tested a chip that, for the first time, can receive images directly from a video camera and perform smoothing via a resistive grid at the video frame rate.

System Description

Physical description of the chip

The chip was fabricated through MOSIS with 2 micron CMOS design rules. It is approximately 5mm by 7 mm in size, and contains a 46 by 47 node hexagonal resistive grid, sample-and-hold circuitry, and switching circuitry. A four by four representation of the chip is shown in Figure 1. Across the top is a row of cells which perform a sample-and-hold operation at the line rate, and select columns of the resistive grid for analog signal input and output. On either side are columns of cells which select the row of the grid which is written to and read from. In the middle is a 2D array of cells which perform a second stage sample-and-hold, smooth the input from this sample-and-hold by way of resistive interconnections, and output a current proportional to the smoothed image.

Input sample-and-hold strategy

A two stage sample-and-hold design is the key to the video rate performance of the chip. In the first stage, the video signal is switched onto one of 46 capacitors through pass transistors. At one microsecond per pixel, 46 of the approximately 52 microseconds of video is utilized per line of the image.(With appropriate video switching, all 52 us could be utilized. The one us per pixel sample time was convenient given the crystals we had at hand.) Followers fed by these capacitors pass the signal to a second sample and hold circuit within each cell of the two-dimensional array during the horizontal blanking interval.

This strategy is essential to the task at hand, since sampling must proceed at high rates (~1 us), but holding must last for at least one video field (~16 ms). This is a difference of four orders of magnitude; the two stage design has circuits that operate independently at each of the two extremes of this range. The settling time of a circuit comparable to our second stage is on the order of 10 us. The horizontal blanking interval lasts for about 9 us in NTSC video; this system makes use of this 'dead time' to charge up the second stage of sample-and-hold circuits inside the two dimensional array.

Since a standard video signal has a resolution of roughly 500 by 500 pixels, split into two fields of roughly 250 lines each, this system samples a video image at a very low resolution. At 1 us per pixel horizontal sampling time, about ten video pixels are averaged (integrated in time by the first stage sample-and-hold circuit) for

input to one resistive grid cell. Similarly, with a 47 row resistive grid, several lines must be averaged to form a single row of resistive grid input. In particular, five rows per field are fed into a given row of second stage sample-and-holds for each field; at two fields per frame, 10 lines are averaged (integrated in time by the second stage sample-and-hold circuit) to form the input to a single line of the resistive grid. Thus each pixel input to the grid for smoothing represents an average of 10 by 10 video pixels.

Resistive net

The resistive net configuration is identical to that of the analog retina [1]. Transverse interconnections are comprised of horizontal resistors. The horizontal resistor circuit utilizes transistors biased by a modified transconductance amplifier; the bias compensates for the body effect so that the resistance is the same at different large-signal levels. As in the analog retina, a single global bias control is routed off-chip so that the transverse resistance can be changed equally at all nodes at once. Also, as in the analog retina, the input resistors consist of transconductance amplifiers configured as followers; a global bias line allows one to vary the input resistance at all nodes simultaneously.

Output circuitry

Signals are read out from the resistive grid as currents. They are sensed by an off-chip high gain current sense amplifier built from a conventional op amp. At any given moment, a single row of the resistive net puts current onto the set of columnar output lines, and a single column is selected to pass its current off-chip through pass transistors. The pixel being charged up in the first stage sample-and-hold at the moment corresponds to the selected column, so that a given pixel is written to and read from simultaneously. (One consequence of this design is that there is some capacitive coupling of the input signal onto the output lines; a simple off-chip compensation circuit serves to cancel this unwanted component from the output.)

System performance

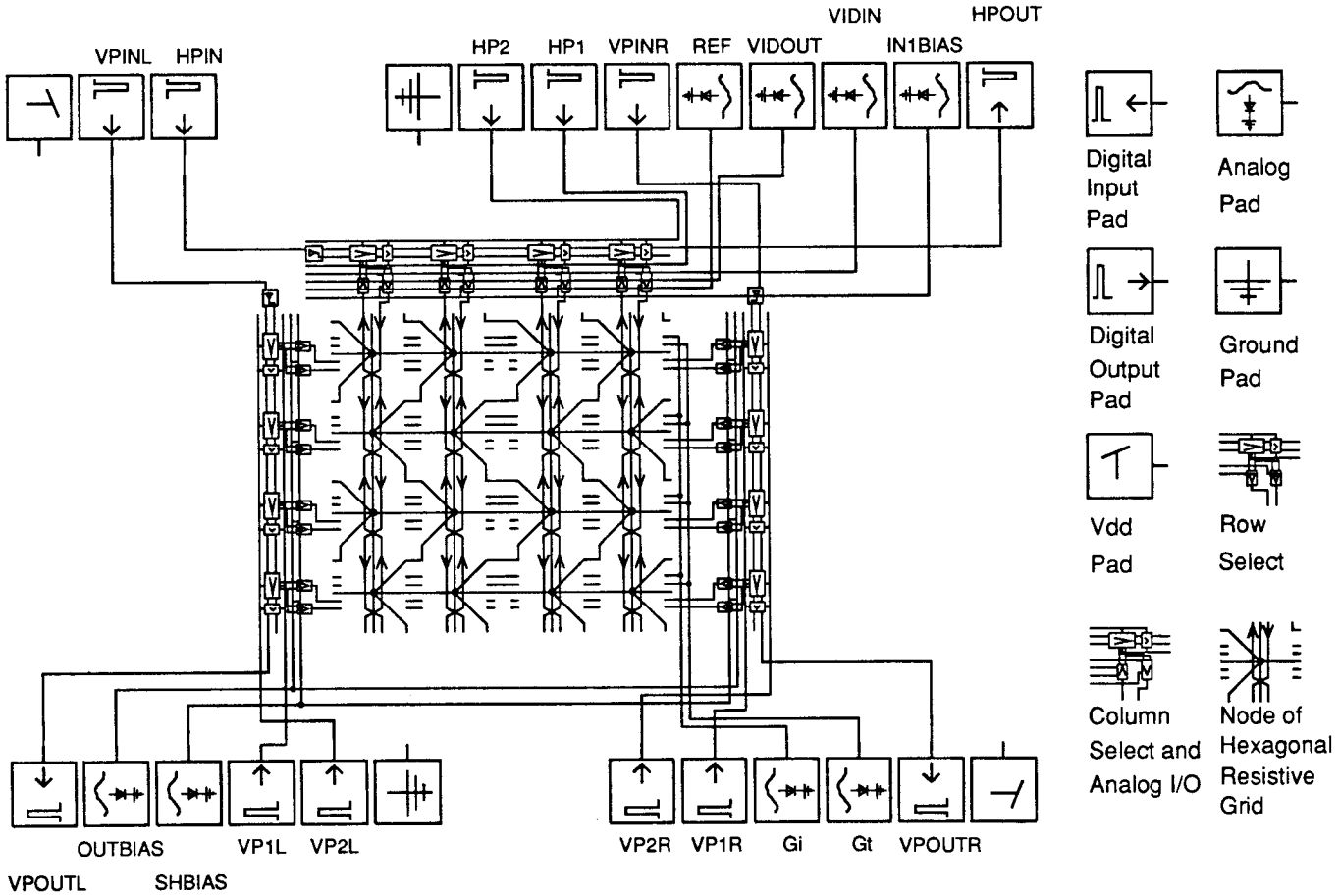

Fig. 1. A 4 by 4 node representation of the 46 by 47 node resistive network chip. Analog pads are used for amplifier biases, for video input and output, and for a reference node used by an external current sense amplifier. Digital input pads are used for video switching signals. Row selection circuitry was built separate for the two halves of the array (columns on either side of the chip); in practice, both columns are driven with identical switching signals.

Horizontal and vertical smoothing are depicted in Figure 2. The space constant of smoothing is proportional to the square root of ratio of the transverse and input conductances, for a one dimensional resistive grid [1]. This relationship also holds in the approximation for a two dimensional grid. For the data presented in Figure 2, the input conductance is fixed and the transverse conductance is varied.

Conclusion

We have described a system capable of image smoothing at video rates. The degree of smoothing is continuously variable across a wide range.

This development is significant since previous analog smoothing techniques have relied upon on-chip photoreceptors (which are noisy and have limited signal range) or upon an external digital system followed by an digital-to-analog converter feeding a single on-chip sample-and-hold stage (which is slow). Our system uses as input a conventional video signal, and so builds upon the man-years of engineering that have gone into building reliable video cameras. Thus, analog VLSI has progressed to the point at which it can take advantage of the highly developed infrastructure that comprises commercial video systems.

Acknowledgements

We are grateful for input from many of our colleagues at Caltech, in particular: M. Mahowald, C. Mead and M. Sivilotti, the inventors of the original silicon retina; J. Harris, J. Luo, and C. Koch, who shared their expertise with resistive grids; D. Lyon (of Apple Computer), M. Mahowald, and S. Ryckebush, for discussions on sample-and-hold design; Hewlett Packard for computing support; and DARPA for support for MOSIS fabrication services. A.M. was supported by fellowships from the Parsons Foundation and the Pew Charitable Trust and by research assistantships from Office of Naval Research, the Joint Tactical Fusion Program and the Center for Research in Parallel Computation. Most of this work was done in Carver Mead's lab - we are grateful for his generous support.

References

[1] Mead, C.A., Analog VLSI and Neural Systems, Reading: Addison-Wesley, 1989.

[2] Luo, J., Koch, C. and Mead, C, "An experimental subthreshold, analog CMOS two dimensional surface interpolation chip," presented at the Neural Information Processing Systems Conference, Denver, November, 1988.

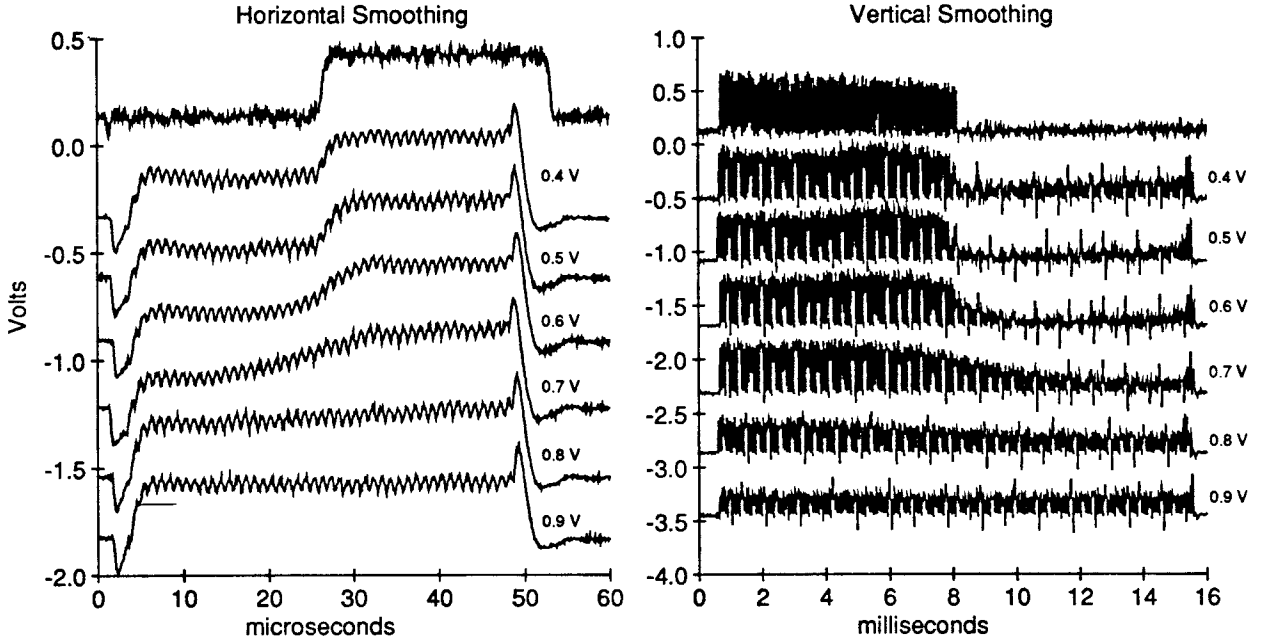

Fig. 2. Chip input (top two traces) and output (bottom traces) at different degrees of smoothing. At left, smoothing is shown at the line rate. Each trace is a horizontal line of the chip output. Bumps of 1 us duration in the output traces result from current switching transients inherent in the current-sense output scheme. At right, smoothing is shown at the field rate.

The bias of the input transconductance amplifier was held at 0.55 Volts while the bias of the transverse resistor circuit was varied. The transverse resistor bias is shown to the right of each trace. The output replicates the input at low transverse conductance values, spatially averages the input at high values, and smooths the input at values in between.