Proceedings of the Second Government Neural Network Applications Workshop,
Huntsville, Alabama, September 1-12, 1991

Image Range Control and Contrast Enhancement via VLSI Masking

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

May 21, 1991

Second Government Neural Network Applications Workshop
Huntsville, Alabama, 10-12 September, 1991
Presentation clearance : **UNCLASSIFIED**
Relevant topics: **Machine perception, satellite surveillance, hardware implementations, sensor fusion**

1 Introduction: the problem with excellent sensors

Various image sensors have been developed for government applications with excellent dynamic range. The maximum level that can be sensed is more than 3 orders of magnitude greater than the minimum (noise) level in CCDs, for example [1]. Video monitors, on the other hand, can only display intensities covering about 1.5 orders of magnitude. The problem with displaying an excellent sensor on a video monitor is that the *incremental contrast* is very small. Detail in the original scene is displayed at low contrast.

This problem is especially obvious in scenes in which one area is much brighter than the rest of the scene. Flames, for example, will overload the monitor range unless the sensor is "turned down" by reducing the iris aperture or shortening the pixel integration time. With the sensor turned down, though, it is hard to see more than a silhouette of what is burning; important detail such as numbers on a vehicle, for example, are lost in darkness.

In the following, the range control problem is approached with a real-time system that performs a special kind of masking, analogous to dodging in film development, with the added benefit of contrast enhancement for video images, automatically at video rates. Masking is performed with a subthreshold CMOS analog VLSI chip at a resolution of 50 by 50.

2 Masking for intelligent range compression

2.1 The need for range compression

The cones of the retina operate over four orders of magnitude of light intensity. Yet the retinal ganglion cell firing rates vary over only 1.5-2 orders of magnitude [2]. The retina must intelligently compress the input range to match the output range. The range of a typical photographic negative is two or more orders of magnitude, while the range of photographic print paper is 1.5 orders of magnitude. (The range of newsprint is even worse - near 1 order of magnitude.) [3] Similarly, the range of a typical CCD exceeds the range of a typical TV monitor. Our artificial imaging systems face a range compression problem analogous to that of the retina.

In the biological system, the solution is to subtract from the input a smoothed version of itself; the horizontal cells in the outer plexiform layer and/or lateral plexes of amacrine cells in the inner plexiform layer serve as the smoothing substrate. The subtraction pulls down the highest inputs and brings up the lowest inputs, so that the resulting image fits within the output range. Since it is a smoothed version of the input that is subtracted, the input detail is preserved. In photography, a similar technique called *masking* serves the same purpose. The original negative is blurred optically to form a positive transparency. When making a print, the positive is put in front of the negative.

For still electronic images, a mask is formed via convolutions or Fourier techniques [4]. Image smoothing at video rates with compact hardware was first achieved in the silicon retina [5-7] with a resistive grid; on-chip photoreceptors form the image. We use a resistive grid also, but use a video signal for input [8]; the smoothed result is subtracted from the sensor in real time.

2.2 The need for intelligent masking

Consider the result of masking in an image with a sharp light to dark transition (Plate 1 a,b). Subtraction of the mask pulls the darks up and the lights down. Near the transition, the mask varies smoothly from light to dark. Upon subtraction, overshoots result on either side of the transition (Plate lc). In psychophysics texts, overshoots are presented as evidence of lateral inhibition in the visual system, and are called *Mach bands.* [9-11] The problem with masking is not overshoots at small transitions, where they sharpen the result (i.e. improve incremental contrast), but at large transitions, where the overshoots may exceed the range of the output. If the visual system "hit the rails" on either side of a transition from sunlight to shadow, as in Plate lc, we would not have survived as well as we have. We would be blind, for example, to a predator sitting near the border of the shadow, and would be eaten!

In photography, such overshoots are avoided by using a masking technique called *dodging.* In the darkroom, an opaque card is moved quickly over the dark regions, to prevent overexposure on the print. If the card is inadvertently moved into the light region, an overshoot results. Thus dodging is a painstaking process and is available for custom work only. In the image processing community, some workers have realized that temperance is called for in electronic masking [12], but techniques analogous to dodging do not exist. Recently, we have found a way to operate our system to take care of the problem of overshoots at large transitions. We take advantage of the fact that the resistors used for smoothing saturate [13]. For small transitions, smoothing occurs normally, but for large steps, corresponding to illumination differences, the electronic mask is step-wise smooth or *quasi-sharp* (Plate ld). (Saturating resistors should not be confused with switching schemes such as resistive fuses [14,15], which require extra processing for annealing and resetting [16].) This approach, then, sharpens detail and masks large steps for range control without introducing overshoots that exceed the range. We believe that the nervous system may utilize a similar strategy.

These results apply to color as well as black-and-white processing (Plate 2). For government applications, the various color channels may be different sensors fused into one image. (In commercial application, this type of processing is known as color constancy [8].) At the workshop, we will show a videotape of both kinds of processing, directly compared to the unprocessed sensor image.

3 References

1. **Seige, P. and Ress, G.**, Application of Texas Instruments TC-104 linear charge-coupled-device arrays in spaceborne camera systems, Optical Eng. 26, 1987.
2. **Creutzfeldt, O., Lange-Malecki, B., and Wortmann, K.**, Darkness induction, retinex, and cooperative mechanisms in vision, *Exp. Brain Res.* 67, 1987.
3. **McCann, J.** *personal communication.*
4. **Oppenheim, A.V., Shafer, R.W., and Stockham, T.G., Jr.**, Non-linear filtering of multiplied and convolved signals. *Proc. IEEE* 56, 1968.
5. **Sivilotti, M.A., Mahowald, M.A., and Mead, C.A.**, Real-time visual computation using analog CMOS processing arrays, in *1987 Stanford Conference on VLSI*, Cambridge, MA: MIT Press, 1987
6. **Mahowald, M.A., and Mead, C.A.**, Silicon retina, in *Analog VLSI and Neural Systems*, C.A. Mead, Ed., Reading, MA: Addison-Wesley, 1989.
7. **Mahowald, M.A., and Mead, C.A.**, The Silicon Retina. *Sci. Am.* 264, May, 1991.
8. **Moore, A., Allman, J., and Goodman, R.**, A Real-time Neural System for Color Constancy. *IEEE Trans. Neural Networks* 2, 1991.
9. **Fiorentini, A.**, Mach band phenomena, in *Handbook of Sensory Physiology: Visual Psychophysics*, D. Jameson and L. Hurvich, Eds., New York: Springer-Verlag, 1972.

10. **Ratliff, F.**, *Mach Bands: Quantitative Studies on Neural Networks in the Retina*, San Francisco, Holden-Day, 1965.

11. **Bekesy, G. von**, Mach- and Hering-type lateral inhibition in vision, *Vision Res.*, 8, 1968.

12. **Schreiber, W.F.**, Image processing for quality improvement. *Proc. IEEE* 66, 1978.

13. **Mead, C.A.**, Aggregating signals, in *Analog VLSI and Neural Systems*, C.A. Mead, Ed., Reading, MA: Addison-Wesley, 1989.

14. **Harris, J., Koch, C. and Luo, J.**, A two-dimensional analog VLSI circuit for detecting discontinuities in early vision, *Science* 248, 8 June 1990.

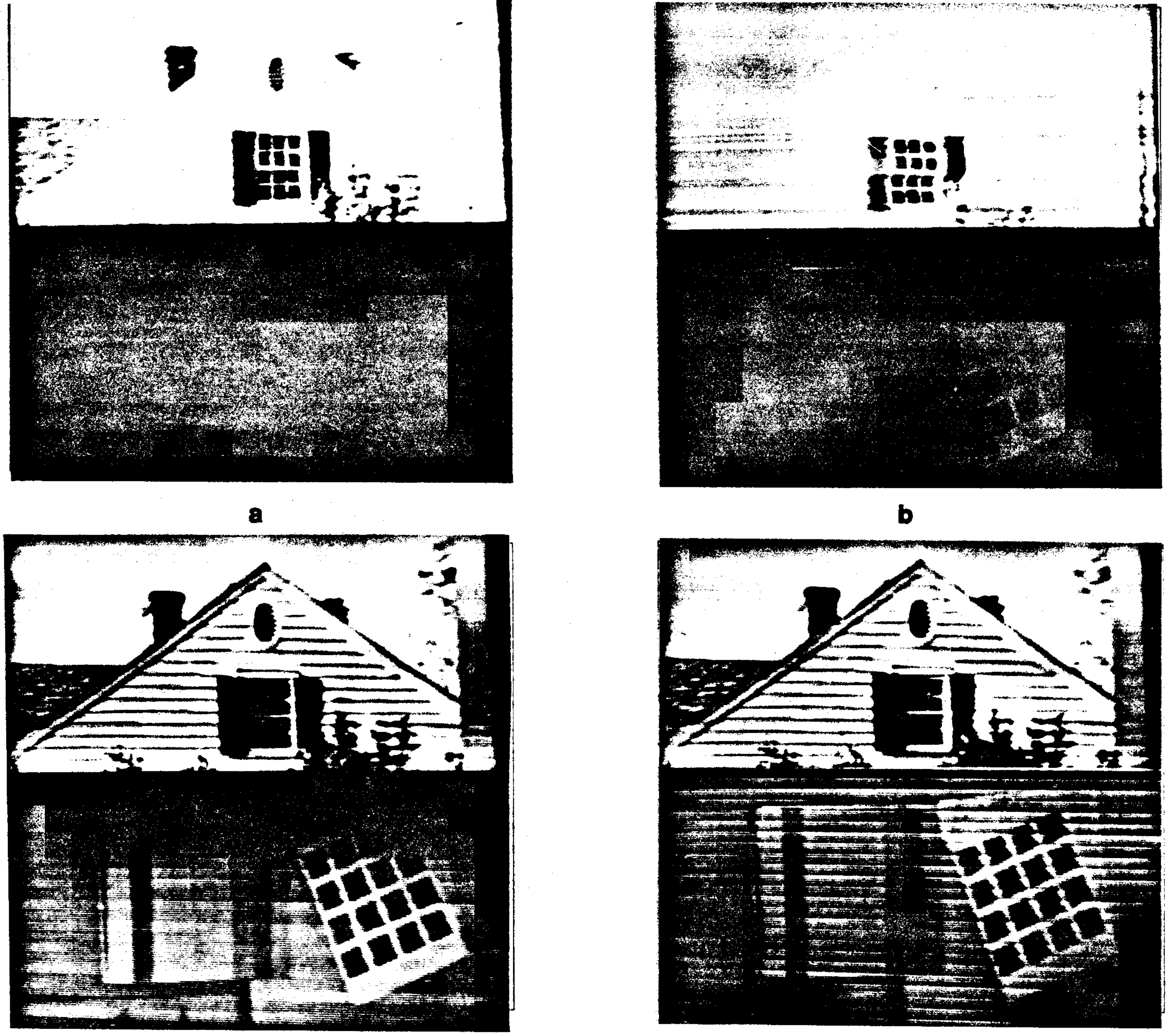
15. **Harris, J., Koch, C., Luo, J. and Wyatt, J.**, Resistive fuses: analog hardware for detecting discontinuities in early vision, in *Analog VLSI Implementations of Neural Systems*, C. Mead and M. Ismail, Eds., Norwell, MA: Kluwer, 1989.

16. **Harris, J.G., Continuous-time segmentation networks**, in *SPIE Technical Symposia on Optical Engineering and Photonics in Aerospace Sensing*, Orlando, FL. *Session on Visual Information Processing: From Neurons to Chips*, Mathur, B., Ed. *In press.*

4 Plate legends

Plate 1. CCD image of a challenging scene: a house outdoors, lit with sunlight, and an office wall lit indirectly. With the camera iris closed down (a), the house is visible outside, but the indoor objects are hard to see. With the camera iris wide open (b) the objects indoors are visible, but the house outdoors is awash in white. Linear masking (c) markedly improves visibility, but also introduces unacceptable distortions at the light-dark border. With nonlinear masking (d), analogous to dodging in film development, the scene is reproduced on a video monitor much as the eye sees it. Range is compressed well and contrast is enhanced with no distortion.

Plate 2. Image of a subject under colored light, before and after masking on the red channel of a color CCD camera. The camera was color-balanced under white (fluorescent) light. When a reddish, incandescent light is turned on, the skin tone is distorted (a). With masking on the red channel (b), skin tone is more normal.



a

b

c

Plate 1

d

a

Plate 2

b