TP 2.1: A Foveated Visual Tracking Chip

Ralph Etienne-Cummings^{1, 2}, Jan Van der Spiegel^{1, 3}, Paul Mueller¹, Mao-zhu Zhang¹

¹Corticon Inc., Philadelphia, PA

²Department of Electrical Engineering, Southern Illinois University, Carbondale, IL ³Moore School of Electrical Engineering, University of PA, Philadelphia, PA

This chip uses both biological motivation and compact focal-plane processing to solve a real commercial problem such as 2D camera pointing [1]. This system is unique in its use of a spatially variant layout and arbitration between smooth pursuit (in fovea) and acquisition (in periphery) to improve the tracking process (similar to primate visual tracking). Furthermore, with the exception of the motor drivers, little additional circuitry is required for successful visual tracking. A micrograph of the chip is shown in Figure 1, highlighting the two functional regions.

At the center of the chip is the foveal region where the photosensitive elements and edge-detection circuits are densely packed. To improve spatial resolution of the fovea, motion-detection circuits are moved away from the fovea and placed at the bottom of the sensing area. The fovea computes divergent 2D velocity. A global vector for the entire fovea is obtained. This vector provides signals for fine motor control to maintain fixation on a moving target.

Surrounding the fovea is the peripheral region, where the centroid of a moving target is computed. In the periphery, highresolution imaging is not required. Hence, all the circuits required to realize the centroid computation are placed within each cell. Moreover, the size of the photo-receptors are also increased so that fewer cells can be used to cover a large area. Each cell must, however, be smaller than the fovea so they can be centered within the fovea. The function of the periphery is to locate new interesting targets and to re-capture targets that escape from the fovea. Hence, an arbitration system decides whether tracking or acquisition should be performed. The signals from the periphery are used for coarse motor control to acquire a moving target.

All computations are performed at the focal plane using parallel, continuous-time mixed analog/digital circuits. The SIMD architecture used in this chip is ideal for compact, real-time focal-plane image processing.

The fovea is composed of a 9 x 9 array of pixels. The photoreceptors are implemented with phototransistors. The inter-receptor area is filled with logarithmic-compression and edge-detection circuits. The logarithmic circuit, implemented using a pair of pMOSFETs operating mostly in sub-threshold, responds to 5-6 orders of magnitude of light intensity. A smoothed version of the compressed image is realized using a passive resistive grid. Edge detection is subsequently obtained by computing the difference between the smoothed and unsmoothed image (a pseudo-discrete Laplacian edge detector).

Most successful motion detection chips are based on local correlation [1,2]. Combining these schemes with time-of-travel measurements produces compact, VLSI-friendly motion algorithms, similar to the approach used in the fovea. A discussion of the technique is found in Reference 3. The algorithm requires a binary image of the edges in the scene. This contrast normalization step, realized with a hysteretic comparator, is limited by the fixed-pattern and temporal noise of the photoreceptor circuits to a minimum contrast of about 10-20%. The target speed is given by the time between disappearance of an edge at a pixel and its reappearance at a neighboring pixel. The direction is signaled by which neighboring pixel receives the edge. In this application, the speed measurement is discarded by OR-gating the direction of motion for the entire fovea. This is sufficient for target tracking and resembles a bang-bang controller, with step sizes calibrated to the angular displacement of the target. The fovea tracks diverging targets moving with 2D velocities ranging from 0.4 to 8500pixels/s. A schematic of the fovea is shown in Figure 2.

The periphery contains similar photoreceptor, edge detection and contrast normalization circuits as the fovea. The phototransistors are enlarged so the field of view covered by each peripheral cell can easily be centered into the fovea, yet a small number of these cells are required to cover a large visual angle. The inter-receptor area is used for centroid computation.

In the periphery, it is not important to measure the velocity of the moving target. Only the location of the target is required. The target must, however, be extracted from the background for the centroid to be correct. Hence, the peripheral cells perform a temporal derivative of the binary edge image, and label the location of arriving edges (+d/dt or ON-set).

When an edge appears at a pixel, it broadcasts its location to the edge of the array by activating a row and column line. This row (column) signal sets a latch at the right (top) of the array. This circuit performs an asynchronous first-come-first-serve function by preventing any other latch from being set. Multiple latches are set only when multiple pixels are activated simultaneously. The location of the triggered row (column) is then given by an analog value read from a resistive divider. When multiple rows (columns) are activated, the centroid is obtained since the various voltages for the rows (columns) are shorted to the output line through identical resistors (CMOS switches). In addition to the analog centroid value, the chip generates a request signal that indicates the current centroid is valid. The chip continues to output this value until it receives an acknowledgment. Acknowledgment tells the chip that its request has been seen and readies the chip for the next centroid computation by resetting latches. Figure 3 shows peripheral circuits.

The chip features and the tracking performance of the combined chip-motor system are summarized in Table 1. The motion perception in the fovea spans a range of over four orders of magnitude in target velocities, and is mainly limited by the settling time of the photosensors, which slows down at lower incident light levels. Still, the chip operates over approximately 5-6 orders of magnitude of ambient light intensity. System performance is limited by the contrast sensitivity of the edge detection circuit. However, smooth tracking in the fovea is observed in contrast as low as 10% and acquisition in the periphery operates even in poorer image contrast conditions.

In a camera pointing application, a small commercial video camera is attached to the tracking unit. This system is able to track walking or running persons under indoor or outdoor lighting conditions at small or large distances. It also has been used to track moving automobiles. Tracking distance range is dependent on the imaging optics. Furthermore, it has been used in linefollowing experiments for autonavigation. Evidently, this chip can be used for many visual tracking tasks in real world environments.

Acknowledgment

The work was supported by Corticon. J. Van der Speigel and P. Mueller have a financial interest in Corticon, Inc..

References:

[1] Koch, C., H. Li (Eds.), Vision Chips: Implementing Vision Algorithms with Analog VLSI Circuits, IEEE Computer Press, 1995.

[2] Etienne-Cummings, R., J. Van der Spiegel, "Neuromorphic Vision Sensors," Sensors and Actuators: A, Vol. SNA056, pp. 19-29, 1996.

[3] Etienne-Cummings, R., J. Van der Spiegel, P. Mueller, "A Visual Smooth Pursuit Tracking Chip," Advances in Neural Information Processing Systems 8, D. Touretzky, M. Mozer and M. Jordan (Eds.), pp. 706-712, 1996.



2-1-1: Micrograph of tracking chip.



2-1-2: Schematic of the fovea.



2-1-3: Schematic of the periphery.

Technology:	2µm n-well, double-poly CMOS			
Size:	6.4x6.8mm ²			
Package:	132 pin DIP			
Array size :	Fovea: 9x9 at 110µm pitch			
Ū	Periphery: 19x17 at 300µm pitch			
Fill factor:	Fovea: 8%			
	Periphery: 48%			
Transistors:	Fovea: 3240			
	Periphery: 679	8		
Ambient [mW/cm ²]		2.5	25	250
Fovea: max speed [pixel/s]		645	2627	8445
Fovea: min speed [pixel/s]		0.36	0.36	0.36
Peri: max. temporal freq. [kHz]		62.5	250	800
Peri: min. temporal freq. [kHz]		(No lower bound)		

2-1-Table 1: Chip characteristics and performance.