

TP 2.2: An Analog VLSI Chip for Estimating the Focus of Expansion

Ignacio S. McQuirk^{1,2}, Hae-Seung Lee¹, Berthold K. P. Horn¹

¹Massachusetts Institute of Technology, Cambridge, MA

²Now with Maxim Integrated Products, Sunnyvale, CA

Attention has recently been given to the design of custom analog VLSI chips for early vision-processing problems [1]. The key features of these tasks are simple operations performed in parallel at each pixel in an image, typically resulting in a description of the scene useful for higher-level vision. This type of processing is well suited to implementation in analog VLSI, yielding compact high-speed low-power solutions. This chip computes the focus of expansion (FOE). As shown in Figure 1, the FOE is the image point that the camera is moving toward. Image features appear to expand outward from the FOE and knowledge of its location provides the direction of camera motion.

Given a pin-hole model for image formation, an equation at each pixel (x, y) can be found that relates the variation in the observed image brightness E to the location of the FOE (x_0, y_0) and the depth of the scene [2]. Special image points, where brightness is instantaneously constant ($\partial E / \partial t \equiv E_t = 0$), provide important constraints on the location of the FOE. At these points, the spatial brightness gradient $\nabla E \equiv (E_x, E_y)$ is perpendicular to the vector to the FOE. With a sufficient number of such points in the image I , the location of the FOE can be estimated by combining their constraints in a least-squares fashion. The resulting location of the FOE is the solution of the linear system shown in Figure 2, where a simple cutoff weighting function W is used to select those points where $E_t \approx 0$. The formulation of this brightness-gradient algorithm makes it possible to calculate the FOE using only nearest neighbor information.

The FOE chip calculates the left-hand side of the equation of Figure 2 and by closing a proportional feedback loop about the chip, the solution of the linear system is found. Due to the complexity of the computation, a row-parallel processing scheme is chosen over a pixel-parallel approach. The system architecture of the FOE chip is comprised of four sections: an interline CCD imager, an array of floating-gate amplifiers, a column of analog signal processors, and a position encoder (Figure 3).

The front-illuminated CCD array acquires the images needed to estimate the brightness gradients. The interline registers have two storage areas per pixel to allow two images to be acquired successively in time and stored in the registers in an interleaved fashion. Once acquired, the image-pair is shifted right columnwise and the array of floating gate amplifiers linearly transduce this charge signal into voltages for the analog signal processors.

The processors also require the estimate of the location of the FOE, driven in from off-chip, and the present (x, y) position of the data, provided in voltage by the position encoder at the far right of Figure 3. The encoder uses the voltage on a resistive chain to encode the y -position up the array and a digital shift register to select the appropriate x value as each pair of columns is processed.

From the image data, the pixel position, and the FOE estimate, each processor in the array computes four outputs that are summed up the column in current and sent off-chip. The overall block diagram for the analog processors is shown in Figure 4.

Finite differencing is used to estimate the image brightness

This work was supported in part by NSF, the ARPA, and by an AT&T Bell Labs Cooperative Research Fellowship.

gradients. Four source-coupled pairs transduce eight pixel voltages into differential currents that are then combined using mirrors to form the brightness gradients E_x , E_y , and E_t . An absolute value circuit computes $|E_t|$ that is summed up in current over the column to form an output of the chip, useful in setting the cutoff of the weighting function.

A local copy of $|E_t|$ is subtracted from a reference current and injected into a latch. If $|E_t|$ exceeds the cutoff threshold, the latch turns off pass gates in-line with the brightness gradient currents preventing signal flow to the rest of the processor. If $|E_t|$ is below the threshold, the E_x and E_y currents are copied using mirroring. One copy is used by a pair of current-mode squarers to compute the squared gradient magnitude. This signal is summed up over the column and forms another output of the chip, useful in setting the gain of the feedback loop.

The FOE chip uses MOS multipliers that output a current approximating the product of a differential voltage and current. In the first layer of multipliers, another copy of E_x and E_y is used to calculate the inner product $(x - x_0)E_x + (y - y_0)E_y$ in current that is then transduced to voltage using MOS triode loads. A second layer of multipliers computes the product of this signal with a final copy of E_x and E_y . The resulting two output differential currents are summed over the column to form the final two outputs of the chip, completing the left side of the equation in Figure 2. Figure 5 shows an example of the y -channel output current for one of the processors in the array where pixel inputs are driven to correspond to E_y only. As predicted by the left side of the system equation in Figure 2, the output is linear in y_0 and quadratic in E_y for such a case.

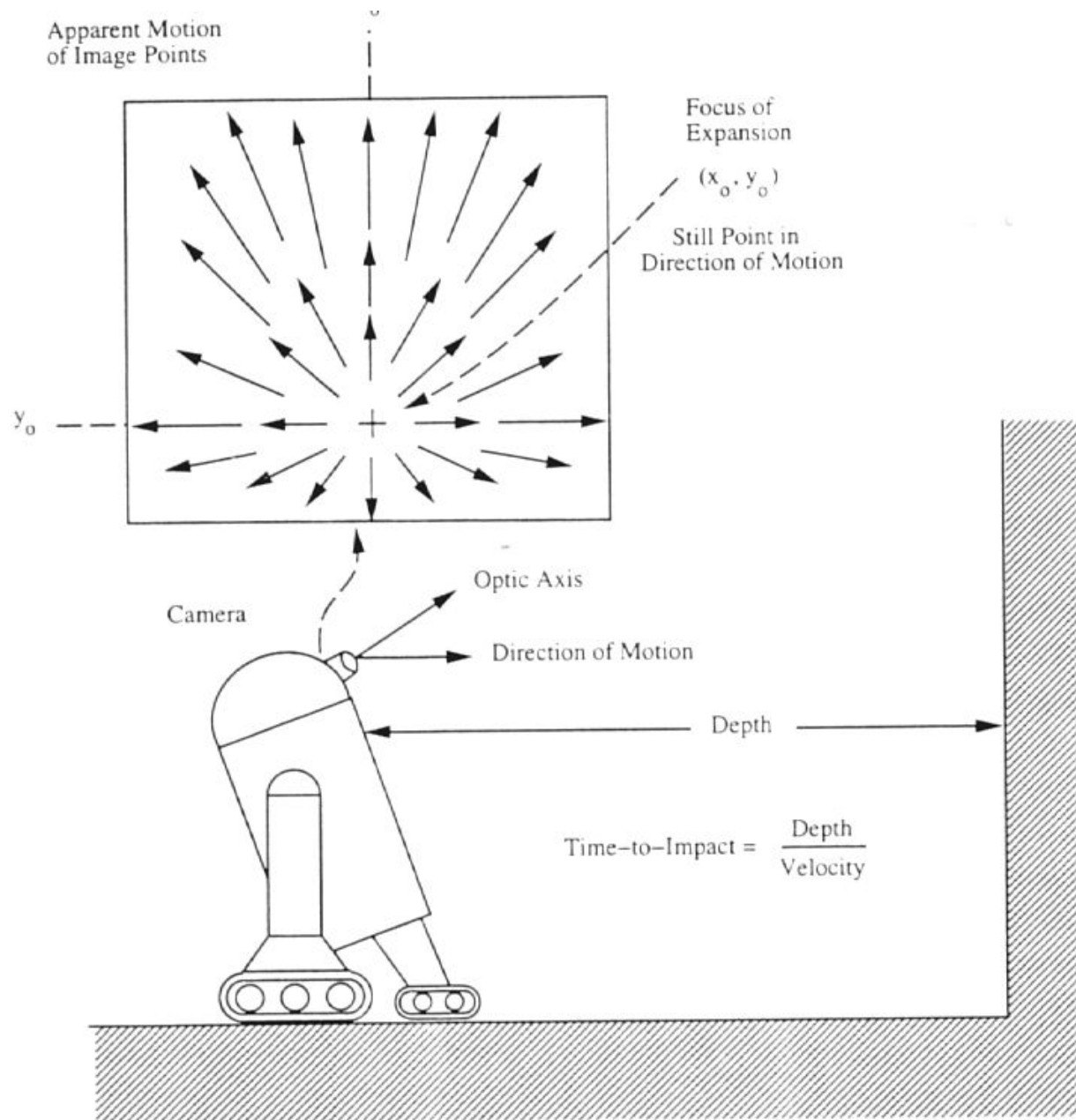
To close the feedback loop, these two currents from the column are accumulated as the image data is shifted out and once a whole frame-pair of data has been processed, the resulting residuals are used to update the FOE estimate.

The chip is tested using a flexible fiber-optic image guide to bring moving images to the stationary chip. One end of the guide is held fixed over the chip, while the other end is moved along a linear track. The viewing direction of the image guide relative to the motion is set by rotation stages, allowing precise FOE positioning.

Experiments were performed for a variety of FOE positions in the image plane. Raw image data from the chip is used to calibrate the system using a rotation-based camera calibration technique. Figure 6 compares the mean output of the feedback loop with the results of the algorithm performed in simulation using the raw image data. The maximum difference in location between the simulation and the chip is <3% full scale. With 170mW power consumption, the chip operates at up to 1000Frame/s but is limited in practice by the optical test setup to 30Frame/s. Table 1 is a summary of performance and Figure 7 is a micrograph of the FOE chip showing the system blocks.

References:

- [1] Horn, B., "Parallel networks for machine vision," *Artificial Intelligence at MIT: Expanding Frontiers*, (P. Winston, S.A. Shellard, eds.), vol. 2, ch. 43, pp. 530-573, Cambridge, MA: MIT Press, 1990.
- [2] Horn, B., E. Weldon, "Direct Methods for Recovering Motion," *International Journal of Computer Vision*, vol. 2, no. 1, pp. 51-76, 1988.

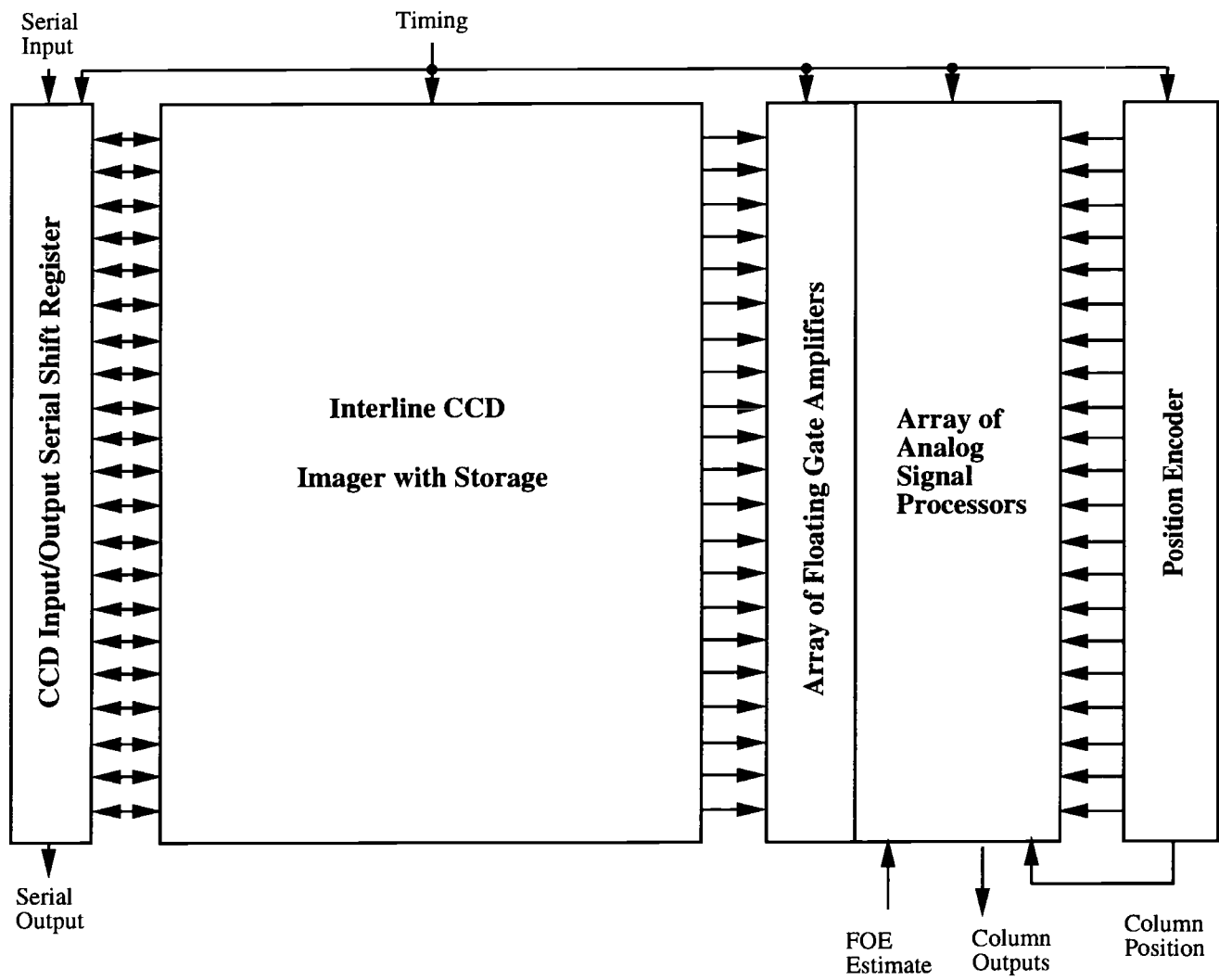


2-2-1: The focus of expansion (FOE) is imagepoint toward which camera is moving.

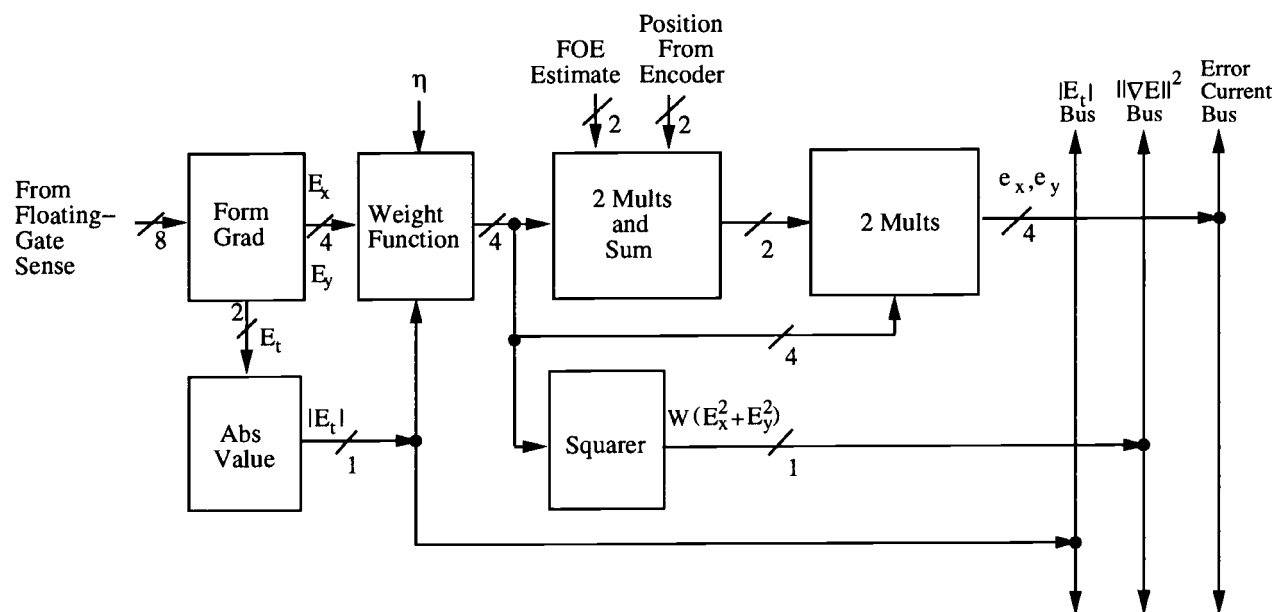
$$\sum_{(x,y) \in I} W(E_t) E_x \left((x - x_0) E_x + (y - y_0) E_y \right) = 0$$

$$\sum_{(x,y) \in I} W(E_t) E_y \left((x - x_0) E_x + (y - y_0) E_y \right) = 0$$

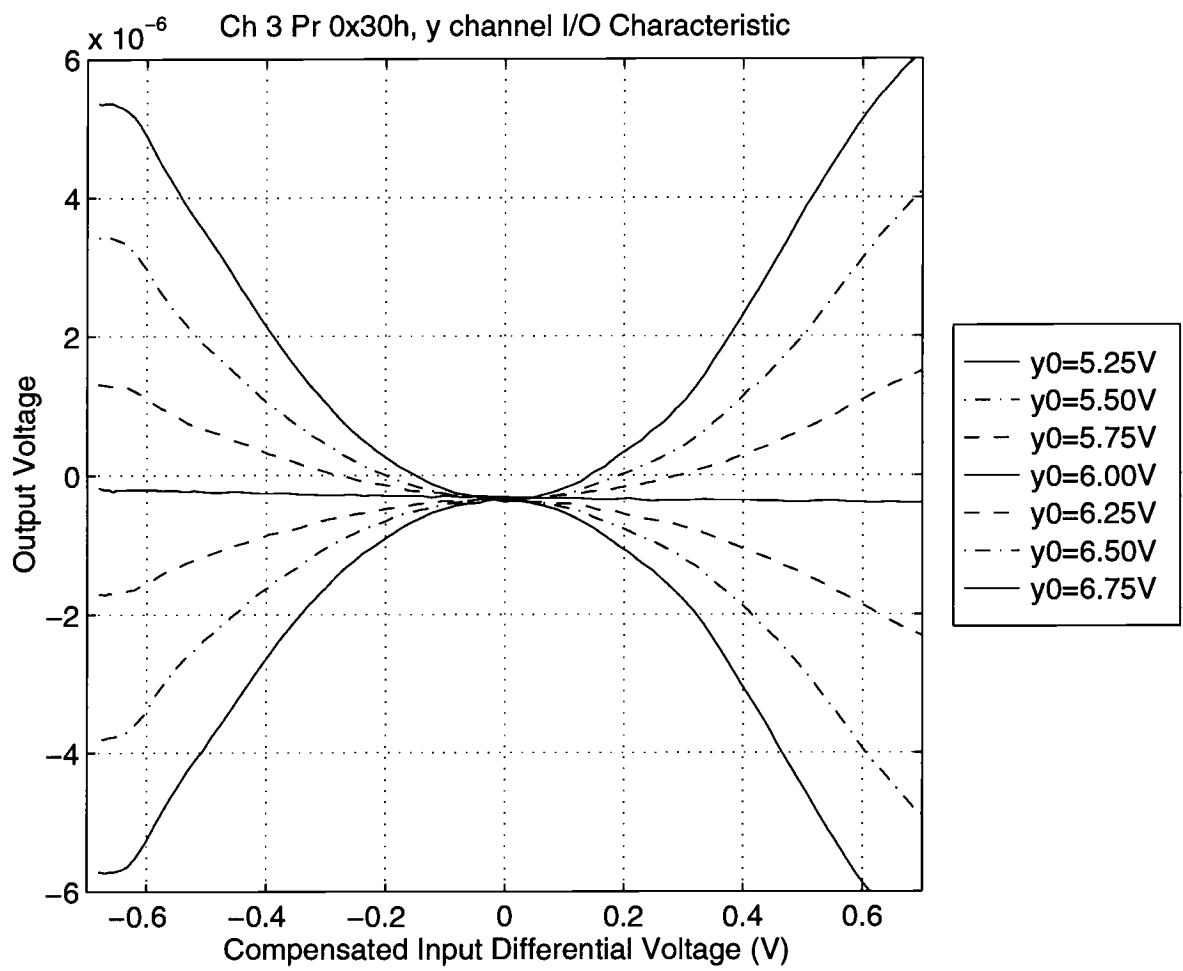
2-2-2: Linear system of equations for FOE location .



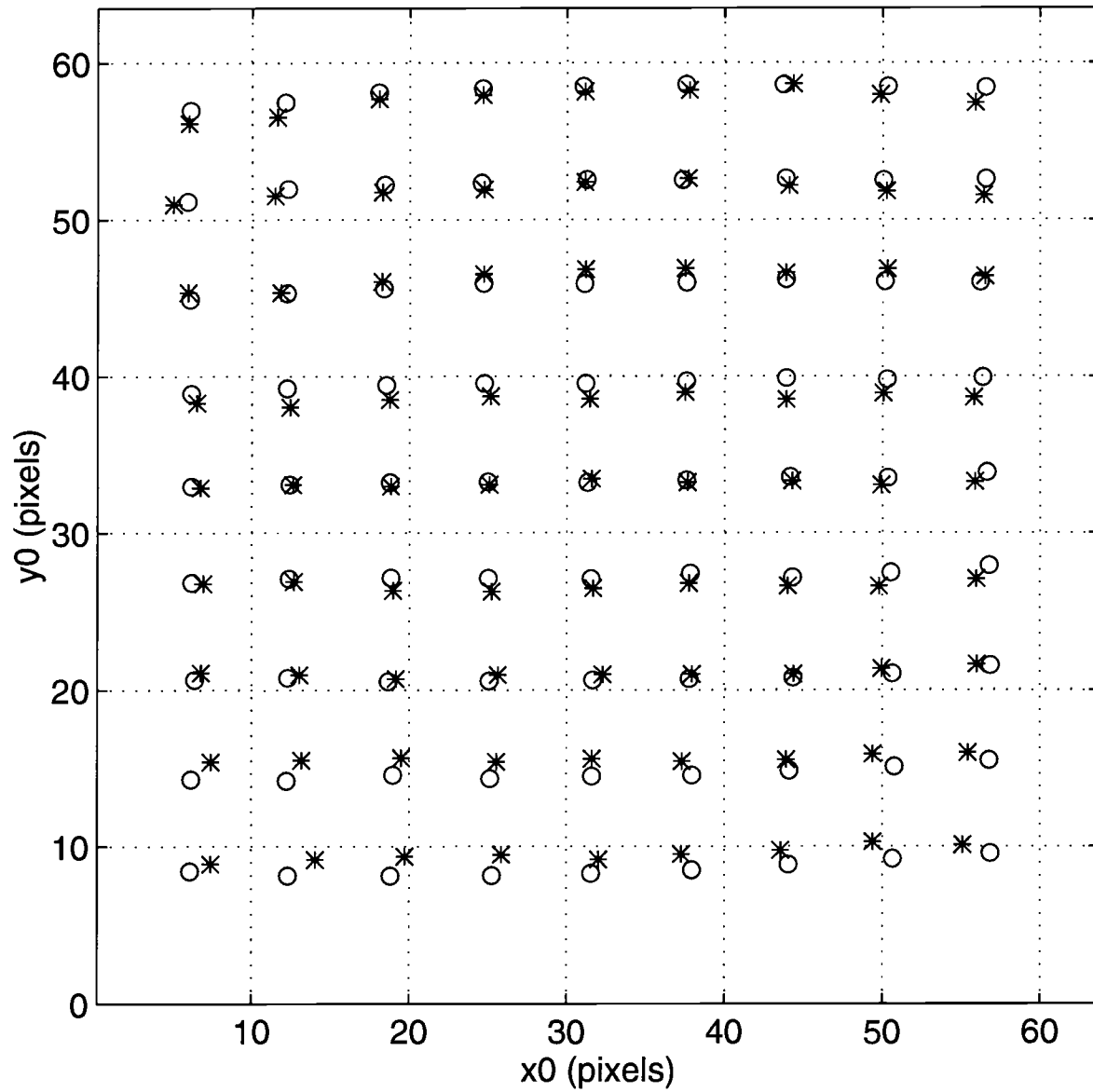
2-2-3: FOE system block diagram.



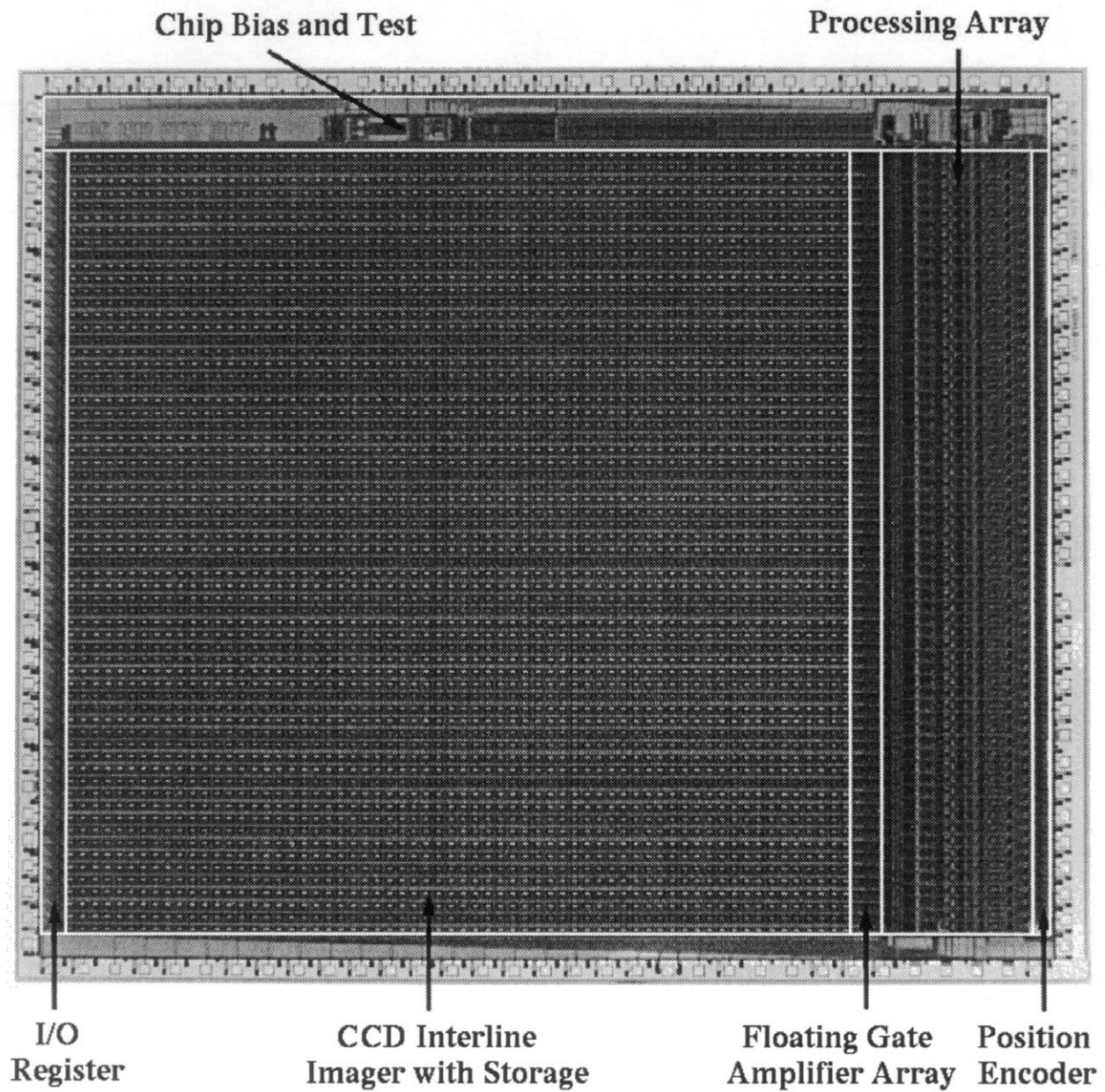
2-2-4: Analog signal processor block diagram.



2-2-5: y output channel from a processor in main array under E_y excitation only.



2-2-6: Comparison of estimating FOE using raw image data (o's) and output from FOE chip (*'s).



2-2-7: FOF chip micrograph with components indicated.

Process	10V <i>n</i> -well 2μm CCD\BiCMOS
Chip dimensions	9.2x7.9mm ²
Imager topology	64x64 interline CCD
Technology	double-poly buried channel CCD
Required clocks	17 CCD clocks, 10 CMOS clocks
Illumination	Front-face
Image sensor	CCD gate
Charge packet size	600k e ⁻
Acquisition time	tested to 1ms
Quantum eff.	30% @ 637nm
Dark current	≤10nA/cm ² @ 30°C
Transfer ineff.	≤7x10 ⁻⁵ @ 500kHz
I/O FGA output	2uV/e ⁻ sensitivity
I/O follower gain	0.97 typical
Sensor nonlin.	< 0.5% (7 b) typical
System frame rate	30Frame/s
Chip dissipation	170mW peak
System settling	20—30 iterations typical.
FOE location	≤ 3% full scale over 80% of field of view

2-2-Table 1: Summary of FOE chip and system performance parameters.