

Pixel-parallel CMOS active pixel sensor for fast object location

Ryan Burns¹, Christopher Thomas², Paul Thomas³, Richard Hornsey²

¹. Electrical & Computer Engineering, University of Waterloo, Canada

². Computer Science, York University, Toronto, Canada

³. Topaz Technology Inc., Toronto, Canada

ABSTRACT

A pixel-parallel image sensor readout technique is demonstrated for CMOS active pixel sensors to facilitate a range of applications where the high-speed detection of the presence of an object, such as a laser spot, is required. Information concerning the object's location and size is more relevant than a captured image for such applications. A sensor for which the output comprises the numbers of pixels above a global threshold in both rows and columns is demonstrated in 0.18 μm CMOS technology. The factors limiting the ultimate performance of such a system are discussed. Subsequently, techniques for enhancing information retrieval from the sensor are introduced, including centroid calculations using multiple thresholds, multi-axis readout, and run-length encoding.

Keywords: active pixel sensor, object location, centroiding, high-speed imaging

1. INTRODUCTION

Advances in CMOS fabrication technology and its use for integrated active-pixel image sensors now allow unprecedented flexibility in the design of application-specific optical sensor systems. Read-out architectures are no longer limited to raster scanning. In-pixel or column-parallel circuitry facilitate the control of all levels of sensor operation. In turn, these developments have extended the range of application of CMOS image sensors to include scenarios in which some scene analysis is performed at the sensor level.

The present paper describes the application of a custom CMOS image sensor architecture to high-speed object location in a specific, but relatively widespread, scenario. Here we define an "object" in its simplest terms to be a cluster of a pre-determined number of pixels with signals above an externally defined threshold value. Hence, while this is not a fundamental requirement of the general approach, we are currently considering cases where the object is well separated from the background. Examples of objects meeting this condition might include a high-brightness laser spot or the identification of contamination during industrial web inspection. In such cases, the rapid identification and location of the object are items of interest, rather than the imagery *per se*. Moreover, the scenario can be further generalized to include situations where there is a large relative velocity between the object and the camera. The location and arrival time of the object within the field of view of the sensor are both arbitrary and unknown. We will also assume, for the present discussion, that the illumination can be controlled such that there is an ample signal-to-noise ratio and that shortage of photons is not an issue.

Primary applications corresponding to the above scenario are those in which the object of interest is a laser spot. Optical inter-satellite links and LIDAR (light detection and ranging), for example, may require a large field of view in order to detect an incoming signal with high probability. Furthermore, the image from a scanning laser illuminator can easily have an apparent velocity across a detector array of 10^6 pixels/s. When it arrives, the signal object has a high illumination intensity but occupies a small fraction of the field of view. In such cases, the spot location is needed so that the sensor can define an area of interest for rapid readout, possibly for detecting modulation of the laser intensity. Figure 1 illustrates an additional application, a 'triangulation rangefinder' that uses the image of a laser spot to determine the range to a moving object. The imaging system need only output the 'centroid location' of the image of the target spot, which can be calibrated to provide the range. With a dithered or multi-spot laser illumination, additional information and self-calibration are possible. Here, the accuracy of the laser spot location directly affects the precision of the system while the update rate influences the motion blur and temporal resolution of the range-finding operation.

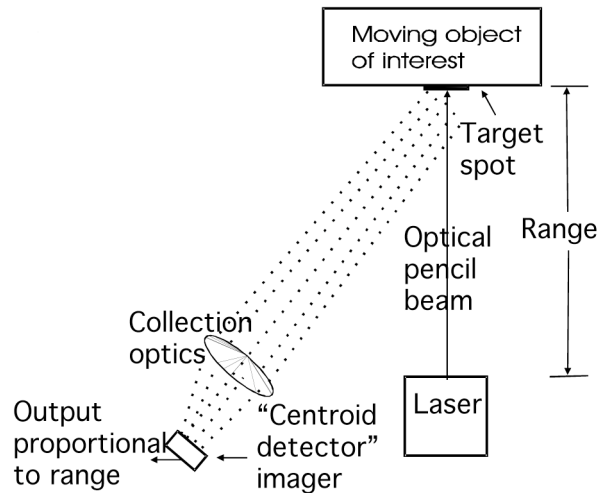


Figure 1. Implementation of a triangulation rangefinder using a custom CMOS optical sensor.

The paper is arranged as follows. Section 2 below outlines the general concept for the custom object location sensors developed in our laboratory, reviews previous related work, and discusses various implementation options. Details of the design, fabrication and operation of an analog version of the sensor will be presented in Section 3, along with the design considerations for an improved digital system. In Section 4, we present three examples of extensions to the basic system to enable accurate centroid determination, facilitate discrimination between multiple spots, and enable further data compression. Concluding remarks are found in Section 5.

2. CONCEPT

Here we describe representative approaches reported elsewhere on custom CMOS sensors for determining the position and size of an object, and present the concept behind the present work.

2a. Background

Sensors for deriving the centroids of imaged objects have been reported previously. One design features a resistive network between adjacent pixels, so that the generated analog signal can be averaged [1]. The centroid is taken to be the maximum of the voltage distribution along the row of pixels in the resistive network. A problem with this architecture is switching noise, and thus a second scheme involving pulse width addition was proposed. In-pixel summation of neighbouring pixel values is used to determine the pulse width. This architecture was capable of reporting multiple centroids for multiple objects in the field of view, with masking flags to eliminate ghost centroids. This algorithm is made possible by in-pixel centroid determination. One drawback of this approach is the necessity for large pixels.

Another design uses an aggregation network of differential transistor pairs at the edge of the array to calculate the centroid in the X and Y directions [2]. The goal is to provide only important information to an off-chip processor. A current-mode readout scheme is used in this design, in which the current is steered into one of three output directions: readout, row centroid, and column centroid. Thus, row and column centroids cannot be calculated simultaneously. This chip does not include the concept of a binary threshold for the array. Rather the centroid of the light energy is found. This result is achieved using a group of differential comparators.

Finally, a centre-of-mass tracker circuit was reported using neuron-MOS transistors [3]. The image intensity is projected onto both the X- and Y-axes. A circuit of neuron-MOS transistors is then able to perform the weighted average required for the centroid computation. A neuron-MOS transistor contains a floating gate, and multiple inputs. The inputs are combined in a weighted average, where the weighting depends on the size of the input gate. This architecture is intended to mimic the operation of a biological neuron. Essentially, this implementation is similar to [2] described above, except the actual centroid calculation is performed with neuron-MOS comparators instead of differential comparators.

Sensors have also been developed using the summation of binary image pixels along rows and columns for centroiding by [4] in a configuration similar to that proposed here. The photo-voltage is compared with a threshold and thus converted into a binary signal, which is stored on a flip-flop. The system then aims to add these bits across the array to provide the sums at the column and row edges. This is implemented with two large transconductance amplifiers in the pixel, which leads to a much reduced fill factor. The system reported exhibits half-pixel centroid accuracy. No information is provided on readout speed.

Previous work in this laboratory has implemented CMOS image sensors with binary readout and pixel-parallel operation for object location. The object is distinguished from the background by application of a simple global threshold. As illustrated in Fig.2, each pixel in the array contains a comparator that determines the logical state of the pixel's output [5, 6]. Rows and columns of the sensor are arranged in such a way as to perform effectively an OR function so that if any pixel in a particular row or column is above threshold, a latch is activated to that effect. The output of this object location system (OLS) is therefore binary. All pixels in the array operate simultaneously, leading to a straightforward and fast determination of an object's presence and dimensions. Global reset and read signals are used in the current versions, although these may also be implemented independently in different regions of the image. For a sensor with N columns and M rows, the information to read out is now of order $(N + M)$ instead of the $(N \times M)$ required for a conventional sensor; for large arrays, this saving can be significant. Simple combinational logic can be added to the row and column latches to discard regions with fewer than a threshold number of contiguous active pixels, thus reducing spatial noise and providing a spatial threshold on object size.

Implementation of in-pixel comparators, while enabling parallel operation, precludes the conventional column-based fixed pattern noise reduction techniques. Hence a compact double-sampling circuit has been implemented with-in the pixel to reduce the effects of fixed pattern noise and hence to allow an increased number of thresholds to be used [7]. As described in Section 4, several successive frames with differing threshold values can be combined to allow determination of the object's centroid to better than 0.5 pixels. There is a smooth trade-off between update rate and centroid precision.

The primary compromise associated with this technique is that detailed object information is lost. A separate "region-of-interest" analog readout of the image pixels in the neighbourhood of the centroid can recover this information. In any determination of centroid with sub-pixel accuracy, the layout of the photo-sensitive region within the pixel is important. These issues are outside the scope of the current paper.

2b. Binary Readout with Cumulative Cross Section

In order to maximize the information recovered from a single frame, the cumulative cross section (CCS) readout scheme

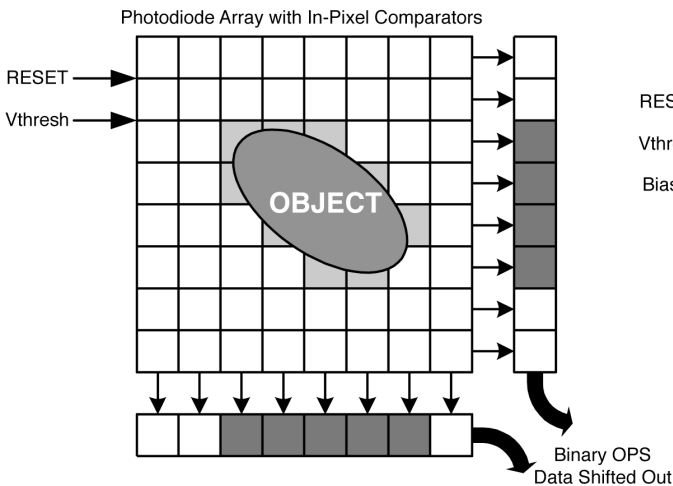


Figure 2. Basic object location architecture in which any row or column containing at least one active pixel is recorded in row and column registers.

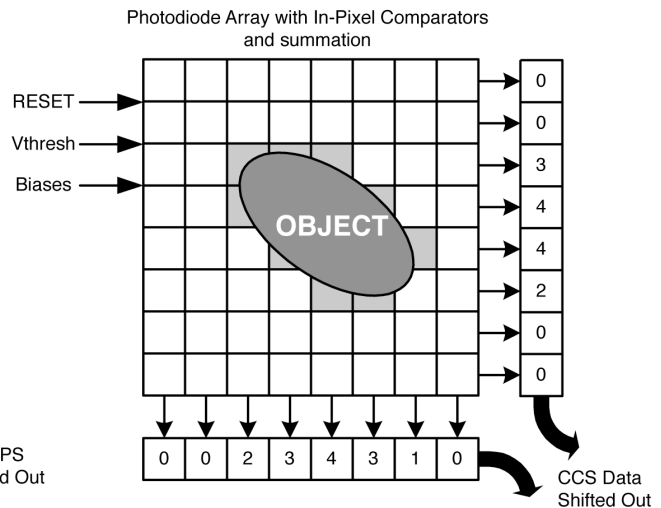


Figure 3. Cumulative cross-section readout technique, in which numbers of active pixels in each row and column are summed.

illustrated in Fig.3 is demonstrated. Instead of a simple binary output, the CCS sensor provides a sum of all pixels active in a row or column. The following section of the present paper will describe the implementation of an analog CCS sensor, along with experimental data. The design of a digital version, currently being fabricated, is also described.

3. IMPLEMENTATION

The prototype analog CCS readout sensor was implemented in standard 0.18 μ m CMOS technology with 80 x 80 pixels on a 3mm x 2mm die. Salient characteristics of this sensor are summarized in Table 1. Each pixel contained a photodiode with a PMOS reset transistor, a comparator based on a standard differential amplifier with active load, an inverter to restore full logic levels, and a current mirror for each of the row and column. For each pixel with a signal above the threshold, $\Delta I = 10\mu$ A was added to the total current in both the row and column. Final currents were converted into voltages by resistors, R_{data} , located at the edges of the array. Each resultant voltage was stored on a sample-and-hold until it could be read out via a shift register to a unity-gain amplifier for transmission off-chip. Further details of the circuit implementation are provided in [8, 9]. It should be noted that, while the present design reads out the rows and columns sequentially; a parallel readout is straightforward. Figure 4 shows a sample output at 100 frames per second from the sensor when a red LED is used as a source of illumination; the LED is moved between the two frames. The background slope on the column data in Fig. 4 might be due to leakage in the sample-and-hold circuits during readout. This figure is intended to provide a qualitative picture of operation only, and uses a simple, uncalibrated optical arrangement of the source. The expected profile of the (extended) light source appears in both the row and column outputs, demonstrating correct operation of the analog CCS system. This version of the CCS readout did not include any non-uniformity correction, such as that described in [7].

| | |
|-------------------|--------------------------------|
| Power | 8 mW (dark), 30 mW (saturated) |
| Frame rate | > 1000 frames/second |
| Die size | 3 mm x 2 mm (80 x 80 pixels) |
| Technology | 0.18 μ m CMOS |
| Pixel dimensions | 12 μ m x 12 μ m |
| Transistors/pixel | 12 |
| Fill factor | 51 % |

Table 1. Characteristics of the analog cumulative cross section readout sensor.

Figure 4. Sample output from the analog cumulative cross section sensor showing the movement of an illumination spot between two successive image frames.

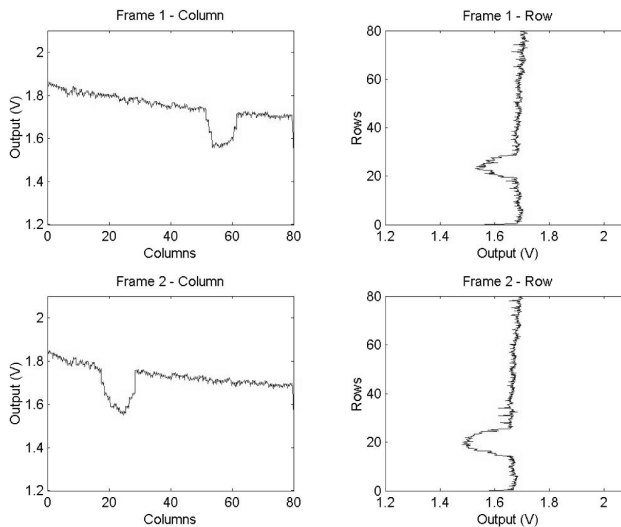


Figure 5 shows that pixel non-uniformities (pattern noise) cause pixels to switch over a range of threshold voltages for a given illumination, while Figure 6 shows the sensor power consumption as a function of frame rate for below- and above-threshold illumination. In order to compensate for the smaller integration time at higher frame rates, the illumination intensity is increased; hence, the horizontal positions of the curves in Fig.5 are shifted for different frame rates. In the absence of pattern noise, the curves in Fig. 5 would be a step function, however, the presence of fixed pattern noise (independent of signal) and a photoresponse nonuniformity (PRNU) that is signal-dependent lead to a finite slope. The similarity of slopes in Fig. 5 for the various frame rates (and hence light intensities) indicates that fixed pattern noise predominates. It is probable that this results from transistor mismatch at the input of the comparator. As expected, the fully active chip displays the higher power consumption since all current mirrors are active (Fig. 6). While the power dissipation for this prototype is too large for it to be scaled to the megapixel level, power consumption can be reduced significantly by selectively enabling the current mirrors on a per-row and per-column basis and by other design improvements.

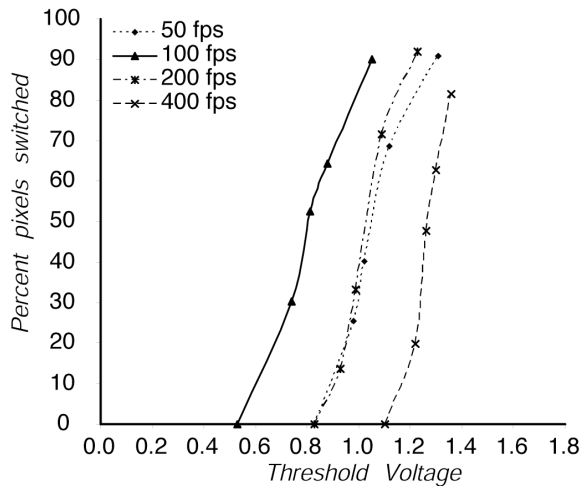


Figure 5. Switching non-uniformity of prototype CCS imaging chip. Different light intensities were used at each frame rate to provide sufficient illumination.

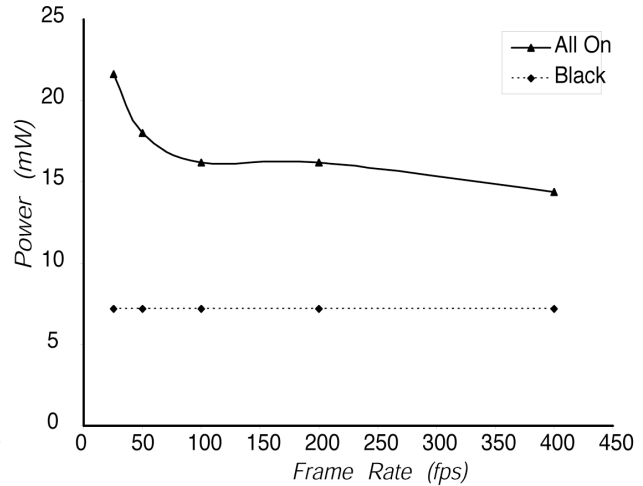


Figure 6. Power consumption of the 80 x 80 array as a function of frame rate for high (above threshold) and low (below threshold) illumination.

The present design did not provide great flexibility for selecting the value of the R_{data} resistor used to convert current quanta to a voltage output. However, it did allow the choice of high or low resistance. By changing the resistance of R_{data} , a quantum of current from an active pixel can cause a very large voltage drop, leading to the ability to operate the sensor as a binary OLS. The choice of R_{data} represents a trade-off between resolution and dynamic range, so a controllable R_{data} could be advantageous when scaling the imager resolution up to more useful sizes, such as in a megapixel imager.

Because of the difficulty of resolving current quanta over a wide range of values, a digital version of the CCS sensor has been designed. A schematic of the pixel is shown in Fig.7 and a layout in Fig.8. Each pixel includes an in-pixel pattern noise reduction circuit and two digital adders. If the pixel is above threshold, the row and column sums are incremented, so that a digital value for the number of active pixels in the row or column is available at the edge of the array. The number of bits required in the adders is determined by the maximum number of active pixels expected in the row or column, and is a trade-off with the desired fill factor of the pixel; here we chose a 5-bit adder for a 64 x 64 pixel array. Region-of-interest and an analog readout path have also been implemented. This sensor is currently being manufactured in a 0.18 μm CMOS process.

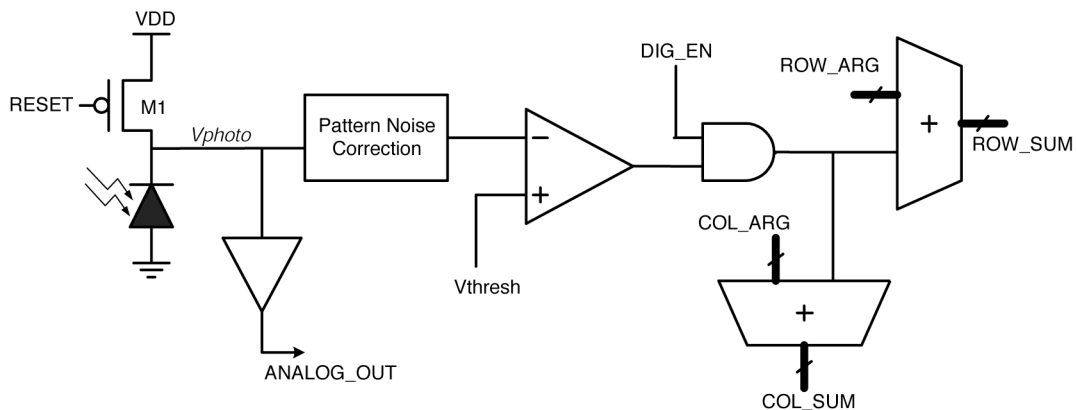


Figure 7. Pixel schematic for the digital implementation of the cumulative cross section readout technique.

Fast operation of the OLS and CCS sensors is achieved by pixel-parallel integration and non-scanned read-out. Table 2 summarizes some of the constraints on the maximum operating speed of these types of sensors. With suitable design and

layout, it is expected that speeds well in excess of 10^6 frames per second are feasible from an electro-optic standpoint. It should be noted that, to first order, the readout rate is independent of the number of pixels because of the parallel operation. In this case, the limiting factor becomes the integration time required to generate sufficient signal for reliable operation of the comparator.

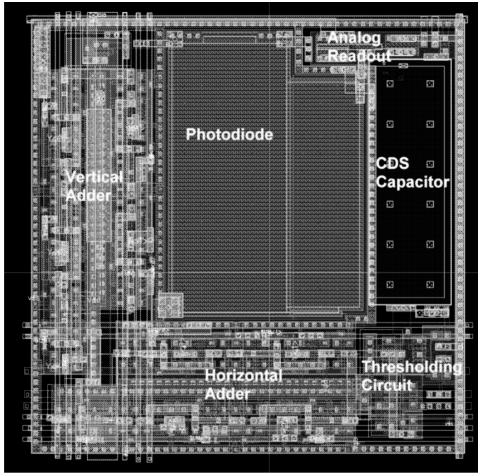


Figure 8. Pixel layout for the digital implementation of the cumulative cross section readout technique.

| Parameter/ Pixel subsystem | Time (Sec) | Constraints, influences |
|-------------------------------|---------------|--|
| Carrier generation | $<10^{-10}$ | Light propagation |
| Charge separation | 10^{-8} | Absorber thickness; carrier diffusion |
| Source follower | $<10^{-8}$ | Design rules |
| Comparator | 10^{-6} | In present design |
| Bus | $<10^{-8}$ | Bus capacitance, power |
| Logic | $<10^{-8}$ | $0.18 \mu\text{m}$ design rules |

Table 2. Estimated values of parameters which could potentially limit readout speed in the CCS sensor.

If we assume that a 10mV signal is required for reliable thresholding, and that the sensor has the modest conversion gain of $1\mu\text{V}/\text{electron}$, a photo-generated charge of about 10^4 electrons is needed to switch the pixel-comparator. For 500nm photons, and assuming an external quantum efficiency of 0.5 , this represents an energy deposition of 10^{-14}J per pixel. For $10\mu\text{m}$ square pixels, the corresponding energy density is 10^{-8}Jcm^{-2} . To achieve MHz operation (10^{-6} s integration time) therefore requires optical power densities on the order of 10mWcm^{-2} , which is fully practical for the rangefinder application discussed above.

Multiple spots in the image can give rise to undesirable artifacts in the centroid value. The removal of such artifacts can be achieved in several ways, including implementation of a region-of-interest partition, discrimination on the basis of cluster size (e.g. by threshold adjustment), or the use of a multi-axis readout. Aspects of the latter two concepts are described in the following section.

4. EXTENSIONS OF BASIC SYSTEM

4a. Centroid determination using multiple thresholds

While the accuracy of centroid determination for a single threshold is limited to half a pixel in binary OLS and a few tenths of a pixel for the CCS readout, the use of multiple thresholds applied sequentially can yield significantly improved accuracies. Hence an accuracy versus update speed trade-off can be made. Figures 9 and 10 present the simulated results for the accuracy of centroid location for spots with a Gaussian intensity profile as a function of the number of combined images with different thresholds used in the calculation. For the binary OLS it was found that around 16 images were required to achieve an accuracy below 0.1 pixel, the level often considered to be the minimum required in satellite guidance applications. Because of the additional intensity information contained in the CCS data, centroid accuracies are improved over the OLS system by about a factor of five. Combining several frames with different thresholds yields some initial improvement in accuracy due to the reduction of the undue influence of pixels with signals very close to the threshold and which may be erroneously included in or excluded from the calculation.

An additional use of the multiple thresholds in the CCS case is to determine the variation of centroid location with threshold value. This yields information about the symmetry of the spot's intensity profile, thereby facilitating optical alignment and diagnosis.

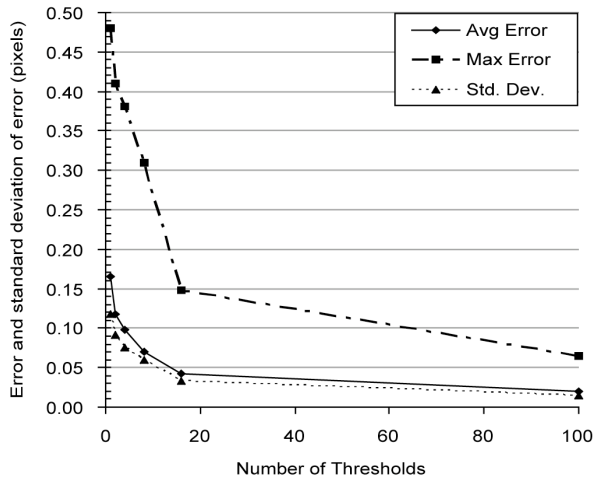


Figure 9. Simulated multi-threshold centroid accuracy for a range of varied Gaussian bright spots using the binary readout.

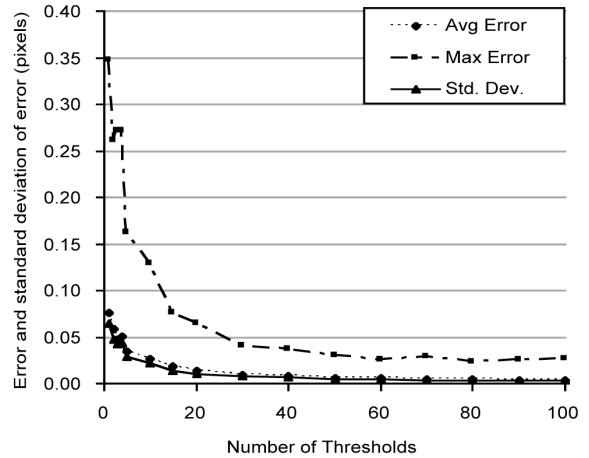


Figure 10. Simulated multi-threshold centroid accuracy for a range of varied Gaussian bright spots using the cumulative cross section readout.

4b. Multiple-axis readout for identifying multiple objects

The object location approach described above projects the image spots onto two orthogonal axes. While this technique is straightforward and unambiguous for single spots, the present of additional active regions, as shown in Fig.11(a), quickly renders even counting the spots, let alone locating them accurately, impossible. One method for reducing the ambiguity is to project the image onto multiple axes, as illustrated in Fig.11(b). In Fig.11, the bounding polygons indicate areas where image reconstruction would incorrectly interpret an object to be. Clearly, as the number of spots to be identified increases, so too does the number of axes required. Results of simulations of this scenario are presented in Fig. 12, where randomly located spots of arbitrary aspect ratio are examined using the binary object location approach on a 256 x 256 pixel array using a variable number of axis directions. Spots were circles stretched vertically and horizontally, with diameters ranging from 2 to 6 pixels. For the same test image and number of axes, the number of valid pixels depends on the precise angles of the axes; this variation is indicated by the error bars on Fig. 12. Here, the fraction of valid pixels is defined as the proportion of reconstructed pixels that accurately corresponds to pixels above threshold in the original image. From these data it is clear that the addition of even a few additional directions of

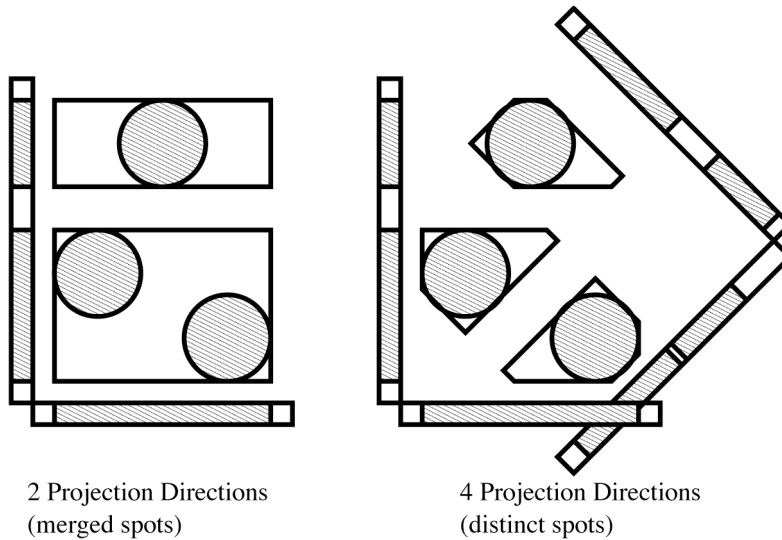


Figure 11. Illustration of the effect of number of readout directions on the accuracy of retrieval of information when multiple spots are present in the object location sensor field of view.

projection can provide a significantly increased reconstruction accuracy. We expect that accuracy can be improved still further if this approach is combined with other techniques, such as multiple thresholds and the cumulative cross sectional readout. Multiple axes of projection can be implemented in at least two ways: physically angled readout buses and ‘soft’ readout patterns of data stored in a frame memory.

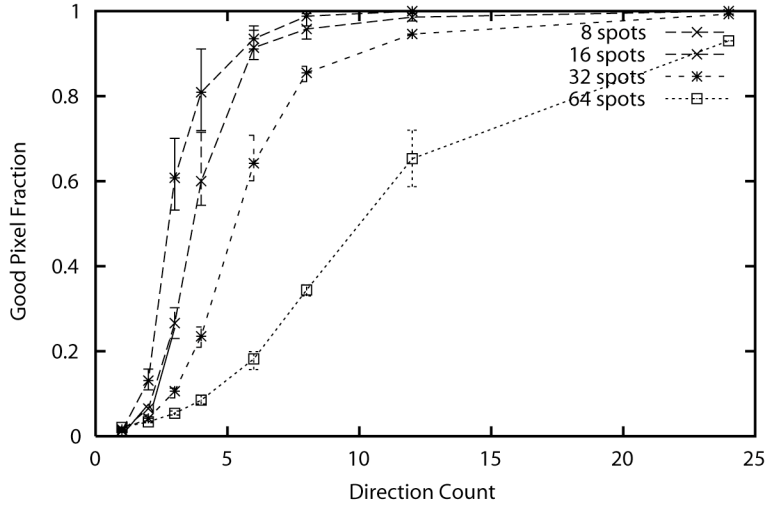


Figure 12. Simulated results illustrating the relationship between the accuracy of the information retrieval using a binary readout object location sensor and the number directions in which the data are read out. Results for various numbers of spots in the image are shown.

Advances in CMOS fabrication technology have been associated with increased numbers of metallization layers for interconnects, going from two or three five years ago to more than six in recent technologies. Hence the flexibility to design an image sensor in which the binary data can be readout in several directions simultaneously now exists. In the limiting case, global reset and threshold signals, as well as Vdd and Vss, can all be supplied to a pixel using a single metal layer. In such a case, pixel fill factor would be sacrificed because these four signals must be routed in parallel connections (although much of this fill factor could be recovered by using microlenses). Hence at most $(M - 1)$ readout directions are possible, where M is the number of metallization layers in the process used; for the TSMC 0.18 μm process used in this work, the limit is thus five readout directions.

A more complex implementation of the multi-directional readout technique would involve a 1-bit frame-store memory which could be read out in predetermined patterns to emulate angled data buses. This approach would have the advantage of flexibility, at the probable expense of silicon real estate and operational complexity. Speed would also suffer as parallel readout would not be possible.

A simple statistical model can be developed which yields a reasonable fit to the simulated data. The fraction of valid pixels is given by

$$P_{\text{valid}} = \frac{P_{\text{spot}}}{P_{\text{spot}} + P_{\text{false}}} \quad (1)$$

where P_{spot} is the probability that the pixel is active. A falsely reconstructed pixel, P_{false} , results from the situation where all projection axes through that pixel contain an above-threshold pixel but the pixel itself is not above threshold. Hence $P_{\text{false}} = (1 - P_{\text{spot}}) \cdot P_{\text{all}}$, where P_{all} is the probability that all lines contain an active pixel. Hence (1) becomes

$$P_{\text{valid}} = \frac{P_{\text{spot}}}{P_{\text{spot}} + (1 - P_{\text{spot}})P_{\text{all}}} \quad (2)$$

In addition, the approximation is made that the probability of each ray (i.e. the line parallel to the read-out axis that passes through a particular pixel) intersecting above-threshold pixels can be considered to be independent of that for any other ray. Hence P_{all} can be approximated as

$$P_{all} = \left(1 - P_{line_clear}\right)^N \quad (3)$$

where N is the number of directions and P_{line_clear} is the probability that the direction contains no active pixels. This assumption holds best when spots are far enough from the pixel under test to overlap at most one line; if this assumption does not hold, an erroneously large value for P_{valid} results.

For circular spots, the probability of a ray passing through a spot is given by the relative angular subtense of the spot:

$$P_{line_spot,j} = \left(\frac{2}{\pi}\right) \sin^{-1}\left(\frac{r_{spot,j}}{d_j}\right) \quad (4)$$

where $r_{spot,j}$ and d_j are the spot radius and the distance of the spot from the pixel for spot j , respectively. For simplicity, it is assumed that the spots have an average size (r_{avg}) and distance (d_{avg}) from the pixel in question, hence

$$P_{line_clear} = \left[1 - \left(\frac{2}{\pi}\right) \sin^{-1}\left(\frac{r_{avg}}{d_{avg}}\right)\right]^N \quad (5)$$

Substituting (4) into (5) and allowing $d_{avg} \gg r_{avg}$, we find

$$P_{line_clear} = \left[1 - \left(\frac{2}{\pi}\right) \left(\frac{r_{avg}}{d_{avg}}\right)\right]^N \quad (6)$$

While the expression in (6) has been found to give a reasonable indication of the observed results (Fig.12) and an insight into the process, it overestimates the fraction of validly reconstructed pixels, especially when spot sizes are large.

4c. Run-length encoding

Run-length encoding of a row or column (with binary output) determines the number of pixels between transitions in that row or column. A limiting feature of the projection approach to identifying spots is an inability to distinguish between multiple spots lying along the projection axis. Run-length encoding directly addresses this limitation by extending the projection mechanism to encode information about “spans” of light or dark pixels present in each projected column. This allows additional geometrical information to be obtained, while not substantially increasing the amount of data that needs to be transferred in most cases (the amount of information is proportional to the number of objects of interest in a column).

The number of spans that can potentially exist in a column is equal to the number of pixels. For spot sizes larger than one pixel, the number of spans that may fit in one column is reduced, but still can be very large. If span information for a column is read in parallel (all spans at once over a wide bus), the number of spans that may be recorded is limited by the amount of chip area that is allocated to the readout bus. If, on the other hand, span information for a column is read out serially (one span at a time over a narrow bus), the number of spans that may be recorded is limited by the time and power constraints on readout. In both cases, the practical limit to the number of spans to be read out will much lower than the worst-case number of spans required to encode the column data. This results in artifacts in the reconstructed image.

Figure 13 shows the effects of run-length encoding on an image containing two spots, for run-length encoding capable of representing one, two, three, and four spans of pixels. Of particular note are the effects of using an even or odd number of spans in the encoding. For the common case where the first span is a span of background pixels, objects have “shadows” of reconstructed-object pixels along the direction of projection. When an odd number of spans is used to encode the image, these shadows are merged with the reconstructed spots. When an even number of spans is used to encode the image, however, the shadow and true object pixels are separated by a one-pixel boundary of background pixels. This reflects the fact that, while the true states of the image pixels in the shadow region are unknown, there must be at least one background pixel to terminate the final span of object pixels that was encoded. The use of an even number of encoded spans can therefore be used to ensure that the presence of unknown pixels does not skew centroiding calculations for objects whose boundaries are precisely known. The criterion for this pattern, namely that the initial span encodes background pixels, can be enforced either by allowing span sizes of zero (recording a span of zero background pixels where an object overlaps the image border), or by encoding the image as if an extra line of background pixels was present at the edge of the image nearest the projection plane.

A simple extension of run length encoding might be to allow a row or column output only when n contiguous pixels are above threshold in order to discriminate against small objects, thereby reducing spatial noise.

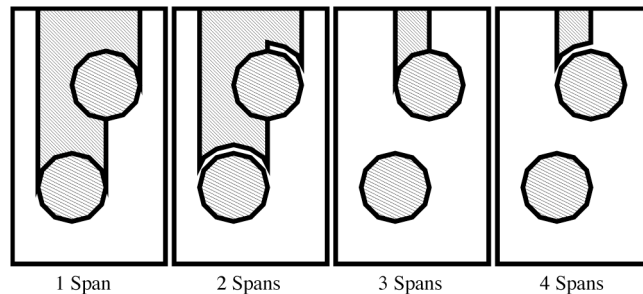


Figure 13. Distinguishing a pair of object spots using run-length encoding for various numbers of spans.

Two extensions to run-length encoding of multi-spot images are considered: double-ended projection, and multi-axis projection. Both involve performing run-length encoding in multiple directions. The first case, double-ended projection, involves projection in opposite directions along the same axis (Fig. 14). This can be thought of in two ways. The first way is to think of it as reducing the effects of projection “shadow” by eroding the shadow from both directions, confining it much more rapidly. The second way is to think of this approach as doubling the number of spans encoded for a column without a time tradeoff, by reading two halves of the span list in parallel. This comes at the cost of the area used by a second readout bus, for the simplest implementation of double-ended projection. Multi-axis projection involves run-length encoding along multiple different projection axes (Fig.15). This can be thought of as an augmented version of the many-axis spot disambiguation approach, discussed above. Using run-length encoding allows the many-axis projection approach to tolerate the presence of a small number of spots in a projection column while still using the projection column to flag non-object pixels.

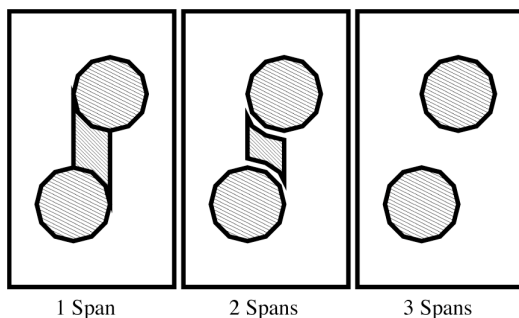


Figure 14. Double-ended run-length encoding technique.

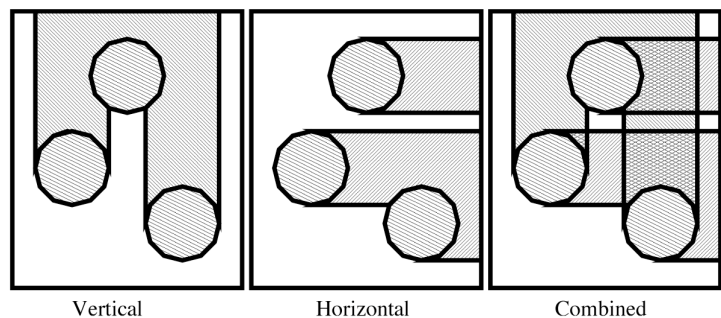


Figure 15. Run-length encoding performed along multiple axes.

5. CONCLUSIONS

It has been demonstrated that, for a range of applications, a custom CMOS image sensor with global thresholding and parallel readout can offer a fast method for determining the presence, location, and size of an object in the field of view. Pixel-parallel operation and straightforward techniques for the removal of spatial noise are advantageous in applications for which high-speed operation are more important than imaging the object, and for which the object can readily be distinguished from its background. A system in which numbers of above-threshold pixels in each row and column are summed is shown to yield more information than the simple binary readout system presented previously. Data for an 80 x 80 pixel analog implementation of the cumulative cross-section sensor are presented and the ultimate performance limits of such a system are discussed. Several possible extensions to this approach are also introduced. Firstly, the possibility of centroid calculation to sub-one-tenth pixel accuracy, using multiple thresholds, is established. Subsequently, a multi-axis readout technique for improved image reconstruction in the presence of multiple objects is described. Lastly, the application of run-length encoding to object discrimination is introduced.

Future developments of this work include the implementation of a digital version of the cumulative cross-section sensor which enables higher-speed readout and on-chip determination of a region of interest for subsequent analog readout. The applications of this system to the detection of intensity modulation will also be investigated.

6. ACKNOWLEDGEMENTS

The authors wish to acknowledge financial support from Topaz Technology Inc. and the Centre for Research in Earth and Space Technology (CRESTech). Support from the Canadian Microelectronics Corporation for chip fabrication is also gratefully acknowledged.

7. REFERENCES

- [1] A. Watanabe, O. Tooyama, M. Miyama, M. Yoshimoto, and J. Akita (2001), "An image sensor with fast extraction of objects' positions - rough vision processor," in *Proc. Int'l Conf. Image Processing*, vol. 2, Oct. 2001, pp. 1105–1108.
- [2] R. Blum, C. Wilson, P. Hasler, and S. DeWeerth (2002), "A CMOS imager with real-time frame differencing and centroid computation," in *IEEE Int'l Symp. Circuits and Systems (ISCAS)*, vol. 3, 2002, pp. 329–332.
- [3] N. Yu, T. Shibata, and T. Ohmi (1998), "A real-time center-of-mass tracker circuit implemented by neuron MOS technology," *IEEE Trans. Circuits and Systems II*, vol. 45, no. 4, pp. 495–503, April 1998.
- [4] G. Erten and S. Hagopian (1999), "Integrated image sensor processor with on-chip centroiding function," in *42nd Midwest Symp. Circuits and Systems*, vol. 1, 1999, pp. 262–265.
- [5] C. S. Hong, R. Hornsey, and P. Thomas (2001), "Single-chip camera modules for mosaic image sensor," in *SPIE Electronic Imaging 2001*, San Jose, CA, January 2001.
- [6] C. Hong (2001), "On-chip spatial image processing with CMOS active pixel sensors," Ph.D. thesis, University of Waterloo, Canada.
- [7] J. Shah (2002), "Applications and implementations of centroiding using CMOS image sensors," MASc thesis, University of Waterloo, Canada.
- [8] R.D. Burns (2003), "Improved object location techniques with CMOS image sensors", MASc thesis, University of Waterloo, Canada.
- [9] R.D. Burns and R. Hornsey (2003), "CMOS Image Sensor with Cumulative Cross Section Readout", presented at the 2003 IEEE Workshop on CCDs and Advanced Image Sensors, Elmau, Germany, May 2003.