Integrated Sensor and Rangefinding Analog Signal Processor

L. Richard Carley†, Andrew Gruss†, and Takeo Kanade‡
Department of Electrical & Computer Engineering†
and School of Computer Science‡
Carnegie Mellon University, Pittsburgh, PA 15213 *

Abstract

Rangefinding, the measurement of the three dimensional profile of an object or scene, is a critical component for many robotic and sensing applications. This paper describes a light-stripe rangefinding system consisting of a 6x10 array of pixels. Each pixel contains a photodiode to detect the light stripe and a primarily analog signal processor to determine and store the time at which the light stripe crosses that photodiode. Incorporating signal processing into each pixel makes it possible to modify the light-stripe rangefinding algorithm to achieve a frame rate over *two orders of magnitude* faster than conventional light-stripe rangefinding methods; i.e., over 100 frames/sec.

1 Introduction

Rangefinding, the measurement of the three dimensional profile of an object or scene, is a critical component for many robotic applications. Many different rangefinding techniques have been developed [2] and light-stripe rangefinding is one of the most common and reliable methods. A conventional light-stripe rangefinder operates in a serial manner — a stripe source is projected onto a scene, a video image is acquired, the position of the projected light stripe is extracted from a video camera image, the stripe is is moved slightly across the scene, and the process repeats until the entire scene is scanned (see figure 1). Range acquisition rates achievable using this method are limited by the the video camera's frame rate, with the number of frames required for a single complete range map increasing linearly with the desired range resolution. Typically, on the order of one second is required to acquire a complete range image in this manner.

The light-stripe range sensor presented in this paper is based on the modification of this basic light-stripe rangfinding technique described by Sato [8] and Kida [4]. In this modified method, the range image is constructed in a parallel manner rather than a serial one. The stripe is swept continuously across the scene and the time at which the light stripe crosses each pixel is detected and recorded as a measure of the range along that line of sight. Although this modified algorithm achieves a great improvement in range data frame rate, it can be severely limited in resolution (the number of pixels), because detection of the time at which the light stripe crosses each pixel is generally performed by separate hardware for each pixel.

In this paper, we propose to overcome this limit by integrating into each pixel the signal processing circuitry to detect and remember the time when the light stripe crosses that pixel. Thus an entire range map is acquired in parallel, and the total time of acquisition is independent of the range map resolution. We have achieved a speed increase of over two orders of magnitude in the frame rate over a conventional camera-based light-stripe rangefinding system.

The novelty of this approach is its use of *smart* sensors, ones which provide signal processing at the point of sensing. In this case, the increase in cell complexity from sensing-only to sensing-and-processing makes practical the modification of the operational principle of rangfinding, which in turn results in a dramatic improvement in performance. Examples of this class of chip exist and include commercial CCD image chips, the *Xerox Optical Mouse* [6], Mead's *Artificial Retina* [9], and an *Optical Position Encoder* done at the CSEM in Switzerland [1]. In order to keep the pixel area small, the necessary signal processing functions have been implemented using primarily analog circuitry.

2 Light-Stripe Rangefinding

In this section we briefly review the conventional light-stripe rangefinding method and the modified method that has been implemented. Figure 1 shows the geometrical principle on which a light-stripe rangefinder operates. The scene is illu-

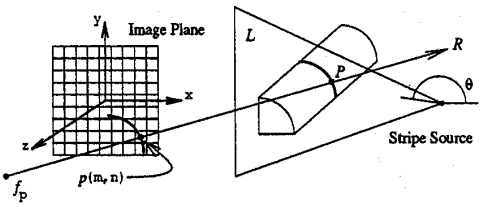

Figure 1: Lightstripe Rangefinder Geometry

*An AT&T Foundation Grant provided partial support for equipment used in this research.

minated with a vertical plane of light. The light is intercepted by an object surface in the path of the beam and, when seen by a video camera placed to the left of the light source, appears as a stripe which follows the surface contour of objects in the scene.

Range data along the contour can be calculated easily using the principle of triangulation. In figure 1, the equation of the plane of light L is known because the projection angle θ is controlled. The line of sight R for each point p on the image of the stripe can also be determined by tracing a line from the image focal point f_p through p. The intersection of the ray R with the plane L uniquely determines the three-dimensional position of P on the surface corresponding to p.

A conventional rangefinder collects range data for an entire scene sequentially — iterating the process of fixing the stripe on the scene, taking a picture, and signal processing each picture to detect the position of the light stripe, until the entire scene has been scanned.

Though practical, the speed of sampling range data by the conventional light-stripe technique is severely limited because the number of pictures which must be taken and processed increases linearly with the desired resolution in the depth dimension.

In the modified (parallel) rangefinding technique, the video camera is replaced by a two-dimensional array of *smart* photosensitive cells. Range data is not acquired in a step-and-repeat manner. Instead, the plane of light is swept across the scene at a constant angular velocity once from left to right. This slight modification to the conventional light-stripe rangefinding algorithm transforms the rangefinding process from a sequential to a parallel one, and results in a qualitative change in the way the range measurement is determined.

Figure 2 shows the array of photosensitive cells including the analog signal processing and readout logic. Each

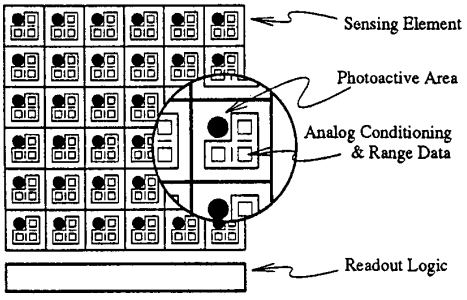

Figure 2: 2D Array of *Smart* Photosensors

pixel detects and remembers the *time* at which it observed the peak incident light intensity during a sweep of the stripe L. Observe that each cell predefines a unique line of sight R and that the time information, t_{cell}, recorded by each cell determines a particular orientation of the stripe $\theta_{t_{cell}}$. Recalling the geometry in figure 1, one sees that this information is sufficient to calculate the three-dimensional position of the imaged object point P, again using triangulation. The data gathered during *one* pass of the stripe in *each* cell of an $M \times M$ array of these smart sensing elements is sufficient to calculate the $M \times M$ range map of an imaged scene. For the new method, the sweep time (T_s) is only limited by the photoreceptor sensitivity at the chosen bandwidth. The resolution, the number of pixels, is limited by integration technology; that is, the size of the sensor and the necessary signal processing circuitry and the yield characteristics of the available technology.

In addition, the resolution of the distance measurement also increases when using the modified light-stripe rangefinding method, even for arrays with a small number of pixels. In the conventional camera-based method, the number of pixels in the horizontal scan line (typically 256 to 512 for a camera) limits accuracy because the range information is derived from the *position* of the peak of the light intensity profile on the scan line. Thus, subpixel peak localization techniques, which can at best provide one tenth of a pixel spacing accuracy, are often employed. It is interesting to note that since the accuracy relies on the interpolation of a spatially sampled profile, an extremely fine light stripe which provides a good $x - y$ resolution, is not necessarily the best for obtaining z accuracy. However, in the modified method the peak of the *continuous* uninterpolated time profile of intensity from the same cell is detected in the *time* domain with much greater accuracy. The spatial position of the light stripe is accurately determined by a shaft encoder on the light stripe projector; hence, a fine light stripe directly contributes to an increase of both $x - y$ resolution and z accuracy.

3 Sensor/Signal Processor Cell

Functionally, each element of the smart photosensitive array converts light energy into an analog voltage, determines the time at which the voltage peaks, and remembers the time at which the peak occurred. In this section we describe how these functions were implemented in our prototype. The implementation of these functions requires that a photoreceptor and mixed analog/digital signal processing circuitry be integrated into each pixel. The photosensor design must take into consideration tradeoffs in pixel size, power dissipation, bandwidth, sensitivity, and accuracy. In order to minimize pixel die area, the pixel design was a primitive as possible (see in figure 3), while still providing the necessary functionality. For example, a simple threshold detector follows the preamp, even though a matched filter should ideally be used before the comparator. Sensor operation consists of two phases — data collection and data readout. Phase one begins by resetting all state within the cells and then each cell collects the data as the image of the light stripe sweeps across the sensor. During phase two, all cell data is offloaded from the chip. The chip is again reset and another scan can be taken.

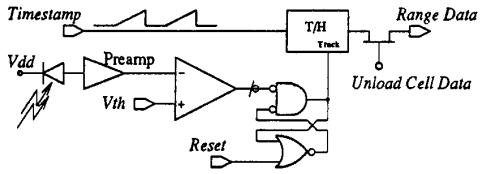

Figure 3: Sensing Element Circuitry

3.1 Integrated Photodiodes and Preamplifier

The photodiodes are critical to the sensitivity and bandwidth of the sensor cells. Current output at a given incident light intensity is directly proportional to the photodiode area as is the junction capacitance. The more area devoted to photodiode structures the better the optical sensitivity of the sensing elements. In the sensor array prototype, the photodiodes are approximately $8,000\ \mu m^2$ in area, one fourth the total area budgeted for a cell.

In a CMOS process, maximum sensitivity photodiodes are built using the well-substrate junction [1]. In a P-Well CMPS process, this vertical photodiode structure is constructed using the n-type substrate as the cathode and the p-type well as the anode, as shown in figure 4. An additional p^+ implant

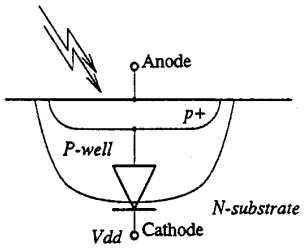

Figure 4: Vertical Photodiode Structure

is driven into the well to reduce the surface resistivity of the anode to which contact is made. Finally, the photodiode structures are surrounded with guard rings to help reduce the contribution of noise from other current in the substrate and to minimize the chance of photocurrent induced latchup. We fabricated photodiodes with and without the passivation layer and found only slight degradation of sensitivity with the overglass in place. Photodiode anodes are high impedance nodes (typical output currents are on the order of picoamps) and will be susceptible to noise coupling from other chip circuitry. For example, digital circuits switching while stripe detection is taking place could inject noise and corrupt the delicate photocurrent measurements. Though techniques exist which can help alleviate these problems [7], the prototype was designed to have no clock signals active during the sweep of the stripe. Only during the readout phase are clock signals switching.

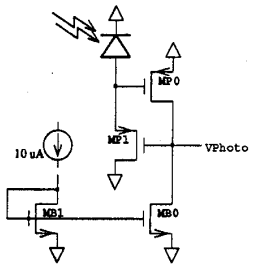

Figure 5: Photodiode and Transimpedance Preamplifier

In addition to the light-stripe range finding array IC, a test IC that included just photodiodes and the transimpedance preamp circuit (see figure 5) was also fabricated using the same process (MOSIS [5] system on a 2μ CMOS p-well double-metal, double-poly process). In this circuit the preamp output was brought to pads so that the performance of the photodiode and preamp could be characterized. Transistor MP1 in the preamp is operating at the DC photocurrent plus the reverse bias leakage current of the photodiode, and is thus in the weak inversion region. The other transistors act as a simple op amp to force the gate-source voltage of MP1 to the necessary value to carry the photocurrent plus leakage current. Since MP1 is operating in weak inversion, the output voltage of this circuit is proportional to the log of the detected photocurrent. Hence, the effective gain of the preamplifier is inversely proportional to the sum of the DC photocurrent and photodiode leakage current. Similarly, the bandwidth of this circuit is directly proportional to this same current. In effect, a uniform background illumination of the sensor IC can be used to make a tradeoff between sensitivity and bandwidth. However, in order to prevent background light from the scene from generating a nonuniform gain across the array, a frequency selective light filter with a center frequency at that of the laser generating the stripe is placed in front of the camera lens.

The preamp test chip was mounted on the camera of the rangefinder set-up and tested. Figure 6 shows the experimental results of one such test. The traces in the figure show the response of two cells to the image of a swept HeNe laser stripe. It demonstrates that the signal has a clear peak whose location moves as the distance to the surface changes. It has been verified that this photodiode/preamplifier combination can successfully detect stripes swept at rates which correspond to $1,000\ frames/sec$ for relatively reflective surfaces and at $100\ frames/sec$ for very dark surfaces.

The shape of the photodiode areas is important. Consider the image of a stripe moving across the photodiode as sketched in figure 7. In the case of a square photocell, output pulse shape depends on stripe orientation — for a vertical stripe, the output pulse observed will be square, but a 45° stripe will produce a triangle wave. This difference in pulse shapes translates to an error in estimating the time at which

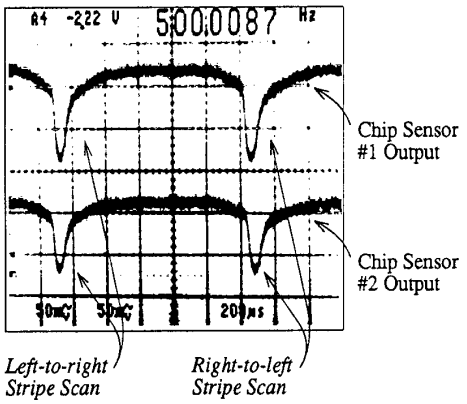

Figure 6: Amplified Photocurrent Oscilloscope Waveforms

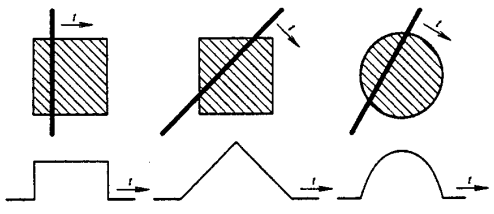

Figure 7: Photodiode Area and Pulse Shape

the peak of the stripe occurs. Therefore, the photodiodes are circular so that the pulse shape will be independent of stripe orientation.

3.2 Stripe Peak Detection and Storage

In the sensor array prototype, light stripe peak detection is done by thresholding. The amplified photodiode signal is compared to a reference voltage set to produce an output when the stripe image passes across the photodiode. The reference input is directly connected to an off-chip voltage source through one of the IC pins. Detection of the stripe trips the RS flip-flop in the cell, indicating that the timestamp value should be held (see figure 3).

The decision to avoid having any digital signals switching during the sweep directly affected the comparator design. The comparator in each sensing element is non-regenerative — essentially an uncompensated two stage opamp circuit [3] with no applied external feedback. Although a clocked cell design would allow the use of regenerative comparator topologies, which are generally faster and more accurate for a given circuit area, that the potential danger of injecting clock noise into the photocurrent measurement outweighed the potential benefits.

A system timestamp is broadcast to each cell from offchip as an *analog voltage ramp*. The timestamp voltage is held on the capacitor of a track-and-hold (T/H) circuit when the latch in the cell is tripped by the detection of the stripe. Representing time in an analog form has several advantages over the digital equivalent of latching the value of a continuously running counter. The analog-only scheme avoids noise problems we have discussed associated with mixing sensitive analog circuitry with digital logic. A multi-bit digital timestamp bussed over the entire chip, combined with transients associated with the latching of timestamp values by sensing elements, are sources of noise with the potential to corrupt the photocurrent measurements. Of course, information stored in analog form is subject to corruption from sources like noise and voltage droop. However, with careful circuit design these errors can be kept to 0.5% or so, yielding useful range accuracy.

In the prototype, a double-poly process was used to provide high quality linear capacitors that exhibit good matching across a die. The matching of these capacitors across the sensor chip will in large part determine the variance in voltage reported by individual sensors for given time values. However, even this is not a fundamental limitation as variations between capacitors can be calibrated out and corrected digitally off-chip.

4 IC Sensor Interface Circuitry

A complete sensor chip must include readout circuitry as well as an array of sensing elements. When a scan has been completed, each cell in the array will be holding its raw range datapoint in the form of a voltage. Readout of held timestamp values proceeds in a time-multiplexed fashion. A shift register, wound through the array, gates the charge held on each cell's T/H capacitor in turn onto a chip pin. This charge is then integrated by off-chip circuitry to produce the raw range data value for the enabled cell.

A die photo of the prototype sensor chip, including an array of 10×6 sensing elements and read-out circuitry, can be seen in figure 8. Each pixel (sensor and signal processor) is roughly $300\ \mu m \times 150\ \mu m$. Looking forward to integrating a large number of pixels on larger dies, the total power consumed by each pixel was kept small: less than $500\ \mu watts$.

Photodiode areas are circular to provide for consistent output pulse shapes as explained above. The capacitor needed for the T/H circuitry has been used to fill in the corners of the round photodiodes to minimize wasted die area. Sensing element support circuitry is sandwiched between the photodiodes. Spatial distribution of photodiode areas in the current version is not even. We plan to rearrange the layout of cell circuitry in future versions and redistribute the photodiodes. Global signals like the analog time ramp, global threshold voltage, array readout line, and the output enabling shift register clocks are bussed in one direction on the second metal layer. Supply rails run perpendicular to these global signals in the primary metal layer.

Another issue is the system interface to control the range image acquisition process. The sensor chip must by cycled

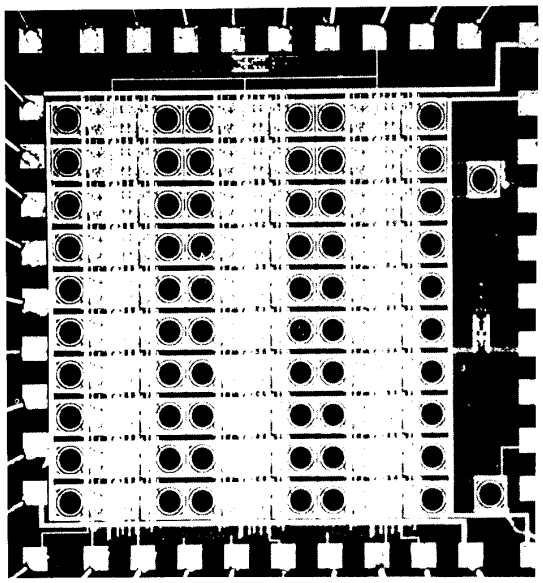

Figure 8: Sensor Chip Layout

and the raw range data read out. The range data is converted into digital form, and downloaded to the host computer over a high bandwidth interface.

Charge accumulated in the sensing elements on the holding capacitors is passed out of the chip on a bus and integrated to produce a voltage as shown in figure 9. We offload sensing

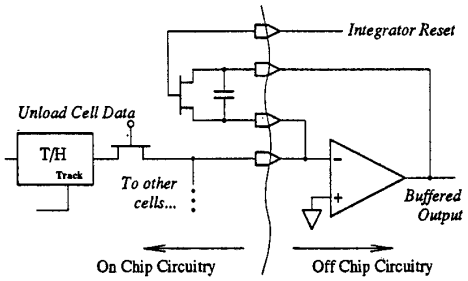

Figure 9: Off Chip Readout Circuitry

element range data in raster fashion, similar to the readout of a CCD imaging chip. Because the voltages held on the T/H capacitors will tend to droop, it is important to offload range samples from the chip immediately after they are taken. A 10-bit 1 MS/s analog-to-digital (A/D) converter is used convert the analog time values into digital form. Simple processing to convert this raw timestamp data into 3D (x, y, z) range map values is performed in the host computer system.

5 Integrated Sensor Test System

We have designed and built a prototype rangefinding system using the integrated sensor/signal processor IC described. Essential components in this system included stripe generation hardware, range sensor IC and support electronics, range sensor optics, and host interface (figure 10). The range sen-

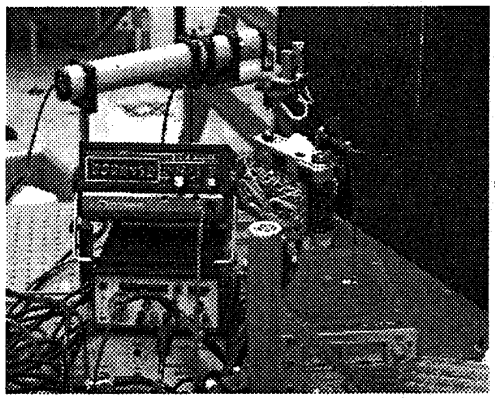

Figure 10: A Rangefinding System Set-up

sor IC was mounted in a 35mm camera body (figure 10) which provided a convenient mechanism for incorporating focusing optics and for sighting the rangefinder. The stripe was generated using a 5 mW helium-neon (HeNe) laser and half-cylindrical lens. A mirror mounted on a galvanometer provided the means to sweep the stripe. Two scans (one from left-to-right, the other from right-to-left) were generated during each period of up to 500 Hz triangle wave used to drive the galvanometer, providing the system with a maximum of 1,000 sweeps per second.

6 Results

Figure 11 shows the waveforms observed during normal rangefinder operation. The top trace shows the output of the op amp integrator that reflects the charge stored on each of the pixel hold capacitors. The second trace from the top shows the triangle wave timing reference waveform. At the left, an upward slope occurs during the sweep of the light stripe. Then, the triangle input is kept constant during the readout phase.

This waveform only shows the readout of 30 out of the 60 pixels. An vertical object (a ruler) was held a fixed distance in from of a reflective background. The effect of the target can be seen in the two different ranges being reported. At first, the less negative voltage level (upper trace) indicates the background. Then, there are two pixels which clearly reflected from the object, and their more negative voltage indicates its distance in front of the target. Note, one or two

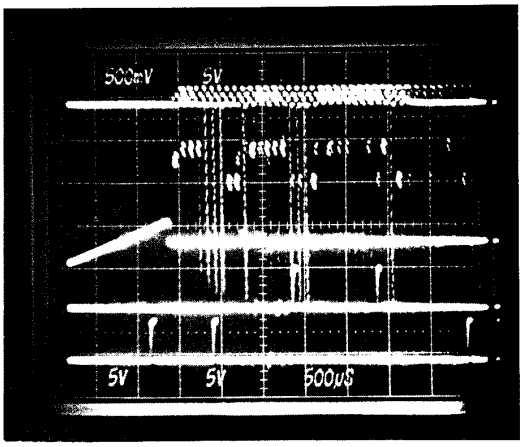

Figure 11: Waveforms During Rangefinder Operation

pixels on either side of the object failed to detect the light stripe, resulting in their voltages being extremely negative. Then, one row of pixels down this process repeats. However, there is a nonuniform strength to the signal returning to the sensor, which results in more of the second and third rows pixels being unable to detect the light stripe.

7 Conclusion

Advances in VLSI technology have made it possible to create *smart* sensors, ones that integrate sensing and signal processing. We have demonstrated a VLSI smart sensor based on light-stripe rangefinding which is capable of acquiring 100 to 1,000 frames of range images per second. One of the most distinguishing features of this approach is that it is not just parallel implementation of known algorithms by VLSI technology to achieve increased speed, such as VLSI chips for convolution. Rather, it demonstrates that integration of sensing and processing can make possible modifications of the operational principles of information acquisition (in our case, range imaging) which results in a qualitative improvement in performance.

References

[1] P. Aubert and H. Oguey. An application specific integrated circuit (ASIC) with CMOS-compatible light sensors for an optical position encoder. In *IEEE 1987 Custom Integrated Circuits Conference*, pages 712–716, IEEE, May 1987.

[2] P.J. Besl. *Range Imaging Sensors.* Research Publication GMR-6090, General Motors Research Laboratories, March 1988.

[3] R. Gregorian and G.C. Temes. *Analog* MOS *Integrated Circuits for Signal Processing.* John Wiley & Sons, New York, 1986.

[4] T. Kida, K. Sato, and S. Inokuchi. Realtime range imaging sensor. In *Proceedings 5th Sensing Forum*, pages 91–95, April 1988. In Japanese.

[5] G. Lewicki et al. *MOSIS User's Manual.* USC Information Sciences Institute, 4676 Admiralty Way, Marina Del Rey, CA 90292-6695, 1988.

[6] R.F. Lyon. *The Optical Mouse, and an Architectural Methodology for Smart Digital Sensors.* Technical Report VLSI-81-1, Xerox Palo Alto Research Center, August 1981.

[7] J.A. Olmstead and S. Vulih. Noise problems in mixed analog-digital integrated circuits. In *IEEE 1987 Custom Integrated Circuits Conference*, pages 659–662, IEEE, May 1987.

[8] Y. Sato, K. Araki, and S. Parthasarathy. High speed rangefinder. *SPIE*, 1987.

[9] M.A. Sivilotti, M.A. Mahowald, and C.A. Mead. Real-time visual computations using analog CMOS processing arrays. In P. Losleben, editor, *Advanced Research in* VLSI — *Proceedings of the 1987 Stanford Conference*, pages 295–312, The MIT Press, 1987.