A VLSI Sensor Based Rangefinding System

Takeo Kanade† Andrew Gruss‡ L. Richard Carley†

School of Computer Science† Department of Electrical and Computer Engineering‡ Carnegie Mellon University Pittsburgh, Pennsylvania 15213

Abstract

Lightstripe rangefinding is one of the most widely used and reliable techniques for the measurement of the three-dimensional profile of a scene. In a conventional lightstripe system, a complete range map is obtained via the step-and-repeat process of projecting a stripe, grabbing a camera image, extracting the projected stripe positions, and stepping the stripe until an entire scene has been scanned. Though practical, the speed of sampling range data with this conventional technique is severely limited.

We present a rangefinder capable of acquiring range information two orders of magnitude faster than conventional lightstripe methods. This rangefinder is based on a specialized VLSI sensor which gathers range data as a scene is swept continuously by a moving stripe. The sensor consists of a two-dimensional array of smart photosensitive cells. Each cell has circuitry that detects and remembers the time at which it observed the peak incident light intensity during a sweep of the stripe. A given cell predefines a unique line of sight and records a timestamp that determines a particular orientation of the stripe. Thus, information sufficient to extract a complete range image is gathered by the sensing elements during a single pass of the light stripe over the scene.

We first developed a working prototype rangefinding system which uses a 4 × 4 array of discrete photodiodes to sense the stripe. Then, VLSI sensing elements, combining both photosensing and analog signal conditioning, have been successfully fabricated and tested. Finally, a unique VLSI multi-element range sensor, which combines photosensing, signal conditioning, and signal processing on one monolithic CMOS chip, has been designed and is in fabrication.

1 Introduction

The goal of our research is to build a compact, fast lightstripe rangefinder using a VLSI smart photosensor array. Our target system will acquire 100 × 100 point range image frames 100 to 1,000 times per second with 0.5% range accuracy. It will be compact and rugged enough to be mounted on the end effector of a robot arm to aid in object manipulation and assembly tasks.

Rangefinding, the measurement of the three-dimensional profile of an object or scene, is a critical component for many robotic applications. Many rangefinding techniques have therefore been developed [Besl, 1988]. Of these, lightstripe rangefinding is one of the most widely used and reliable methods available. A conventional light stripe rangefinder operates in a step-and-repeat manner – a stripe source is projected on an object, a video image is acquired, the position of the projected light stripe is extracted from the image, the stripe is stepped, and the process repeats. Range acquisition rates achievable using this method are limited by the time needed to grab and process the video images, increasing linearly with the desired horizontal resolution. Typically, it takes one to several seconds to acquire a whole range image by this method.

The fast VLSI range sensor we are building is based on the modification of this basic lightstripe ranging technique in a manner described by Sato [Sato, 1987] and Kida [Kida, 1988]. In this modified method, a conventional camera is replaced by a smart sensor. Each cell of the smart sensor detects and remembers the time when it sees the light stripe while the stripe source sweeps continuously across the scene. Thus an entire range map is acquired in parallel, and total time of acquisition is independent of the range map resolution. We can expect a speed up of a few orders of magnitude in the frame rate over a conventional camera-based system.

The novelty of this approach is the use of smart sensors, sensors which provide processing at the point of sensing. Each smart cell independently senses and processes the signal. The slight increase in cell functionality from sensing-only to sensing-and-processing allows for the modification of the operational principle of ranging, which in turn results in drastic improvement of performance. The practical realization of the new rangefinding technique, however, requires the use of VLSI technology – the ability to integrate photoreceptors, analog circuitry, and digital logic on a single CMOS die. This paper presents our progress in developing a fast VLSI sensor based rangefinder.
2 Rangefinding by a Smart Sensor

A slight modification to the conventional lightstripe 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. The following discussion illustrates the difference in the two techniques.

Figure 1 shows the geometrical principle on which a lightstripe rangefinder operates. The scene is illuminated 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 \( \theta \) 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 *step-and-repeat* rangefinder collects range data for an entire scene sequentially – iterating the process of fixing the stripe on the scene, taking a picture, and processing the resultant image until the entire scene has been scanned.

Though practical, the speed of sampling range data by the conventional light stripe technique is severely limited. Assume that a video camera image has \( N \) rows. Since from one image at each step we can obtain up to \( N \) data points, the maximum speed of sampling is \( S_{max} = \frac{N}{T_f} \) where \( T_f \) is the time required to acquire and process an image frame. Typically, \( N \) ranges between 256 and 512 samples and \( T_f \) ranges between 0.3/50\( \text{second} \) and 1/10\( \text{second} \). Thus, sampling speeds of camera based systems are limited to

\[
S_{max} \approx 2.5K \sim 15K \text{samples/second}.
\]

In the parallel rangefinding technique, the video camera is replaced by a two-dimensional array of smart photosensitive cells. In addition, 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.

The array of photosensitive cells, shown in figure 2, is smart for the following reason. Each has circuitry that can detect and remember 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 \( \theta_{cell} \). 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.

![Figure 2: 2D Array of Smart Photosensors](image)

For an \( M \times M \) array of these cells, the sweep time \( T_s \) of the lightstripe and \( M \) determine the sampling speed. The sweep time \( T_s \) will be limited by photoreceptor sensitivity and \( M \) by integration technology. We are predicting values of \( T_s \) on the order of one to ten milliseconds and values of \( M \) ranging between 40 and 100. The rates at which the smart sensor array generates range data will be

\[
S_{max}^{Smart} = \frac{M^2}{T_s} \approx 0.2 \sim 10 \text{Msamples/second}.
\]

This is a speedup of a few orders of magnitude over that of a conventional camera-based system.

In addition, the resolution of the distance measurement also increases. In the basic method, the number of pixels in the horizontal scan line (typically 256 to 512) limits accuracy because the range information is derived from the position of the peak of the light intensity profile on the scan line. Thus, quite often subpixel peak localization techniques, which provide at most one tenth of a pixel spacing accuracy, are 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. In the modified method, however, 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 of the light stripe projector. With the modified method, a fine light stripe directly contributes to an increase of both $x-y$ resolution and $z$ accuracy.

3 Discrete Photodiode Based System

3.1 Implementation

We have designed and built a prototype rangefinding system using a sensor constructed out of a $4 \times 4$ array of discrete photodiodes. Essential components in this system included stripe generation hardware, range sensor, range sensor optics, and host interface (figure 3). This implementation is similar in spirit to systems built by others [Sato, 1987,Kida, 1988] and served as groundwork for our VLSI sensor based system.

The photodiodes were mounted in a 35mm camera body (figure 4) which provided a convenient mechanism for incorporating focusing optics and for sighting the rangefinder. Photodiodes were chosen for the sensing elements because they possess bandwidth sufficient to meet our sweep rate specifications.

Analog signal conditioning circuitry for each of the discrete photodiodes was designed to provide a digital transition when the reflection of the stripe passed through a diode’s field of view. It consisted of a high gain transimpedance amplifier and simple thresholding stage as shown in figure 5.

The design of the photocurrent amplifier was crucial. For purposes of small signal analysis, reversed biased photodiodes can be modeled as the parallel combination of a current source of a few nanoamps, a resistance, and a capacitance. Though the internal capacitance is moderate (on the order of a few picofarads), the internal resistance is ten gigaohms or more. This parallel RC combination creates an undesirable pole at a frequency of a few hertz which suppresses high frequency output signal components. The photocurrent amplification circuitry must provide a large photocurrent to voltage gain while minimizing the effect of the internal photodiode capacitance on bandwidth. The amplifier shown in figure 5 provides $18 M\Omega$ of transimpedance gain and employs negative feedback to servo the photodiode anode to a constant voltage, increasing available output signal bandwidth.

Amplified photodiode signals were highpass filtered before reaching the comparator. This was done to make the comparator threshold level independent of photodiode dark current and ambient light levels and to remove low frequency circuit noise. The digital output from the comparator in each cell was passed directly to the host interface where the sixteen comparator outputs were sampled into a local dual-ported memory. Host access to the data was provided via a memory-mapped VME port.

The stripe in this discrete implementation 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

![Figure 3: A Rangefinding System Set-up](image)

![Figure 4: A $4 \times 4$ Photodiode Array Mounted in a 35mm Camera](image)

![Figure 5: Photodiode Signal Conditioning Circuitry](image)
from right-to-left) were generated during each period of a 500 Hz triangle wave used to drive the galvonometer, providing the system with the desired 1,000 sweeps per second.

3.2 Results

The photodiode based rangefinding system with discrete components was able to generate and record over 1,000 4 x 4 range images each second. System software, running on a SUN 3/160 workstation, slowed the rate of processed image data to about 100 frames a second. Range data, encoded in the time from a scan origin to when a sensing element sees the flash, was continuously displayed on the monitor of the host workstation.

4 VLSI Sensor Based System

From a speed and sensitivity standpoint, our discrete photodiode based system is a successful implementation of the fast rangefinding algorithm. However, a 4 x 4 array does not provide enough range points to be useful. As one considers building larger and larger arrays out of discrete photodiodes, implementation problems quickly become apparent. The cost of wiring individual photodiodes to interface circuitry is prohibitive for arrays much larger than 10 x 10. Support circuitry, built out of off-the-shelf analog and digital IC components for each sensing element, will be bulky and impractical.

A VLSI implementation of the range sensor shows the greatest promise for increasing the range image density of the system. The essential advantage gained through the use of VLSI technology is the ability to integrate the photoreceptor, analog circuitry, range memory, and signal processing into each smart cell. Examples of this class of chip exist and include commercial CCD image chips, the Xerox Optical Mouse [Lyon, 1981], Mead's Artificial Retina [Sivilotti, 1987], and an Optical Position Encoder done at the CSEM in Switzerland [Aubert, 1987].

4.1 Sensor Element Design

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. The implementation of this functional specification requires that photoreceptor and signal conditioning circuitry be integrated into a unique hybrid sensor cell. The sensing element design must consider tradeoffs in cell size, power dissipation, bandwidth, sensitivity, and accuracy. Our initial VLSI design, sketched in figure 6, provides this functionality. We aimed at simplicity in the initial design - for example we use thresholding for peak detection. In future designs we may begin to implement more intelligent and accurate stripe peak detection.

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 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. Pipelining the storage within the cells could have been used to overlap the two chip phases, but again we chose not to do this in the initial design.

4.1.1 Integrated Photodiodes

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. The more area devoted to photodiode structures the better the optical sensitivity of the sensing elements. Our photodiodes are approximately 8,000 μm² in area, one half the total area budgeted for a cell.

In a CMOS process, maximum sensitivity photodiodes are built using the well-substrate junction [Aubert, 1987]. 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 7. An additional p⁺ implant 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. Only the anode of the photodiode is accessible for reversed biased operation. The cathode of the diode is the substrate and will be at the substrate voltage.

The shape of the photodiode areas is important. Consider the image of a stripe moving across the photodiode as sketched in figure 8. 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 a frequency characteristic which is a function of imaged surface orientation. The accuracy with which the time of peak signal amplitude can be determined will be affected by the phase differences changing pulse shapes represent. We have chosen to make our photodiode areas circular so that pulse shape is independent of stripe orientation.

4.1.2 Amplification and Stripe Peak Detection

In charge current design, 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 light sensitive area. 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.

Special care must be taken with the photodiode amplification stages. Photocurrents induced by incident light from the stripe are on the order of a nanoamp and must be amplified to reasonable voltage levels. In addition, the high rate of range map acquisition supported by our system implies high bandwidth photodiode signals. A 100 × 100 element sensor gathering 1,000 range images per second requires an amplifier that can provide a gain out to 100 KHz.

Filtering of the photodiode signal at the input to the comparator, as was done in the discrete photodiode system, would have been advantageous. We know that the frequency content of signals generated by a continuously moving stripe will be found above the base scanning frequency. Thus, the lowpass nature of amplified photocurrents should be combined with a highpass filter stage to yield an amplifier with an overall bandpass frequency response. The cells that result will be most sensitive to frequencies generated by the image of a moving stripe. The amount of interference caused by ambient light and signal conditioning circuitry 1/f noise will be reduced.

Thus, a single highpass pole placed to yield a 1.2 KHz corner frequency would have made the cell less sensitive to frequency components found in standard room lighting. However, a filter section built from a passive RC network turns out to be large in chip area. Consider a simple single pole highpass stage built out of a series capacitor and a resistor to ground. The integrated circuit process we are using makes good floating 0.5 fF/m² capacitors available. A 10 fF capacitor would be 140 μm on a side — as big as the photodiodes we are using. However, even with this "large" capacitor, a 13 MΩ resistor is needed to give the desired filter cutoff. Passive resistances of this magnitude are very difficult to fabricate at all on an integrated circuit and, in any case, would require large amounts of chip area.

The desired filter could be constructed in a reasonable die area if we were to employ switched capacitor (SC) techniques [Allen, 1984]. An SC circuit simulates a large valued resistance with an active circuit consisting of switches and a capacitor. We are, in fact, considering the use of SC filter stages for use in future cell implementations. However, any SC design would require that a constantly running digital clock be present. Photodiode anode points are high impedance nodes and will be susceptible to noise coupling from other chip circuitry. Digital circuits switching while stripe detection is taking place have the potential to inject noise and thus corrupt the delicate photocurrent measurements. Though techniques exist which can help alleviate these problems [Olmscheid, 1987], passive analog circuits, like those of the current cell design, do not suffer from this drawback.

Another implementation detail affected by the choice of an unclocked cell is the comparator design. The comparator in each sensing element is non-regenerative – essentially an uncompensated two stage opamp circuit [Gregorian, 1986, pages 168-172] with no applied external feedback. A clocked cell design would allow the use of a regenerative comparator topology. Regenerative comparators use positive feedback when clocked making them faster and more accurate for a given circuit area than the non-regenerative variety. However, again we felt that the potential for injecting clock noise into the photocurrent measurement outweighed the potential benefits.

4.1.3 Timestamp Representation

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. Corruption from effects like these is not an issue for information represented in binary form. We feel, however, that with careful circuit design these errors can be kept to 0.5% or so, yielding useful range values.

The chip area needed for timestamp broadcast and latching will be smaller for the analog scheme when compared with the circuit area required by the digital scheme. An analog timestamp can be broadcast over the entire chip on a single wire and the circuitry to record an analog time value consists of a holding capacitor and a switch. The eight bits of digital timestamp necessary for 0.5% resolution would have to be broadcast over a bus and each sensing element would need a eight bit latch.

In addition to the structures that can be built using a standard p-well CMOS process, the double-poly process we are using provides high quality linear capacitors. These capacitors exhibit good matching across a die and are needed by sensing elements on the range sensor IC to store analog voltages. 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. At capacitive densities of 0.5fF/μm², a 1pF T/H capacitor will be 45 μm on a side.

4.2 Results of Sensing Element Fabrication and Test

A chip with cells which include photodiodes integrated with transimpedance amplifier circuitry has been fabricated and tested. The fabrication was done through the MOSIS [Mosis, 1988] system using a 2μ CMOS p-well double-metal, double-poly process. The size of the photodiode area was 20,000 μm².

The fabricated chip was mounted on the camera of the rangefinder set-up and tested. Figure 9 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 the cells are successfully detecting stripes swept at rates which correspond to 1,000 frames/sec for relatively white surfaces and at 100 frames/sec for very dark surfaces.

4.3 Sensor Array Design, Fabrication, and Chip Interface

A complete sensor chip must include readout circuitry as well as 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.

The first prototype sensor chip, including an array of 10 × 6 sensing elements and read-out circuitry, has been designed, and is in fabrication. A block diagram of this sensor chip, showing sensing element layout, can be seen in figure 10. Total sensing element area is roughly 300 μm × 150 μm for the current design. We used a small die for the initial sensor fabrication, roughly 1/4 cm on a side, and thus were limited to a 10 × 6 element array. We will fabricate a denser array, probably 40 × 30, on a larger die if testing of the prototype chip goes well. Power dissipation becomes a problem as sensing element densities approach 100 × 100. In order to keep the total power used by the chip to a few watts, cell current must be budgeted at around 100 μamps.

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

![Figure 9: Photodiode/Photoamp Oscilloscope Waveforms](image-url)
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 we have to consider is the system interface to control the range image acquisition process. It cycles the sensor chip, retrieves the raw range data, converts it to digital form, possibly translating it from time values into a 3D range points, and makes this data available to a host computer over a high bandwidth path.

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 11. We plan to offload sensing element range data in raster fashion, similar to the readout of CCD imaging chip. For a 40 × 40 (1,600) element sensor, we can spend 500 ns on each cell if offloading time is to be comparable with the target 1 ms range acquisition rate. Future versions of the sensor chip could take advantage of multiple data pathways to reduce the time necessary to dump stored range data. However, each data pathway will require its own charge integration and A/D conversion hardware. Data could also be pipelined with range acquisition if two sets of T/H circuitry were built into each cell. Initially we have decided not to do this for two reasons. First, cell area would grow mainly due to the size of the two additional T/H capacitors needed. Second, time multiplexing of the on-chip busses is essentially a digital process which has the same noise pitfalls as any other digital circuitry on the chip. By separating acquisition and offloading phases, we can ensure that no digital switching will be occurring while range measurements are taking place. In any case, because voltages held on the T/H capacitors will tend to droop, it is to our advantage to offload range samples from the chip as fast as possible.

A high speed analog-to-digital (A/D) converter will be necessary to convert the analog time values into digital form. The A/D converter will need to have a 2 MHz conversion rate and better than ten bits of accuracy. Simple processing here may be used to convert this raw timestamp data into 3D (x, y, z) range map values before presenting it to the host computer system.

5 Conclusion

Advances in VLSI technology make smart sensors possible by integrating sensing and processing. Our goal is to build a VLSI smart sensor based rangefinder which is capable of acquiring 100 to 1,000 frames of range images per second. So far, we have proved the concept by building a small prototype sensor with discrete components. We have designed and fabricated basic smart sensor cells which integrate photodiodes and amplifier circuitry using a monolithic CMOS process. We have demonstrated the capability of the integrated cells to detect the light stripe. Currently a complete prototype chip with 10 × 6 cells is in fabrication.

One of the most distinguishing features of this program is that it is not just parallel implementation of known algorithms by VLSI technology for speed up, such as VLSI chips for convolution. Rather, it demonstrates that integration of sensing and processing can allows for modification of the operational principle of information acquisition (in our case, range imaging) which results in a qualitative improvement in performance.
References


