Abstract

In this paper we present a very fast lightstripe rangefinder based on an IC array of photoreceptor and analog signal processor cells which acquires 1,000 frames of range image per second—two orders of magnitude faster than currently available rangefinding methods. Unlike a conventional lightstripe rangefinder which obtains a frame of range image by the step-and-repeat process of projecting a stripe and grabbing and analyzing a camera image, the VLSI sensor array of this new rangefinder gathers range data in parallel as a scene is swept continuously by a moving stripe. Each cell continuously monitors the output of its photoreceptor, and detects and remembers the time at which it observed the peak incident light intensity during the sweep of the stripe. Prototype rangefinding systems have been built using a 28 × 32 array of these sensing elements.

1 Introduction

Rangefinding, measurement of the three dimensional profile of an object or scene, is a critical component for many robotic applications. Numerous rangefinding techniques have been developed [3], and of these lightstripe rangefinding is one of the most common and reliable methods. In this paper we present a very fast range finder based on an IC array of photosensor and analog signal processor cells, which acquires up to 1,000 frames of lightstripe range data per second.

A conventional lightstripe rangefinder operates in a step-and-repeat manner—a light stripe is projected onto a scene, a video image is acquired, the position of the projected stripe in the image is extracted, the stripe position is stepped, and the process repeats until the entire scene has been scanned. The rate at which video frames of range data can be acquired using this method is limited by the time needed to acquire the video image and process each video image. Typically, on the order of one second is required to acquire a complete range image in this manner.

The VLSI-based rangefinder presented in this paper is based on a modification of the basic lightstripe rangefinding technique described by Araki and Sato [4] and Kida et. al [5]. Imagine that the light stripe is swept continuously across the scene. The time at which the image of the stripe crosses each cell of the range sensor indicates the range along that photodiode's line-of-sight. If all the cells perform this task independently, an entire range map is acquired in parallel during a single continuous sweep of the stripe across the scene.

This rangefinding technique requires cell-parallelism, where each cell detects the timing when the light stripe passes across its photosensor's field of view. Accurate detection of the timing requires a high bandwidth connection between the photosensor and the signal processor that detects the time of the peak. Araki and Sato [4] constructed a small array of discrete photosensors connected to separate peak detection hardware. For large arrays, however, the only practical way to provide a high bandwidth link between each photosensor and the signal processing electronics is to integrate photosensors and signal processing circuitry into a dense array of cells.

In this paper we present the development of a prototype rangefinder using an array of cells, each of which contains a photodiode and the signal processing circuitry needed for the modified lightstripe rangefinding described above. In order to keep the cell area small, the necessary signal processing functions have been implemented using primarily analog circuitry. Prototype chips were fabricated through MOSIS [6] in a 2 μ CMOS p-well double-metal, double-poly process. This design builds on some of the ideas that have been developed for ICs that integrate signal processing circuitry with photosensors; e.g., an optical mouse [7], an artificial retina [8], and an optical position encoder [2].

2 Conventional vs. Cell-Parallel Lightstripe Rangefinder

Fig. 1 shows the geometrical principle on which a lightstripe rangefinder operates. The scene is illuminated with a vertical stripe...
plane of light. The light is intercepted by an object surface and, when seen from 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 using the principle of triangulation. In Fig. 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 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. Both conventional and cell-parallel light stripe rangefinding rely on this principle, but their operations differ slightly.

2.1 Conventional Lightstripe Rangefinding

A conventional lightstripe 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 maximum rate at which frames of range data can be acquired by this technique is severely limited because the number of pictures which must be taken and processed increases linearly with the desired horizontal (i.e., $N$, the number of light stripe positions) resolution. Typically, $N$ ranges between 100 and 200 and time required to acquire and process an image frame ranges between 1/30 second and 1/10 second. Thus, this camera-based conventional rangefinder typically requires between 3 seconds and 20 seconds. Through the use of very high speed cameras or coded multi-stripe patterns, the maximum speed of conventional lightstripe rangefinders has been increased to a few to several frames of range data per second (see [3] for survey).

2.2 Cell-Parallel Lightstripe Rangefinding

In the modified cell-parallel rangefinding technique, the ordinary video camera is replaced by a two-dimensional array of smart cells consisting of photosensitive and signal processing circuits as illustrated in Fig. 2. Range data is not acquired in a step-and-repeat manner. Instead, the plane of light is swept once across the scene from left to right at a constant angular velocity.

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. 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 current cell design implements each of the functions in the simplest possible manner (see in Fig. 4), while still providing the necessary functionality. A photodiode is used rather than a phototransistor in order to maximize the bandwidth. To detect the time at which the light stripe crosses the photodiode, we use a threshold detector which follows the preamp, even though a matched filter should ideally be used before the comparator. Storage of the time information is implemented by a track-and-hold circuit with the analog ramp voltage.

Figure 4: Sensing Element Circuitry

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.

3.1 Integrated Photodiodes and Preamp

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. In the prototype array, the photodiodes are approximately 8,000 \( \mu \text{m}^2 \) in area, one fourth the total area budgeted for a cell. In a CMOS process, one of the highest sensitivity photodiode structures is the well-substrate junction [2]. In a p-well CMOS 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 Fig. 5.

The shape of the photodiode area is also important. Consider the image of a stripe moving across the photodiode as sketched in Fig. 6. The amount of light collected varies because the length of the intersection between the stripe and the photodiode area varies. In the case of a square photocell, the output pulse shape depends on stripe orientation – for a vertical stripe, the output pulse observed will be square, but a 45\(^\circ\) stripe will produce a triangle wave. This difference in pulse shapes translates to an error in estimating the time at which the peak of the stripe occurs. Therefore, the photodiodes are made circular so that the pulse shape will be independent of stripe orientation.

Figure 5: Vertical Photodiode Structure

Finally, the photodiode structures are surrounded with guard rings to help prevent noise currents introduced into the substrate at other points from contaminating the photocurrent, and to prevent photocurrent induced latchup at high light levels.

Figure 7: Photodiode and Transimpedance Preamp

The performance of photodiodes and the transimpedance preamp circuit (see Fig. 7) was tested by a separate test IC which was also fabricated using the same process. The preamp test chip was mounted in the film plane of a standard 35mm camera and a moving stripe was used to generate light stimuli. Fig. 8 shows the experimental results of applying a swept infrared laser stripe across two adjacent cells.
It demonstrates that the signal has a clear peak whose location moves as the distance to the surface changes. The bandwidth of these test sensors is compromised by the connection to a bonding pad. However, the ability of this photodiode/preamplifier combination to successfully detect stripes swept at rates corresponding to 1,000 range frames/second has been experimentally observed for relatively reflective surfaces (higher light input signal level). The maximum rate for very dark surfaces drops to 100 frames/second.

3.2 Stripe Peak Detection and Time Storage

In the prototype array, lightstripe peak detection is done by thresholding. The amplified photodiode signal is compared to a reference voltage, ranging between 2.5-3.0 volts, 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 bonding pads. Detection of the stripe trips the RS flip-flop in the cell, indicating that the lightstripe has been detected and that the timestamp value should be held (see Fig. 4).

A global time signal indicating the position of the lightstripe is broadcast to each cell from an off chip reference as a 0-5 volt analog ramp voltage synchronized with the stripe rotation. Sweeps (ramp duration) of down to 1 msec were employed. 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. The primary reason for choosing analog storage over digital was die area. In a digital implementation, the multi-bit (at least 8 bits) digital timestamp signal bussed over the entire chip and the multi-bit digital latch within each cell require a large amount of the die area. An analog timestamp voltage can be broadcast over the entire chip on a single wire and the circuitry to record an analog time value consists only of a holding capacitor and a switch.

Unfortunately, information stored in analog form is subject to corruption from sources such as switch charge injection and voltage drop. However, with careful circuit design errors due to charge injection, for example, can be kept to about 0.5%, yielding acceptable range accuracy. In the prototype, a double-poly process was used to provide high quality linear capacitors that exhibit good matching across the die. The final layout of a cell is shown in Fig. 9.

3.3 Stored Time Readout

When a scan has been completed, the hold capacitor in each cell of the array is at the voltage corresponding to the time at which the stripe was detected along its line-of-sight. Readout of these stored time values proceeds in a raster scan fashion, like a standard CCD camera except that it is range information, not intensity, which is being read out. A shift register, wound through the array, gates the charge held on each cell's T/H capacitor sequentially onto a single analog readout bus. The accumulated charge is integrated by an on-chip op-amp integrator and presented to an output pin as a buffered output voltage using the circuit shown in Fig. 10. Because the voltages held on the T/H capacitors will tend to drop due to substrate leakage currents, it is important to offload range samples immediately after the data acquisition.
4 Prototype Rangefinder

A prototype rangefinding system has been built that uses the integrated sensor/signal processor arrays described above. Fig. 12 shows the system set-up. Essential components in this system include stripe generation hardware, sensor IC and support electronics, sensor optics, and a host computer interface. The range sensor IC used to be mounted in the film plane of a 35mm camera body which provided a convenient mechanism for incorporating focusing optics and for sighting the rangefinder. Recently we have designed and fabricated a special optical head for the mounting sensor chip which provides a better field of view. The stripe is generated using a 10mW infrared (780nm) laser-diode based projector. A mirror mounted on a galvanometer provides the means of sweeping the stripe. Two scans (one from left-to-right, the other from right-to-left) are generated during each period of the 500Hz triangle wave used to drive the galvanometer, providing the system with a maximum of 1,000 sweeps per second. As an alternative, we also have a spinning mirror to generate sweeping stripes. The experimental system has a 50cm baseline between the stripe generator and the sensor head.

5 Rangefinder Performance

This section presents the experimental results obtained so far to evaluate the performance of the range finder.

5.1 Raw Signals

Fig. 13 show waveforms observed during the operation of a 5 \times 5 range sensor prototype. The middle trace shows the analog timestamp ramp being broadcast to all elements during the range data acquisition phase, at a rate corresponding to approximately 1,000 frames/second. The bottom trace is an index pulse generated by the stripe generation assembly.
that indicates the stripe is positioned at the beginning \((t = 0)\) position. When a scan is completed, range data readout starts.

The top trace shows the output of the op-amp integrator that reflects the charge stored on each cell. The five groups of five pulses that appear are the output from the cells of a \(5 \times 5\) sensor array being read out in a column-wise raster scan. For this experiment, a posterboard (plane) was held inclined vertically at approximately 50cm from the sensor. Referring to Fig. 3, the voltage representing the time stamp \(t\) for each cell is proportional to the stripe angle \(\theta(t)\). Thus, since the read-out integrator inverts the output, low values represent stripe events recorded late in the sweep of the stripe (and close in range): say, for cell \(S_1\), smaller the value \(t_1\) is, the object surface is closer along its line of sight.

5.2 Calibration

Calibration of the parallel-cell range sensor poses a unique challenge. The typical calibration procedure of a camera-based light-stripe range system first determines camera parameters through the analysis of reference object images. Once camera calibration parameters have been found, light-stripe projector geometry is again determined through analysis of reference object images illuminated by the projector at various stripe angles. The current cell-parallel IC-based rangefinder does not provide intensity images of the scene that can be used to aid in the calibration process. For this reason, new algorithms for calibrating the system had to be developed.

The present method for calibrating the cell-parallel IC range system relies on an accurate 3D positioning device for the independent determination of range sensor geometry, together with a reference object. The reference target used in the calibration procedure is a planer surface (posterboard) out of which a triangular section has been removed (Figure 15). As indicated in the figure, the triangular "hole" is manipulated in the field of view of a sensing element so that the line-of-sight of a cell passes the vertex of the triangle. However, it is very difficult and unreliable to detect that moment. Instead, we use the edges. When the line-of-sight of a sensing element changes from being "occluded" to "not occluded" by an edge of the hole, a discontinuity in range measurement is reported by the cell and the position of the target is recorded. The points on the edges of the hole are systematically mapped out by moving the target in the fashion shown in Figure 15. The intersection of the lines resulting from the upper and lower edge points yields one 3D coordinate position on the line-of-sight of the sensor. By moving the target forward in \(Z\) and repeating the process, the coordinates of several points on the sensor cell's line-of-sight ray can be determined.

The stability of such measurements can be shown in the histogram plot of Figure 16. In the figure, each Y-axis value is the number of times (out of 1000 trials, plotted logarithmically) the range sensor yielded a timestamp of that value when the \(Z\)-position to the target was kept constant (timestamp data ranges in value from 0 to 4095). Data from such measurements at several target \(Z\)-positions are plotted. The sharpness of the peaks is an indication of the repeatability of the range data produced by the cell-parallel range system.

Figure 17 shows (projections of) upper and lower edges detected for various \(z\) values. Line fitting to those intersection points results in the line-of-sight equation.
6 Significance of Sensor/Processor Integration

Advances in VLSI technology have made it possible to create smart sensors for high performance rangefinders by integrating sensing and signal processing. Analog signal processing is well suited for performing the kinds of computations necessary to achieve intelligent sensors. This section will discuss the advantages of this approach, and where they come from.

6.1 Speed and Resolution

As noted in section 2, with a conventional camera-based lightstripe rangefinder the number of video images required to build a complete range map increases linearly with the desired spatial (x) resolution. In contrast, for the cell-parallel lightstripe rangefinder the total time of acquisition is independent of the range map resolution. If the geometry of the optics and the speed of the lightstripe remain fixed, increasing the size of the sensor array just increases the spatial (x – y) resolution or the visual angle that is imaged.

6.2 Accuracy of the Range Measurement

The cell-parallel lightstripe rangefinding method actually increases the accuracy of the range measurement. As illustrated in Fig. 18, the accuracy of the conventional camera-based method is limited by the number of pixels in the horizontal scan line (256 to 512 for a typical camera) because the range information is derived from the position of the peak of the light intensity profile on the scan line. Subpixel peak localization techniques can at best provide one tenth of a pixel spacing accuracy. It is interesting to note, however, 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 cell-parallel method with analog circuitry, on the other hand, the peak of the continuous uninterpolated time profile of intensity from the same cell can be detected in the time domain with much greater accuracy: even when the time profile of intensity for a cell photodiode becomes square, it is still reasonable to determine the middle point as the time when the light stripe passes the cell, as long as the surface has a locally constant orientation. 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.

6.3 Space- vs. Time-Domain Operation

Though both conventional and cell-parallel lightstripe rangefinding techniques share the same optical and geometrical principle, their operational principles are actually very different. Fig. 19 explains this fact. Imagine a static scene and a single scan line along the spatial x axis. Brightness along the scan line has a particular pattern, and if no
active illumination were applied to the scene, the intensity pattern would remain the same over the time. So, the spatial-temporal intensity function \( I(x, t) \) would be constant along the time axis \( t \). When a light stripe is swept over the scene, the pattern \( I(x, t) \) will include an added component due to the light stripe, as indicated in the figure by a ridge moving diagonally across the \( x - t \) plane.

In the conventional technique, range data is extracted by analyzing an intensity pattern along a scan line of the video image. We can interpret this as slicing the spatial-temporal intensity function \( I(x, t) \) along a fixed \( y \) and detecting the light stripe timing \( t \) in \( I(x, y, t) \), as shown. In this sense, the conventional technique is space-based, whereas the cell-parallel technique is time-based. As was demonstrated by the developed prototype sensor, the latter technique allows for much faster range-finding. Moreover, we can see that it is fundamentally easier to detect a peak in the time domain, since its signal contains only one peak with constant background—except cases for split pixels (a cell’s field of view covers two surfaces of different depth) or multiple reflections. It should be noted that this modification of the operational principle was made possible by the use of smart sensors which integrate sensing and signal processing.

7 Conclusion

We have demonstrated a VLSI smart sensor based on light-stripe range-finding which is capable of acquiring 100 to 1,000 frames of range data per second. One of the most distinguishing features of our integrated sensor is that it is not just a parallel implementation of known algorithms using VLSI technology to achieve increased speed or compactness. Rather, it demonstrates that integration of sensing and signal processing can make it possible to modify the operating principles of the information acquisition method (in our case, light-stripe range-finding) which can result in a qualitative improvement in performance.

References