Paper ID R28

Smart pixels for future 3D-TOF sensors

M. Lehmann, T. Oggier, B. Büttgen, Chr. Gimkiewicz, M. Schweizer, R. Kaufmann, F. Lustenberger[1], and N. Blanc
CSEM SA, Photonics Division, Badenerstrasse 569, CH-8048 Zurich, Switzerland
[1]phone: +41 44 497 1444; fax: +41 44 497 1400; email: felix.lustenberger@csem.ch

Abstract

The Time-of-Flight (TOF) principle is well-known and is suitable for determining the distance of objects. By implementing all the circuits on a single chip, the technology can be extended from single-point distance measurements to imaging surveillance, where the 3D-TOF sensor provides a 3D representation of the surrounding scenery in real-time. One of the main challenges of such an imaging architecture is that the sensor has to deal simultaneously with a large signal dynamic range and additional background illumination. In this article, two different pixel architectures are presented, which overcome these limitations. In addition to the functional description of these new pixel architectures, measurement results are provided showing background illumination suppression up to full sunlight intensity.

1. Overview

The phase-measurement principle using actively modulated light waves as an information carrier is well suited for mid-range applications at distances of up to 10 meters: The wave front that is diffusely reflected from the scene is demodulated in the receiving pixel matrix and the phase delay is measured, hence providing direct distance information of the entire scene (see Figure 1). With the approach chosen in [1], the photo-electric demodulator consists of a small CCD element integrated within each pixel of the smart 3D-TOF sensor. The read-out circuitry is similar to conventional CMOS active pixel sensor (APS) imagers. These basic principles, complemented by advanced approaches in information processing, help to improve the dynamic range and to reduce the effect of background illumination.

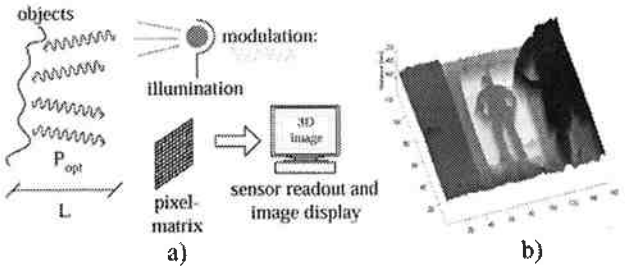

Figure 1: a) System overview and b) B/W-coded 3D image

In this paper, section two explains the TOF measurement principle and the formulae used to calculate the distance values and then, section three presents details of the system implementation. Sections four and five describe two novel pixel architectures, which have been implemented and thoroughly tested. They both overcome limitations in the dynamic range of a standard pixel. Section six compares the 2 new architectures directly. Finally, the various fields of application for which the pixels can be used in the very near future are proposed and the article is concluded.

2. The Time-of-Flight Principle

For distance measurements utilizing the TOF principle, an active illumination source emits sinusoidally modulated light waves normally in the non-visible near infrared spectral range. These waves travel with the speed of light. They are diffusely reflected at the objects and some of the optical power finally returns to the detector (see Figure 2). The detector determines the phase difference between the outgoing light wave and the reflected one, which is proportional to the distance of the object (equation 2).

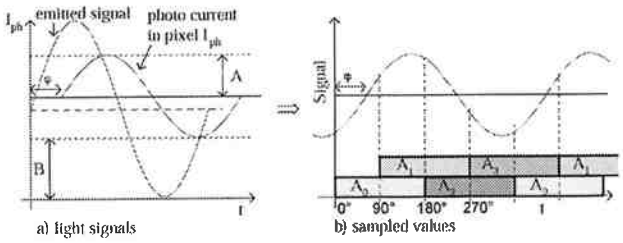

Figure 2: Optical waveforms and sampling process
φ: phase delay; B: background signal; A: amplitude

Several different approaches for phase measurement exist today [2]. In the present case, using simple CCD structures, the photo-generated electrons are separated and finally stored in four different so-called taps. The separation is controlled by a sampling signal synchronized with the modulation frequency. The collection of the photo-generated electrons for one tap is performed for half of the modulation period; the sampling period is shifted by 90° with respect to the modulation signal for each of the other taps. Although only a few electrons are collected per cycle, the signal becomes large enough during the accumulation over thousands of cycles and eventually surpasses the electronic noise floor of the circuitry. Hence, by applying the almost noise-free integration process using optimized CCD structures, the signal-to-noise ratio (SNR) is drastically enhanced.

Bearing in mind the basic charge-separation principle described above, it is clear that the sensor provides four output tap values A_0 to A_3 (see Figure 2). The phase shift of the incoming light with respect to the transmitted light wave and thus the distance D of the object can be calculated by equations (1) and (2):

$$\varphi = \arctan\left(\frac{A_1 - A_3}{A_0 - A_2}\right) \quad (1); \qquad D = \frac{\varphi}{2\pi} \cdot \frac{c}{2f}, \quad (2)$$

where φ is the phase delay, D the distance between the object and the camera system, c is the speed of light and f the modulation frequency (actually 20 MHz in the present examples). The theoretical background can be found in more detail in references [1] and [3].

3. CCD-based Demodulation Pixel

As described in the previous section, the sampling of the four taps is done by a small CCD element in each pixel. A 0.6µm CMOS process with two poly and three metal layers is currently used which also provides a CCD option with overlapping poly gates and a buried-channel. This CCD element is responsible for the fast charge separation and hence the demodulation of the incoming light.

3.1. Demodulation and integration

A cross section of the basic CCD element is shown in Figure 3. It is built up symmetrically with the main photoactive area, i.e., the photo gate (*pgm*) in the middle. Toggle gates (*pgl* and *pgr*) are located on both sides. During the demodulation process, their potentials are switched between a high and a low value synchronously to the modulated light. Together with *pgm*, they build staircases in the potential distribution and thus the charge is separated by relying on drift and diffusion processes. These photo-generated charge carriers are collected under the integration gates (*intg*) next to the two toggle gates on both sides. An output gate (*outg*) separates the *intg* from the sense node of the regular APS architecture.

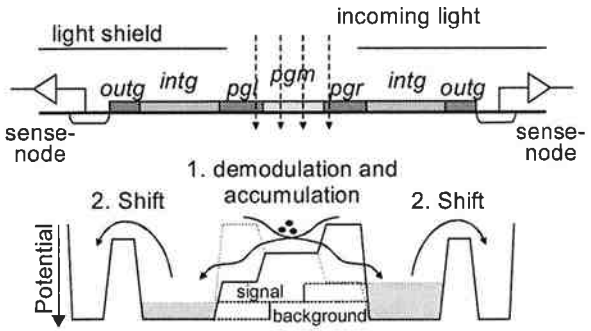

Figure 3: CCD element and its functional principle

When the *integration phase* has finished, the *shift phase* starts. The photo gates (*pgl*, *pgm*, *pgr*) are all set to a low voltage, such that they build up a potential barrier. Then, the integration gate potential is lowered and the collected electrons flow over the *outg* to the sense node, decreasing the potential of this preset node (see also Figure 4). The sense node is then read out through a source follower similar to conventional three-transistor pixel APS imagers.

3.2. Two-tap and four-tap architectures

The basic circuit described above can only accumulate two opposing taps at once, i.e., A_0 and A_2 or A_1 and A_3. The remaining two taps have to be acquired either in a second integration period, where *pgl* and *pgr* are phase shifted by 90°, or through a second CCD element in the pixel with the toggle gates also shifted by 90° with respect to the original illuminating wave front. The first solution is called a 2-tap sensor; whereas the second solution is called a 4-tap, or more accurately a double 2-tap sensor.

3.3. Dynamic range

The main challenge of this kind of ranging sensor is the enormous intra-scene dynamic range (DR) that has to be supported. A simple estimation shows the problem: Due to an active illumination source, the received light intensity diminishes with the square of the object distance, i.e., from 0.5 m to 7.5 m by a factor of $15^2 = 225$ (equal to an attenuation of 47 dB). Additionally, the objects may exhibit very different reflection coefficients of, e.g., 5 % to 95 %, providing another factor of about 20 (equal to 26 dB). Furthermore, to reach a minimum distance resolution of, e.g., 10 cm, the signal must have a SNR of at least 25 dB, resulting in a minimum total dynamic range of 98 dB.

Despite the active light source, the background illumination also degrades the SNR and calls for a high DR. In a first approximation, the background signal can be assumed to be constant. During one integration cycle, an equal amount of photo-generated electrons is integrated on each tap, thereby generating a common mode signal on these taps. As shown in Figure 2 b), the background signal can easily exceed the signal amplitude. Nevertheless, the sensor has to detect the signal amplitude out of the collected common mode signal.

To improve the dynamic range, one solution is to grab the image with various different integration times. An algorithm then chooses the most suitable result for each pixel. However, when using several integration cycles, only a low frame rate can be achieved. The following pixel concepts have been designed to overcome these limitations: the minimum charge-transfer pixel tries to suppress the common mode signal already at the pixel-level [4] whereas the pixel-wise integration approach optimizes the integration time for each pixel.

4. Minimum Charge Transfer (MCT)

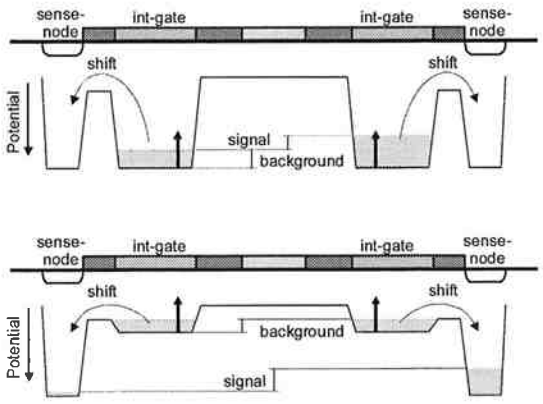

Figure 4: Shift stops in MCT pixel as soon as all taps receive charge, common mode charge remains under *intg*

The charge packet collected underneath the integration gate always consists of the common mode part and the signal part. However, only the signal part is useful for distance calculation. The idea behind the MCT

functionality is to shift only the relevant signal part of the accumulated charge carriers from under the integration gate to the sense node.

This selective transfer of charge carriers is performed by lowering the *intg* potential with a given slew rate, while a monitoring circuit controls the outputs of the pixels. As soon as all taps become lower than a certain threshold voltage, then the transfer is stopped since the minimum amount of charge has reached each tap. The *intg* potential is then frozen and no further transfer takes place. The remaining charge carriers underneath the integration gates is equal for all taps, so this background charge is suppressed from flowing over the *outg* barrier into the sense node (see Figure 4).

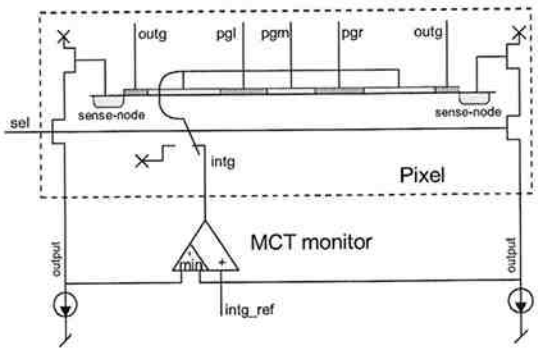

Figure 5: Pixel with MCT monitoring circuit

In order to preserve a high optical fill factor, the task of checking the levels is done column-wise outside of the pixel field (see Figure 5). In the presently selected row, the *intg* potential is switched from a high state to the output of the MCT monitoring circuit, whereas all of the *intg* potentials in the other rows remain high. Then the monitoring circuit lowers the *intg* potential until the threshold is reached. After this MCT phase, the outputs can be read out conventionally through the regular APS circuitry.

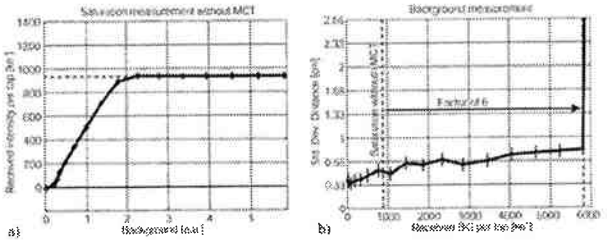

Figure 6: Measurement results on the MCT pixel

The measurement results on the MCT pixel show the ability to handle much more background signal than the sense node can store. In Figure 6 a), the MCT option was switched off and the background illumination has been increased. The taps in the pixel exhibit saturation at a level of approximately 1 Me$^-$, which corresponds to the full well capacity of the sense node. When the MCT mode is switched on, the background illumination could be increased to a level of almost 6 Me$^-$. The background charge handling capability is increased by more than a factor of 6 (see Figure 6 b).

5. Pixel-Wise Integration (PWI)

The second topology proposed within this article adapts the integration time on a pixel-by-pixel basis. In this approach, the entire integration time is divided into small steps. After each integration step, a charge-shift operation is performed until a comparator decides that enough photo-generated charge carriers have been collected. Afterwards, any further shifting operation is suppressed in order to avoid saturation of the sense node. The comparator actually detects if one of the taps of a pixel has obtained more than half of the allowed signal swing (see Figure 7). If the pixel taps are all below this level, then the integration time of the next step is doubled, gaining 6dB of dynamic range at each repetition. Each individual pixel has therefore the optimum integration time, and hence, the dynamic range of the sensor is dramatically increased.

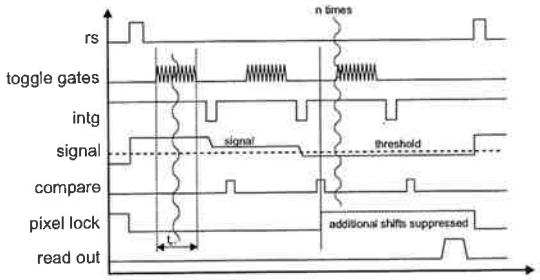

Figure 7: Timing diagram of the PWI pixel

Since each pixel requires its own optimum integration time, the comparator has to be included inside each pixel (see Figure 8). Furthermore, because the integration time must be the same for all taps, only 4-tap pixels are suitable for this approach. The comparator and the storage element have to be as simple as possible in order to still provide a tolerable optical fill factor. Since the comparator is directly checking the sense node level, the pixel must be laid out very carefully with respect to parasitic components on these nodes.

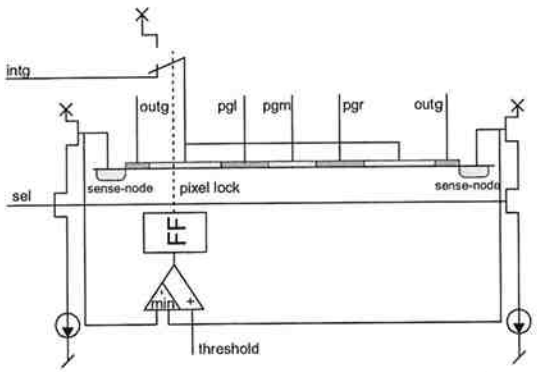

Figure 8: Block diagram of PWI pixel; only one of the two CCD elements of the double 2-tap pixel is drawn

The potentially achievable integration times are indicated in the measurement results (see Figure 9). The intensity level, i.e., the mean value on all four taps, has a saw-tooth shape when measuring with varying distances and hence with varying illumination levels. Every time the pixel switches from one integration time to the next,

the intensity curve makes a step; nevertheless, the calculated distance values remain linear.

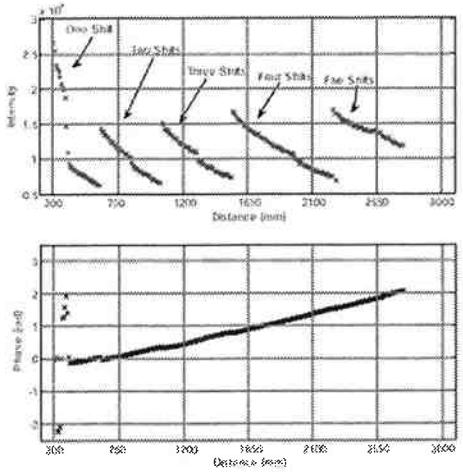

Figure 9: Measurements on the PWI pixel reveal the different integration times in the intensity values (upper diagram)

6. Comparison between MCT and PWI

The two pixel architectures with the minimum charge-transfer and pixel-wise integration approaches have been compared in depth. Their ability to suppress the background illumination and their achieved dynamic range has been measured and contrasted. Some important parameters of the MCT and PWI pixels that were realized on a test chip are listed in Table 1.

Table 1: Comparison of MCT and PWI pixel

Parameter	MCT pixel	PWI pixel
Topology	double 2-tap	double 2-tap
Pixel pitch	60µm x 70µm	60µm x 90µm
Optical fill factor	11%	8%
Full well	0.95 Me^-	1 Me^-
Charge handling at *intg*	5.8 Me^-	N/A

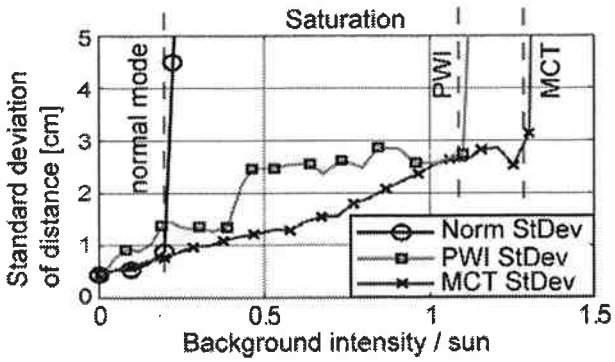

Figure 10: Comparison of the MCT and PWI pixels with fixed target and integration times. The background is normalized to the sun's illumination equivalent.

In Figure 10, the background illumination was increased from zero to the equivalent intensity of the sun on a bright day. With higher background illumination, the photon shot noise also disturbs the accuracy of the measurement. Instead of saturating at about 0.3 times the sun's brightness as a normal pixel does, both the MCT and PWI pixel can still deliver reliable distance values with good accuracy up to a background illumination power level larger than the equivalent of the sun.

7. Discussion and Conclusions

As the measurement results already indicate, the MCT pixel has significant advantages in the background light suppression, whereas the PWI pixel, due to its automatic adaptation of the integration time achieves a larger distance measurement range. Therefore, the PWI technology may be used advantageously in applications where no or only a small amount of background light is present. However, the distance measurement range has to be large. Examples can be found in mobile in-door robots or machine surveillance systems. On the other hand, the MCT technology is more practicable in outdoor situations where the potential interference by the sun or other bright light sources can occur.

This article provided a brief overview of the 3D-TOF technology and its current limitations. Two novel 3D-TOF pixel architectures were presented which overcome these limitations. The functional principles as well as measurement results were presented; both demonstrate the improved performance of the advanced 3D-TOF pixels with respect to their achievable dynamic range.

Using such advanced pixel architectures, the 3D time-of-flight technology will further improve and open up new applications, such as outdoor environments where the illumination is hardly controllable.

8. Acknowledgment

This work was supported by the Swiss Federal Commission for Technology and Innovation (CTI) and by the strategic industrial partners IEE, Luxembourg and CEDES AG, Landquart, Switzerland.

9. References

[1] T. Oggier et al., "An all-solid-state optical range camera for 3D-real-time imaging with sub-centimeter depth-resolution (SwissRanger)", Proc. SPIE Vol. 5249, pp. 534-545, 2003

[2] R. Lange et al., "Demodulation pixels in CCD and CMOS technologies for time-of-flight ranging", Proc. SPIE Vol.3965A, pp. 177-188, 2000

[3] B. Büttgen et al., "Demonstration of a Novel Drift Field Pixel Structure for the Demodulation of Modulated Light Waves with Application in Three-Dimensional Image Capture", Proc. SPIE Vol. 5302, pp. 9-20, 2004

[4] R. Kaufmann et al., "Novel Pixel Architecture with Inherent Background Suppression for 3D Time-of-Flight Imaging", Proc. SPIE Vol. 5665, pp. 1-8, 2004