# 3D-Imaging in Real-Time with Miniaturized Optical Range Camera

T. Oggier*, R. Kaufmann*, M. Lehmann*, P. Metzler*, G. Lang*, M. Schweizer*, M. Richter*, B. Büttgen*, N. Blanc*
K. Griesbach+, B. Uhlmann+, K.-H. Stegemann+, C. Ellmers+

*CSEM SA ; Centre suisse d'électronique et de microtechnique SA, Badenerstrasse 569, 8048 Zurich
+ZMD, Zentrum für Mikroelektronik Dresden, Grenzstrasse 28, 01109 Dresden

## *Abstract*

The ability to capture the entire environment in three dimensions represents completely new possibilities for many kinds of applications. The availability of the scene's distance information in real-time renders algorithms much less complex and, for many applications, allows problems to be approached in a new, cost-efficient way or even new solutions to be found. A new miniaturized 3D-camera, SwissRanger$^{TM}$, will be presented.
The SwissRanger camera is based on the time-of-flight principle. The emitted signal is amplitude modulated in time, reflected by the objects in the scene and backscattered onto the sensor. Each pixel on the sensor demodulates the incoming light signal and recovers the sine wave function. The phase delay of the recovered signal compared to the emitted signal is directly proportional to the distance from the captured object to the camera; the offset represents the conventional black-and-white image and finally the amplitude is a measure of the quality of the acquired distance information.
An overview of the theoretical background of the phase-delay distance measurement will be given. The distance, offset and amplitude calculation using four sampled integrations will be derived.

The demodulation principle of the SwissRanger is based on the so-called 2-tap architecture. Every pixel contains two different storage locations. Therefore, this procedure allows the two opposing samples to be integrated simultaneously. In the applied algorithm, a total of four samples are captured to rebuild, reliably, the incoming signal characteristics - namely the phase, the amplitude and the offset. A closer look will be given on the 2-pixel architecture. This includes the presentation of simulation results using the ISE-TCAD tool. The simulations are performed to determine the precise geometries as well as the optimal doping profile. The main criterion of the simulation is given by the demodulation ability of the pixel. Different parameters are taken into account, such as the driving voltage of the gates and the wavelength of the NIR-light source. Performance simulations for different buried channels will be discussed and analyzed.
The different components of the SwissRanger and their functionalities within the system will be briefly described, before presenting the test results achieved. The accuracy of the distance measurement will be plotted as a function of a given target's distance. Different sensor characteristics, such as storage capacity and demodulation contrast, will be presented. It will be proven that accuracies in the sub-centimeter range can be achieved.

## 1. Introduction / Technological Challenge

The ability to capture the environment in three dimensions brings tremendous new opportunities in several application fields. Our approach is based on a time-of-flight (TOF) distance measuring principle. Other approaches, such as triangulation methods (e.g. stereo vision), interferometric systems and radar systems, are studied in detail in various papers. Due to the rapid progress in microtechnologies, mainly in microelectronics, a system using the TOF-principle to capture entire environments has become increasingly precise and, nowadays, can even fulfill the specifications of the first industrial applications. Systems based on the TOF-approach profit mainly from the tremendous progress in microtechnology.
The main difficulty, and the reason that the TOF-system needs even further progress in microtechnology, is the fact that measuring distances using the speed of light requires very fast and precise devices. As an example, by using a counter with a clock frequency of 1 GHz, a distance resolution of at best $\Delta L = c/(2\cdot F) = 15$ cm can be achieved.

* Send correspondence to Thierry Oggier, e-mail: thierry.oggier@csem.ch, phone: +41 1 497 1415, fax: +41 1 497 1400

This example shows the high demands placed on time of flight measurements.
By implementing more sophisticated approaches than a simple counter, the requirements on the can be lowered and thus higher distance resolution can be achieved.
In this paper, we present the approach developed at CSEM, based on the phase measurement of an incoming wave front [1]. An intensity-modulated signal is emitted by the system, is reflected by the objects in the scene and returns to the phase measuring detector. The phase of the incoming wave represents a direct measure of an object's distance in the scene. Theoretical background information is presented and the physical limits to distance resolution using the phase measurement algorithm are sketched. These estimates again show the need for further improvements in microelectronics.
The state-of-the-art of the integrated circuit design of such a phase measuring system is presented in detail. Different simulations and models are described, whereby the importance of a very close collaboration between design house and foundry is emphasized.
The system design of the complete distance-measuring device is presented. The measurements show, that such a TOF system reach performances that are appropriate for use in industrial markets.
Finally, an overview is given of anticipated improvements in TOF-cameras arising from further advances in microtechnology.

## 2. Theoretical Background of the Phase Measurement Technique

The camera system emits a sinusoidally modulated light wave. This light impinges on different objects in the scene and is reflected back to the camera system. Through an appropriate optical system, the electromagnetic wave front is imaged onto a demodulating sensor. The emitted light can be modeled as ideally sinusoidal:

$$e(t) = e \cdot \left[1 + \sin\left(\frac{F}{2\pi} \cdot t\right)\right] \tag{1}$$

*F: Modulation frequency*
*e: emitted mean power*

The wave front impinging on the sensor appears as:

$$s(t) = BG(t) + e \cdot k \cdot \left[1 + \sin\left(\frac{F}{2\pi} \cdot t - \phi\right)\right] \tag{2}$$

*BG(t): Background illumination power*
*k: Attenuation factor including target (distance, reflectivity) as well as the optics (lens, filter)*
*$\varphi$: Phase delay arising from the object's distance*

In conventional environments, the background illumination changes with much lower frequencies than the emitted light and, therefore, can be approximated by a constant BG(t) = BG.
If we demodulate the incoming signal by sampling four times ($A_0$, $A_1$, $A_2$, $A_3$), each sample delayed by a quarter period of the modulation frequency, the incoming signal can be computed according to the following formulae:

$$\varphi = \operatorname{atan}\left(\frac{A_3 - A_1}{A_0 - A_2}\right) \tag{3}$$

$$B = \frac{A_0 + A_1 + A_2 + A_3}{4} \tag{4}$$

$$A = \frac{\sqrt{[A_3 - A_1]^2 + [A_0 - A_2]^2}}{2} \tag{5}$$

*$\varphi$: Measured phase delay*
*B: Measured offset*
*A: Measured amplitude*

These three parameters allow the incoming signal to be completely recovered. The offset B represents the intensity information, as is obtained by conventional black and white imaging. The amplitude A is a mean of quantifying the incoming light derived from the illumination unit. Finally, the distance from the target can be calculated directly by equation 6:

$$L = \frac{L_0}{2\pi} \cdot \varphi \qquad (6)$$

$L_0$ represents the non-ambiguity range of the phase measurement. Because the distance accuracy of the measurement is dependent on the acquired amplitude and offset, these two parameters provide a very fast and cost-efficient way qualifying the distance measurements performed.
The ultimate physical limitations of such a distance measuring device are described in more detail in [2] and [3].

## 3. System Description

The 3D SwissRanger camera has a size of 14.5 cm x 4 cm x 3 cm and can be connected to any PC equipped with a standard USB interface. Both USB 1.1 and 2.0 are supported. Apart from the custom designed sensor IC itself, all components of the camera system are commercially available.

*Figure 1: Picture of the SwissRanger™ 3D-camera.*

The camera is composed of the illumination unit and the control board with the sensor. The illumination unit consists of 48 LEDs, emitting a sinusoidal wave front with a total mean power of 800 mW. The wavelength of the emitter is in the NIR, viz. 870nm. The light is modulated with a frequency of 20 MHz, resulting in a non-ambiguity range of 7.5 m.
If large distances or very high frame rates are required for a given application, the illumination unit can be modularly expanded.
The core of the control board is the sensor, which has been manufactured in a 0.8 µm CMOS/CCD technology from ZMD. The CCD option renders a noise-optimized photon detection, whereas the CMOS circuitry provides for complete readout flexibility of the sensor. The sensor has been optimized using ISE-TCAD tools for both process and device simulation during each step of the manufacturing process. The technology and the simulation methods are described in detail in section 4.
Each pixel is based on the so-called two-tap architecture. This means that each pixel contains two storage capacitances and two analog outputs. The sensor is readout at 10 MSamples/s.
The two analog outputs of the sensor are digitized by a 12-bit A/D converter. This data is then passed to an FPGA, where the data is processed and buffered in a RAM. The FPGA controls the complete camera, generating the different sensor control signals by either driving the signals or addressing the 8-bit D/A converter. The initial data processing is also performed by the FPGA. It reconstructs the incoming light wave by calculating the phase delay (distance), the offset (intensity) and the amplitude. Various parameters influencing the data processing can be directly set through the USB port. For example, the integration time can be adjusted in real-time according to the intensity output, signal thresholding can be performed and regions of interest can be chosen and altered during operation of the camera. This flexibility allows customers easy access to the main parameters and enables them to adjust the camera to their specific application.

## 4. Sensor Modeling / Simulation / Technology

The integrated 3D-camera chip was manufactured using a modular standard 0.8 µm CMOS process for 3 / 5 V with CCD option. The modular process integration concept guarantees a cost effective, flexible and reliable solution of sensor function, electronic control and readout circuitry. There is a possibility to mix and match all analog options without affecting performance and parameters of standard devices and CCD. Among the optional devices that can be incorporated are low and high voltage transistors (NMOS, PMOS), bipolar transistors, linear resistors and poly-poly capacitors (see figure 2). The CCD option itself offers a wide spectra of optical sensor applications.

The aim of a CCD is to store photo-generated charges in potential wells and to transport the charge packets through the CCD. Its common to realize a CCD by a sequence of overlapping poly gates with minimal space.

The ZMD CMOS/CCD process uses two poly silicon layers for CCD gates, a poly 2 / poly 1 overlap, a small poly gate space as well as a good poly / poly isolation. One CCD gate is a Poly 1 gate at the 33 nm gate oxide that is normally used for HV transistors. The other CCD gate is realised using a Poly 2 gate at a split gate oxide (33 nm) consisting of gate oxide 2a (only for CCD) and the normal LV gate oxide (for LV transistors). The growing of the two gate oxides lead to a Poly 1 oxide of about 70 nm which guarantee a reliable poly / poly isolation for CCD gate voltages up to 10V. Beside this the poly 1 oxide represents the small space between the Poly1- and Poly 2 CCD gates (see figure 3).

In order to improve the CCD performance in terms of quantum efficiency, speed and demodulation contrast there is a buried channel implant option. The advantage of the BCCD is a maximal value of fringing fields to speed up the charge transport and the absence of any interaction of charge carriers with the interface states, thus minimizing the transfer inefficiency. On the other hand, compared to surface-channel CCD, the BCCD can handle only about one-third of a charge packet.

The buried channel implant as well as the substrate doping level were determined and optimised for the 3D-camera by extensive simulations utilizing the ISE-TCAD tools. It could be shown that the buried channel indeed increases the ability to separate and detect the photo-generated electrons, while improving the performance of the sensor in terms of noise and sensitivity. Figure 4 illustrates a simulation result of a CCD-gate structure as implemented in the SwissRanger$^{TM}$ 3D sensor. In this specific example, the electrostatic potential in the silicon is displayed for selected voltages applied to the CCD gates. The potential distribution within the silicon is shown using a grayscale representation.

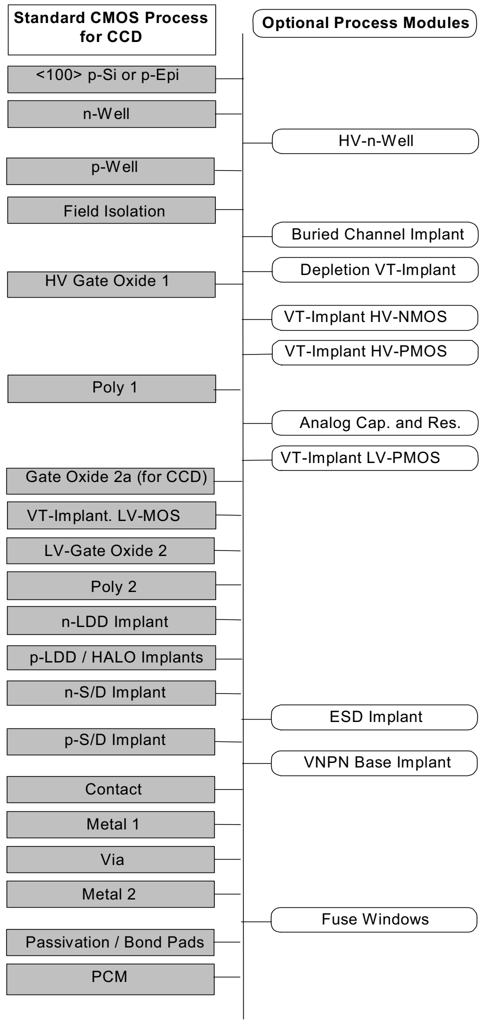



*Figure 2: CMOS/CCD modular process flow*

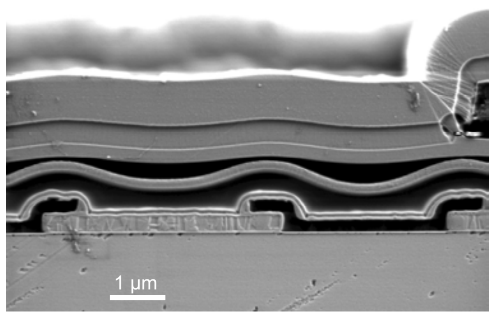



*Figure 3: SEM cross section of a CCD*

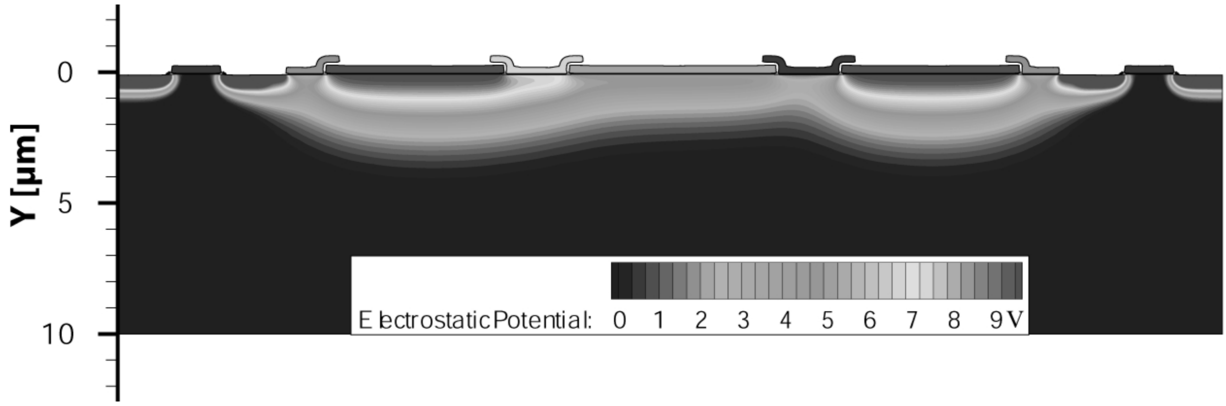



*Figure 4: Example of an ISE-TCAD electrostatic potential simulation*

## 5. Tests Results Discussion

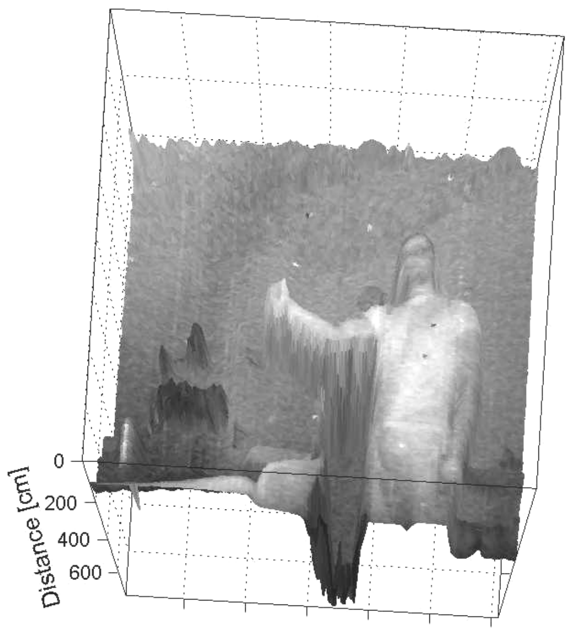



*Figure 5: Three-dimensional represented scene with intensity overlaid picture.*

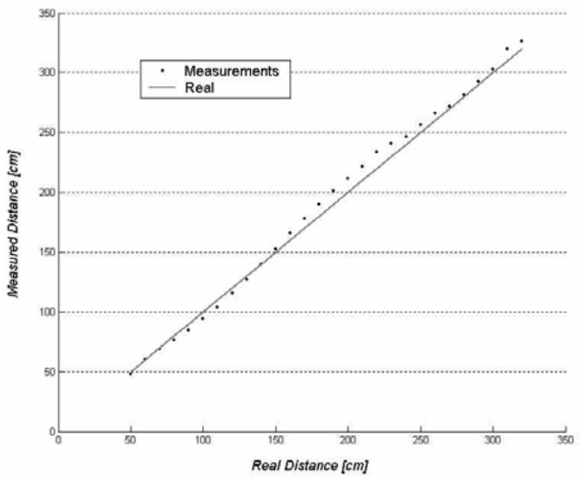
Figure 5 shows a sample 3D-image acquired with the SwissRanger camera. The measured distance map is represented in x-y-z coordinates, where z represents the measured distance, x and y coordinates are directly given by the sensor, the pixel dimensions and the optics used.

The distance scale starts at 0 cm and reaches the non-ambiguity range of 750 cm. As shown in section 2, “Theoretical Background of the Phase Measurement Technique”, the intensity can be calculated simultaneously with the distance. In the 3D-map, the intensity information is overlaid on the distance map. The third item of reconstructed data, the amplitude, is not taken into account in Figure 2, since no amplitude filtering has been performed.

In order to characterize the camera more specifically, an automated distance-measuring test is carried out. A target hangs on a linear rail and can be moved forwards and backwards with sub-millimeter resolution. This target mover is programmed in such a way that a distance measurement evaluation is performed every 10 cm, starting at 50 cm and reaching 320 cm. The target itself is a white board. The camera captures 200 images, with an integration time of 2.55 ms and the linearity of the central pixel of the sensor is examined. In Figure 6, the measured distance value is plotted against the real distance. The x-axis represents the real distance in centimeters, the y-axis the measured distance. The line plots the ideal distance and the dots represent the measurements. Apart from the distance calculation, no additional calibrations are performed.

In order to compute the depth resolution, the standard deviation of the 200 sequential measurements is calculated. The evaluation has been performed for a central pixel field of the sensor consisting of 10x10 pixels.

The results are plotted in Figure 7. The x-axis represents the target’s distance, the y-axis the standard deviation of the 200 measurements. For each distance, 100 values are plotted, corresponding to the evaluated 10x10 pixel field. Figure 7 shows that distance resolutions in the sub-centimeter range can be obtained.



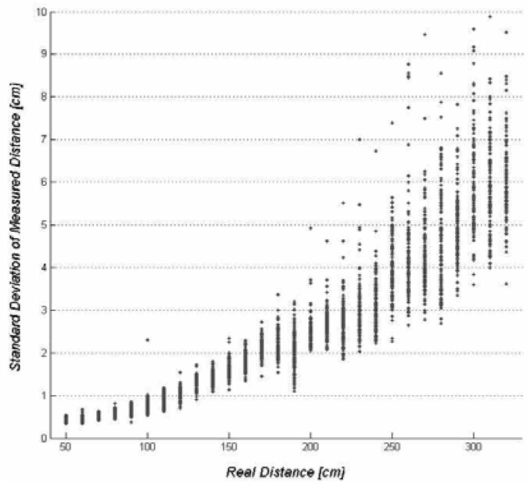




*Figure 6: Linearity measurements of the SwissRanger camera (line = real distance, dots = measured distances).*

*Figure 7: Standard deviation for 100 pixels as a function of the target distance (integration time = 2.55 ms)*

## 6. Outlook / Acknowledgements

In this work, the 3D TOF SwissRanger camera with a sensor resolution of 124 x 160 pixels has been presented. It has been shown that in laboratory environment depth resolutions in the sub-centimeter range can be achieved.
A more aggressive technology would allow the integration of more functionality into each pixel or an increase in the sensor resolution. ZMD is currently preparing an optimized 0.6 µm CMOS/CCD process, into which TOF-sensors will be integrated, enabling them to profit from the design options of e.g. stacked vias or an additional metal layer. At the same time, the development of a complete system-on-chip (SOC) is foreseen. The A/D converter and first algorithm are intended to be integrated into a one-chip solution.
New pixel structures [4] are in development, targeting a four-tap pixel architecture (synchronous integration of the four samples) and a denser pixel design.
Further improvements in sensor sensitivity can be achieved by the use of microlenses, which may be used to increase the fill factor.

The authors would like to acknowledge everyone participating in the fruitful cooperation between ZMD and CSEM.

### *References*

[1] R. Lange, P. Seitz, "Solid State Time-Of-Flight Range Camera", IEEE Journal of Quantum Electronics, 37 (2001) 390

[2] T. Oggier et al., “An all-solid-state optical range camera for 3D real-time imaging with sub-centimeter depth resolution (SwissRanger™)”, SPIE Proc Vol. 5249-65, St. Etienne (2003)

[3] R. Lange, P. Seitz, A. Biber, and S. Lauxtermann, “Demodulation pixels in CCD and CMOS technologies for time-of-flight ranging”, Proceedings of the SPIE, Vol. 3965A, pp. 177-188, San Jose, (2000).

[4] B. Büttgen, T. Oggier, R. Kaufmann, P. Seitz, N. Blanc, „Demonstration of a novel drift field pixel structure for the demodulation of modulated light waves with application in three-dimensional image capture”, Proceedings of the SPIE Vol. 5302 No. 2, San Jose (2004).

輸入販売元：有限会社日本クラビス
〒275-0012　千葉県習志野市本大久保5-2-9
(Tel)047-403-1678 (Fax)047-403-1655
Web : www.j-clavis.co.jp e-mail : sales@j-clavis.co.jp