In Proceedings of the 2011 IEEE International Electron Devices Meeting
December 5-8, 2011 Washington DC USA

A 192×108 pixel ToF-3D image sensor with single-tap concentric-gate demodulation pixels in 0.13 µm technology

T.Y. Lee, Y.J. Lee, D.K. Min, S.H. Lee, W.H. Kim, S.H. Kim, J.K. Jung, I. Ovsiannikov†, Y.G. Jin, Y.D. Park, E.R. Fossum, and C.H. Chung

Semiconductor R&D Center, Samsung Electronics, Hwasung, South Korea,
†Samsung Semiconductor Inc., Samsung Electronics, Pasadena, USA
TEL:+82-31-209-6352, FAX:+82-31-209-6300, E-mail:ty.lee@samsung.com

Abstract

A single-tap concentric photogate pixel of 28 µm pitch is developed for application to time-of-flight (ToF) three dimension (3D) image sensors. The 198×108 ToF pixel array exhibits demodulation contrast higher than 50% and distance error less than 1%, over 1 to 7 m range using 20MHz modulation of 850 nm light emitting diode (LED) illumination.

Introduction

Over the past 10 years, ToF-3D image sensors using high frequency modulation of near-infrared light (*e.g.* 850 nm at 20MHz speed) have emerged as a viable alternative to stereo and structured light imaging for capturing range information [*e.g.*, 1]. Quantum efficiency (QE) at these wavelengths is low and approximately proportional to active layer thickness – typically 3-10 µm. Larger pixel sizes are useful for improving signal-to-noise ratio (SNR), which in turn degrades the spatial resolution of the image.

Improvements in 3D imaging performance are critical to mass adoption in consumer electronics applications such as gesture-control interfaces, safety applications, and generating content for 3D media displays. A key metric for 3D imaging performance is distance error (DE), which depends on SNR, and the demodulation contrast (DC). DC improves with decreasing carrier transit length, typically proportional to pixel pitch, which is at odds with SNR improvement. Recent comparable work reported a DC around 38% at 20MHz with 10 µm pixel pitch [2].

Besides significantly improved DC and higher SNR due to the use of a larger pixel, a different system approach can be utilized. Most 3D ToF sensors deliver both in-phase (0^o) and out-of-phase (180^o) signals relative to the modulated light during one integration cycle. A second field is used to collect the quadrature signals at 90^o and 270^o phase difference. This is the so-called two-tap approach. While two-tap allows efficient signal capture, the pixel pair for two-tap operation requires careful matching of the two tap outputs. Otherwise, inter-pixel non-uniformity itself can degrade the performance of 3D imaging.

Concentric-Gate Pixel Architecture

In this work a concentric photogate (CG) pixel with a single-tap structure, and four "fields" are used to collect the four required signals. Matching requirements are eliminated and photo-response non-uniformity (PRNU) is reduced but at the cost of 30% reduction in SNR.

The device, shown in Fig. 1, consists of a 3-10 µm thick p- photodetector layer on a p+ substrate. Two concentric gates are formed. In the center is a floating diffusion which integrates the signal. The inner gate, IG is separated by 0.2 µm from the surrounding concentric outer drain gate DG that in turn is surrounded by a drain diffusion (DD). The latter can be in common with adjacent pixels. In the device, IG and DG are buried-channel devices. It was found that a buried channel yielded improved high- frequency DC characteristics compared to other structures including a pinned-diode structure. This can be expected from the fringing field effect which has been used in buried-channel charge-coupled devices (CCDs) and was also utilized in previous reports [*e.g.*, 2].

The effect of the fringing field on the demodulation performance of the concentric-gate pixel can be estimated from 3D TCAD simulation, as shown in Fig. 2. Generation of photoelectrons at 850 nm is fairly uniform throughout the photoactive layer. Carriers generated in the p+ substrate are largely removed by recombination. Carriers generated in the

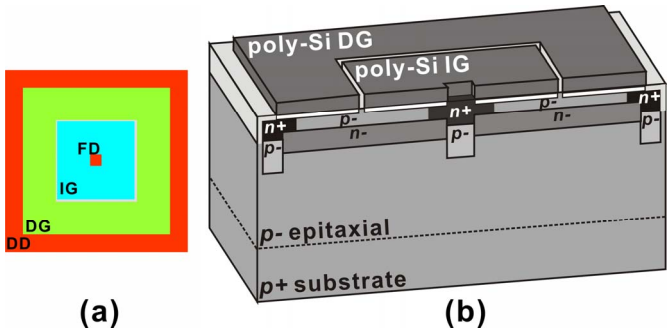

Fig. 1. (a) Schematic layout of concentric-gate pixel with floating diffusion FD, inner gate IG, drain gate DG, and drain diffusion DD. (b) Cut away view of pixel showing buried-channel structure.

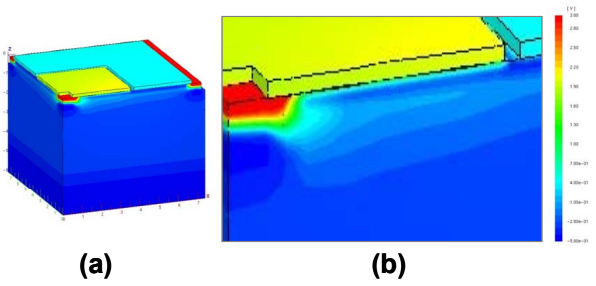

Fig. 2. Simulated potential diagram of the quarter vertical portion of the pixel after 10 nsec application of 1.6 V bias to inner gate and 0.0 V bias to drain gate (a), calculated from 3D TCAD. The potential diagram zoomed out near the FD region (b) shows that buried channel with lateral electrical field is formed so that high speed demodulation is expected.

photoactive layer are collected primarily by drift. Carriers first drift vertically to the surface into the buried n-channel under the device. There they are steered to either the FD or DD depending on the biasing of IG and DG.

Fabrication and Experimental

A 0.13 μm CMOS frontside-illuminated (FSI) image sensor process was used for the test chip. Devices with various p- and n- doping profiles and several p-epi layer thicknesses for the structure shown in Fig. 1 were fabricated and characterized. The test chip is comprised of a 192×108 array of pixels and simple on-chip circuitry. The signal chain is shown in Fig. 3 and resembles circuits used for characterization of

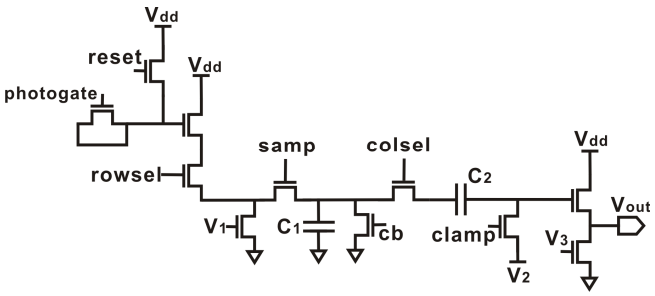

Fig. 3. Schematic diagram of signal processing chain used to obtain field images from the test chip [3].

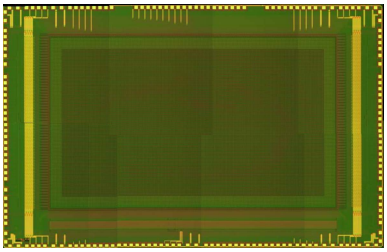

Fig. 4. Chip photograph of the 198X108 pixel ToF-3D image sensor.

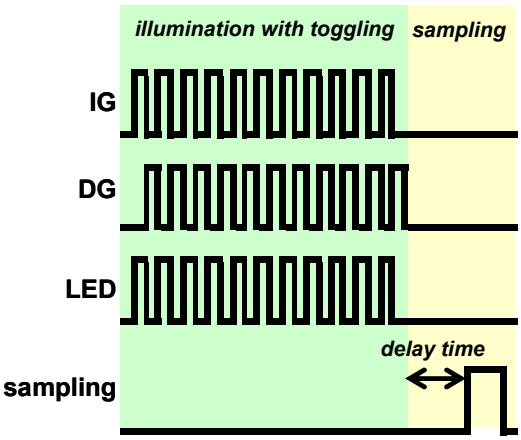

Fig. 5. Timing diagram of CG pixel operation. Normal 3D image is obtained with delay time less than 50 nsec.

early CMOS image sensors arrays. A chip photograph is shown in Fig. 4.

For characterization of the test chip, a timing diagram shown in Fig. 5 was used. Both IG and DG signals toggle out of phase from each other. The four fields are obtained by adjusting the LED timing phase as 0°, 90°, 180°, and 270° with respect to the IG timing. One field corresponds to measurement of one of the four phases so that four fields are needed to create one 3D image frame. With signals from each frame sig_0, sig_{90}, sig_{180}, and sig_{270}, distance R and DC were obtained from the following equations:

$$R = \frac{c}{4\pi f} \cdot \tan^{-1}\left(\frac{sig_0 - sig_{180}}{sig_{90} - sig_{270}}\right) \quad (1)$$

$$DC = 2 \cdot \frac{\sqrt{(sig_0 - sig_{180})^2 + (sig_{90} - sig_{270})^2}}{sig_0 + sig_{90} + sig_{180} + sig_{270}} \quad (2)$$

where c is the speed of light and f is the modulation frequency.

The sensor operates at 25 frames/sec. The reported sensor uses a fully-flexible open architecture so that timing control and analog-to-digital convertor (ADC) is implemented off-chip. Distance error (DE) as well as other performance parameters are characterized by images obtained from a custom built camera, which is comprised with field programmable gate array (FPGA), an image capture board, LED array, appropriate lens and optics, as well as our chip. Full electronic rolling shutter (ERS) timing is used to adjust for lighting.

Error correction and ambient light cancellation are implemented at the system level, but beyond the scope of this paper.

Results and Discussion

Frequency dependence of DC using 850 nm wavelength LED for various doping conditions of the photodetecting layer is shown in Fig. 6. It is evident that formation of a proper buried-channel structure improves demodulation performance over the whole observed frequency range, and DC value reaches as high as 53% at 20 MHz frequency.

Part of such high frequency performance is facilitated by rapid charge transfer. Figure 7 shows signal increase measured for various signal levels as a function of the delay time as defined in Fig. 5. When a non-optimal doping profile is used, the generated photoelectrons fail to arrive at FD during the illumination, especially at high frequencies, which degrades demodulation contrast. For instance, around 15 mV signal increase accounts for the remnant photoelectrons under CG. On the other hand, essentially no signal increase is observed for the pixel with a more optimized doping profile. All the results following Fig. 7 are obtained from chips with the same underlying p- and n-layers.

The improved demodulation performance leads to lower distance error. Figure 8 shows extracted distance measured from varying the camera to object distance of 1 to 7 m. It is found that precision is higher than 99%.

At each distance, DE is characterized from statistics with 100 images captured independently, and dependence of DE on distance is obtained as in Fig. 9. It is noted that DE is less than 0.9% over the range of 1 to 7 m.

It is generally accepted that DE is critically dependent on

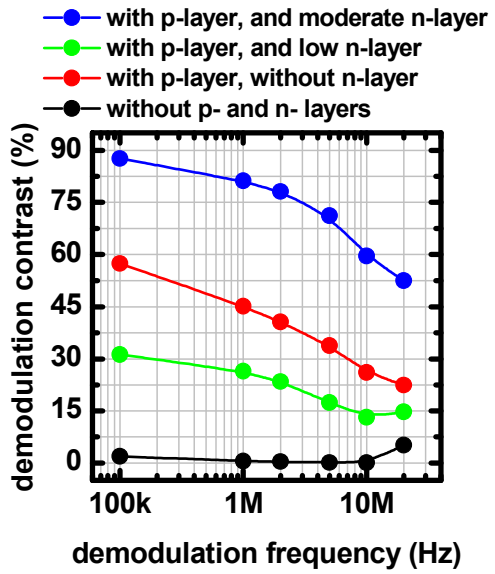

Fig. 6. Frequency dependence of demodulation contrast for various photodetecting layer formation conditions.

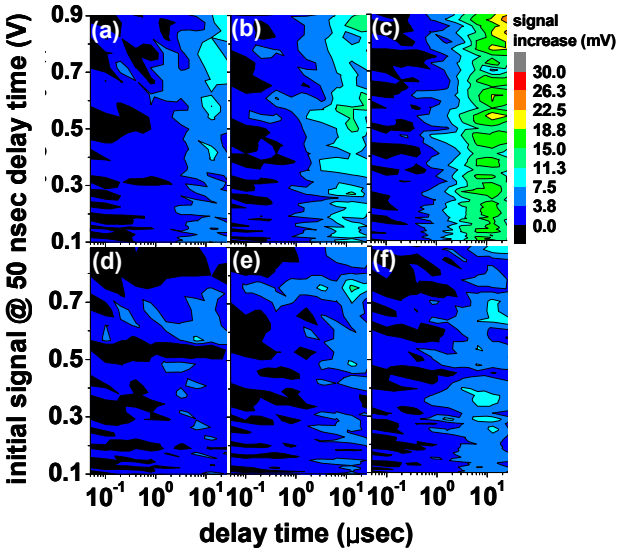

Fig. 7. Signal increase as a function of delay time measured from doping with optimal condition (panels below), and less optimal condition (panels above) at various demodulation frequencies of 1MHz (a, d), 5MHz (b, e), and 20MHz (c, f).

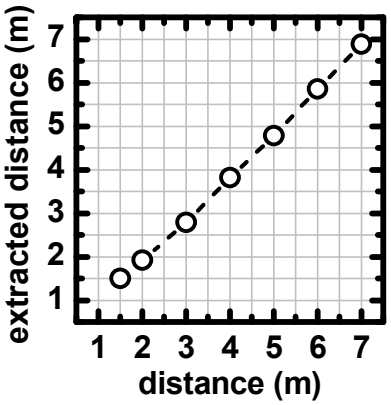

Fig. 8. Extracted distance from varying camera-to-white paper distance ranging from 1 to 7 m. Dotted line is obtained from the Monte Carlo simulation.

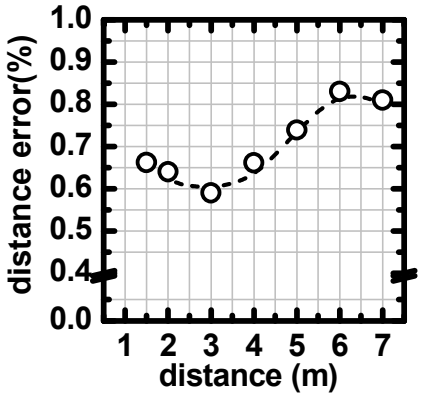

Fig. 9. Distance error measured in the range of 1 to 7 m. Dotted line is obtained from the Monte Carlo simulation.

DC, QE, demodulation frequency, as well as various noise sources. It is noted that temporal and spatial noises as well as non-uniform response of each pixel significantly deteriorates DE. A Monte-Carlo simulation of DE was carried out from the independently generated 100 frames of artificial images, in consideration of various temporal and spatial noise sources as well as PRNU. The simulation is found to fit detailed behavior observed from the measurements, as shown in Fig. 9.

Our simulation results in Fig.10 highlight the significant effect of PRNU on DE especially at shorter distances, in comparison to the other noise sources. In contrast to the case when PRNU is negligible, small PRNU increases DE dramatically, especially at near distance. Single-tap pixel architecture facilitates reduced inter-pixel non-uniformity and PRNU. Thus, it can be concluded that single-tap pixel structure and demodulation operation ensures enhanced performance, especially at shorter camera-to-object distances.

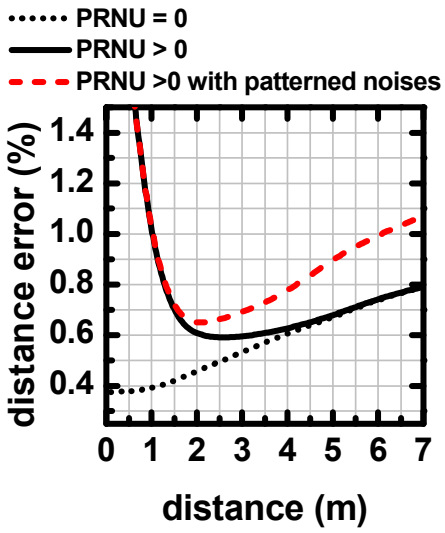

Fig. 10. Relation between distance error and distance expected from Monte Carlo simulation in consideration of various noise sources and PRNU.

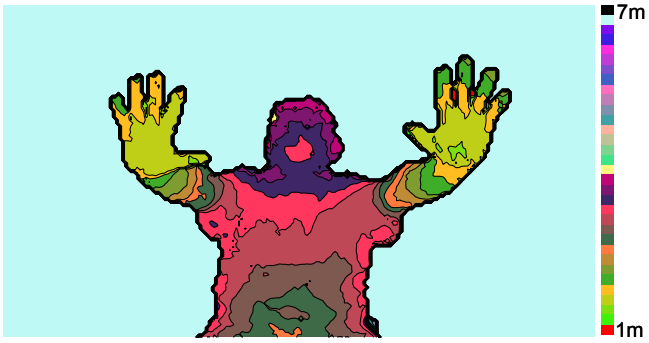

Fig. 11. A three dimensional image obtained from a human being. Not only is the distance information clearly visible, but also the image resolution is enough to resolve each finger at 1 m distance, which is important for gesture control.

TABLE I. PERFORMANCE SUMMARY FOR THE CONCENTRIC GATE DEMODULATION PIXEL

Parameters	Values
Chip size	7.8mm × 5.0mm
Pixel array	192(columns) × 108(rows)
Pixel size	28 μm × 28 μm
V_{pp}	2.8 V
Pixel conversion gain	7.2 μV/e
Signal to noise ratio	44.6 dB
Linear full well cap	28,700 e-
Dark current at 35°C, 20MHz	68.4 e/msec
DC at 20MHz	over 50%
DE at 7m	0.8%

A 3D image obtained by the sensor and subsequent signal processing is shown in Fig. 11. It is noted that every detail of the object along with the distance information is clearly visible.

Summary and Conclusions

A 3D-ToF FSI image sensor using novel concentric photogate pixels with single-tap operation is described. Through the use of CG structure, we are able to achieve high DC at larger pixel pitches. The new CG pixel structure substantially improves DC to 53% at 20MHz at 28 μm pixel pitch. Recent initial results from a backside-illuminated (BSI) implementation of the same sensor show further improved performance and will be reported elsewhere.

The sensor performance from the CG pixel array is summarized in Table I, which readily shows the significantly improvement over the previously reported devices. The size effect of the constituent photogates and possibility of tuned structure are under investigation.

References

(1) T. Oggier, "Image sensor technologies for 3D time-of-flight range imaging", *Proc. 2011 Int. Image Sensor Work*, June 2009.

(2) D. Stoppa, N. Massari, L. Pancheri, M. Malfatti, M. Perenzoni, and L. Gonzo, "A range image sensor based on 10-μm lock-in pixels in 0.18-μm CMOS imaging technology", *IEEE J. Solid-St Circ.*, vol. 46, pp. 248-258, January 2011.

(3) S. Mendis, S. Kemeny, R. Gee, B. Pain, C. Staller, Q. Kim, and E. Fossum, "CMOS active pixel image sensors for highly integrated imaging systems", *IEEE J. Solid-St Circ.*, vol.32, pp.187-192, February 1997