

Proposal of a Single-Shot Multi-Frame Multi-Frequency CMOS ToF Sensor

Peyman F. Shahandashti, P. López
V.M. Brea, D. García-Lesta

Centro Singular de Investigación en Tecnoloxías Intelixentes (CiTIUS)
University of Santiago de Compostela
15782, Santiago de Compostela, Spain
E-mail: peyman.fayyaz@usc.es; p.lopez@usc.es

Miguel Heredia-Conde

Center for Sensor Systems (ZESS)
University of Siegen
Paul-Bonatz-Str. 9-11, 57076 Siegen, Germany
E-mail: heredia@zess.uni-siegen.de

Abstract—A novel ToF image sensor architecture for simultaneously acquiring raw images at different modulation frequencies is introduced in this paper. The proposed approach constitutes a disruptive paradigm as it allows to obtain Fourier coefficients of the scene response function in a single-shot and free from harmonic distortion. Electrical simulations demonstrate the validity of the proposal.

Index Terms—3D image sensor, time-of-flight, single-shot, CMOS image sensor.

I. INTRODUCTION

Time-of-Flight (ToF) sensors work by measuring the round-trip time delay or phase difference of active light moving between the sensor and the target. The ability to measure, not only along two spatial dimensions, but also along the time dimension allows for estimating depth or distance between the camera and the scene on a per-pixel basis. To attain this, ToF pixels typically feature two or more accumulation wells, also called taps, and the integration process in each of them is controlled by signals that are typically binary and mutually exclusive. In other words, photo-generated carriers are accumulated in a specific well depending on their photon arrival time within the exposure, differently from conventional imaging sensors, which accumulate photo-generated carriers independently from their photon arrival times. Control of both the modulation of the illumination signal and the demodulation of the reflected light received by the ToF pixels allows for implementing custom measurement functions in the time domain. Typically-implemented sensing functions are pulse-shaped and sinusoidal. This gives rise to the two major operation modes of ToF imaging, namely, pulsed mode and continuous wave (CW) mode. In the first case, measurement diversity is obtained by shifting the ToF pixel control signals w.r.t. the illumination control signal (or the other way around), while in the second case, measurement diversity can be obtained by varying the

modulation frequency and the phase shift. In both cases, several acquisitions are required.

A ToF camera in CW mode with the appropriate harmonic cancellation mechanisms can directly extract Fourier coefficients of the scene response function [1]. Alternatively, the same information can be extracted from measurements obtained in pulsed mode [2], [3]. M-sequences and custom codes can be used as control signals to obtain sharp pulses as effective sensing functions [4]. In all these cases, the unknown parameters of simple scene response functions, such as a weighted sum of Dirac delta functions, can be retrieved. In other words, sufficient measurement diversity allows for disentangling multiple return paths per ToF pixel. Furthermore, using an appropriate regularization, one can attempt to reconstruct non-parametric scene response functions from the ToF measurements, e. g., using the sparsity-promoting ℓ_1 regularizer [4]. CW-ToF measurements at different frequencies and different phase shifts can also be used to attain transient imaging [5], or even imaging through highly scattering media. [6]. Similarly, the same type of CW-ToF measurements at multiple frequencies and phase shifts can be used to attain non-line-of-sight (NLOS) imaging [7].

The aforementioned applications unveil the potential of the measurement diversity provided by inexpensive ToF cameras to solve classically ill-conditioned problems. Nevertheless, taking profit of this potential comes at the cost of many sequential raw image acquisitions. This translates into abnormally long acquisition times (e. g., the 90 s of data acquisition for generating a transient image in [5]), which precludes the real-time operation of the aforementioned approaches. In order to bring these novel ToF-based computational imaging methods closer to real-time operation, simultaneous acquisition of multiple ToF raw images becomes necessary. In the limit, one would desire to have a single-shot ToF camera [8], able to acquire the full stack of required raw images within a single exposure. This is typically attained by different types of spatial multiplexing, which can be roughly classified in:

- Multi-tap ToF pixel, [9]–[11]. Provided that the photo-generated carriers cannot be integrated in more than one tap at each time instant, the control signals regulating each tap are mutually exclusive. Consequently, single-

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 860370. This work has also received funding from the Spanish government project RTI2018-097088-B-C32 (MINECO/FEDER), the Consellería de Cultura, Educación e Ordenación Universitaria (accreditation 2016-2019, ED431G/08 and ED431C 2017/69) and the European Regional Development Fund (ERDF).

shot multi-frequency CW-ToF cannot be directly implemented, provided that it implies overlapping control signals.

- Clustering of sets of neighboring ToF pixels forming macro-pixels, [12]. Each macro-pixel features a number of taps that is k times the number of taps of the underlying ToF pixels, where k is the number of the latter clustered per macro-pixel. Crucially, taps belonging to different ToF pixels within the macro-pixel can be controlled with partially-overlapping signals.
- Multi-aperture systems, [13]. In this kind of implementation, the ToF image sensor is divided into multiple sub-imagers of equal size, each featuring its lens. A major weakness of multi-aperture systems is the imperfect registration of the different raw images for objects close to the camera, giving rise to artifacts at their borders.

All these alternatives yield simultaneous acquisition of multiple raw images at the cost of compromising the spatial resolution of the sensor. So far, no existing ToF image sensor architecture has been proposed for simultaneously acquiring raw images at different modulation frequencies, despite this type of data has proven to be a key enabler in recent works on ToF computational imaging. Ideally, one would like to have a ToF imager able to directly extract several Fourier coefficients of the scene response function in a single-shot and free from harmonic distortion. In this work, we introduce a tailored ToF sensor design to accomplish this ambitious task.

II. SINGLE-SHOT MULTI-FRAME MULTI-FREQUENCY RESONANT DEMODULATION CONCEPT

In this work, we considered the 2-tap pixel architecture shown in Fig. 1, [14]. The design consists of a symmetrical structure containing a pinned photodiode photo-demodulator (PPD) and the corresponding readout circuit. The two transmission gates, TG_A and TG_B are used to direct the photo-generated charge acquired by the PPD to the floating diffusions FD_A and FD_B , respectively. First, the floating diffusion nodes are reset through transistors $M_{p3,4}$. Then, during the exposure time, the modulating signal is applied to the transmission gates, and the photo-generated charges are directed to the floating diffusions. A global shutter mechanism to ensure synchronous sampling of the demodulated signal through all pixels is commonly used in these kinds of structures ($M_{p1,2}$). The final values for each of the taps are read out through OUT_A and OUT_B signals.

The single-shot multi-frame acquisition can be achieved by grouping several ToF pixels into a macro-pixel, as seen in Fig. 2a. In this example, each macro pixel is formed by 4 sub-pixels, each of which will take the form in Fig. 1. Every sub-pixel will then have a 2-tap structure, and each tap will be controlled by one demodulating signal applied at TG_A and TG_B . These modulating signals will have the same frequency with a phase shift between them. Single-shot multi-frequency operation can be attained by controlling each of the ToF pixels inside each macro-pixel with an adequate control signal generated by on-chip electronics. This would

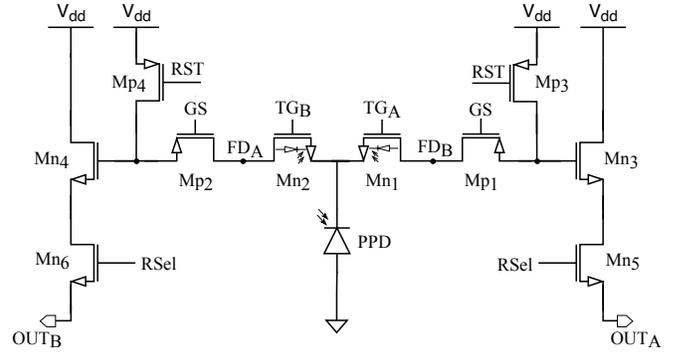


Fig. 1: ToF pixel schematic considered.

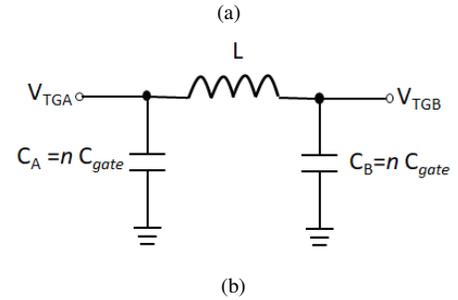
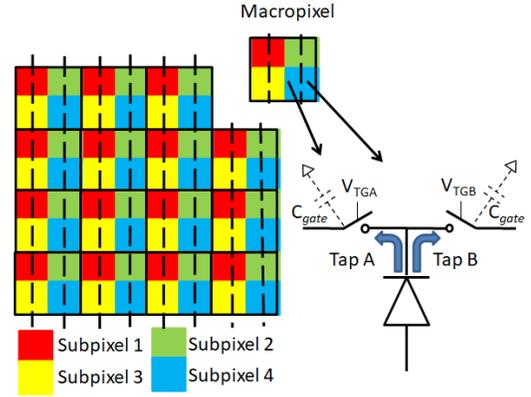


Fig. 2: (a) Single-shot multi-frame multi-frequency demodulation concept; (b) Resonant demodulation concept.

mean that, for example, a set of subpixels of one type, e.g., all the red subpixels in a row/column or predefined area, will be modulated at the same frequency (with a phase shift for each of its two taps) whereas the rest of the sub-pixels in the macro pixel (green, blue, and yellow in the example in Fig. 2a) will be modulated at different frequencies. In this way, one-shot multi-frame and multi-frequency 3D imaging can be achieved.

The proposed modulation paradigm has the additional advantage of achieving close-to-zero harmonic distortion by implementing a resonant circuit for each control signal, as seen in Fig. 2b. To do so, ToF pixels of the same type, that is, demodulating at the same frequency, or a subset of them, are controlled together (e.g., per row, per column). The parallel-

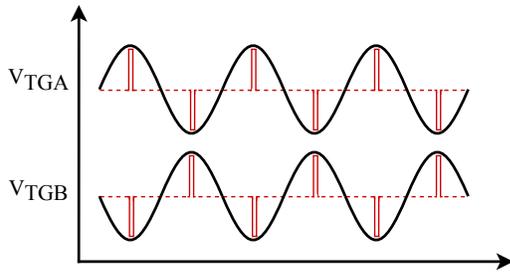


Fig. 3: Modulating signals applied at the modulating gates T_{GA} and T_{GB} . When the cycle repetition rate matches the resonant frequency, harmonic distortion reaches a minimum.

connected control gates present an intrinsic capacitance of adequate value to generate a resonant circuit at the desired frequency. As each subpixel has a 2-tap structure, we will refer to the lumped capacitance values resulting from connecting a set of same-type subpixels from different macro-pixels as C_A and C_B for the A and B taps, respectively. As the two taps are identical, $C_A = C_B = nC_{gate}$, where C_{gate} is the gate capacitance at each of the individual modulating gates and n is the number of connected subpixels. When an inductor, L , is connected across these two capacitors, the resonant frequency can be approximated as,

$$f_{res} = \frac{1}{2\pi} \sqrt{\frac{2}{nLC_{gate}}} \quad (1)$$

In order to achieve close-to-zero harmonic distortion, the modulating signals V_{TGA} and V_{TGB} , shown in Fig. 3, are applied at transmission gates T_{GA} and T_{GB} , respectively. First, taps A and B are pulsed high and low, respectively, and left to oscillate. Then, pulses in the opposite direction are applied, and the cycle is repeated. When the cycle repetition rate matches the resonant frequency, in-pixel resonant demodulation occurs, and harmonic distortion reaches a minimum.

III. SENSOR ARCHITECTURE

Fig. 4 shows the system architecture of the ToF image sensor, which consists of an array of macro-pixels, the row and column decoders, the column filter stage (CFS), and the double-delta-sampling (DDS) stage as well as digital control circuitry to program the number and position of the subpixels to be connected. The CFS block processes and filters pixel signals and can be configured to allow two separate sensor readout modes: direct and differential readout. A DDS approach is also used to minimize the column Fixed Pattern Noise (FPN). Additional control circuitry is added in order to define the regions within which subpixels of the same type will be connected together to form a resonant unit.

IV. SIMULATION RESULTS AND DISCUSSION

To demonstrate the proposed demodulation concept, we performed electrical simulations in an XFAB 180 nm optical process CMOS technology. In order to demonstrate the

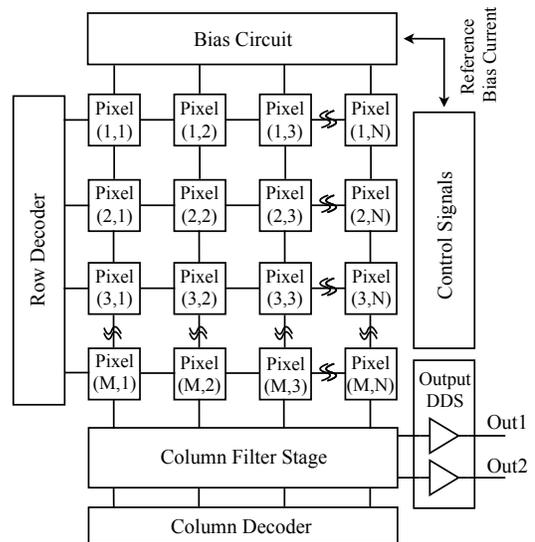


Fig. 4: System architecture of the ToF image sensor.

feasibility of the proposed approach, we considered an ideal inductance element. We studied different configurations of the array of pixels of the ToF sensor with macro-pixels of size 2×2 consisting of 4 sub-pixels with 2-taps each of them. Each of the subpixels was implemented as in Fig. 1. The gate capacitance value for each modulating gate was estimated by the simulation to be $C_{gate} = 7$ fF. First, we analyzed the most straightforward case with just one demodulating frequency. We considered a case with 4 macro-pixels and connected one subpixel from each macro-pixel, making a total of 4 connected nodes. The simulated lumped capacitance, in this case, is $C_A = C_B = 28.1$ fF, that is, roughly 4 times the gate capacitance of the individual subpixels, as expected. Then, a modulating signal of $f = 20$ MHz was applied, which, according to (1), means that resonance should occur for an estimated inductance value of $L = 4.5$ mH. We then performed a parametric simulation of the circuit for different values of the inductance. As can be seen in Fig. 5, a well-defined sine-wave signal is obtained at resonance when the inductance value is near the estimated one for that frequency. For other values of L , the retrieved signal shows a distorted behavior, as expected.

In order to prove the multi-frequency demodulation capability of the system, we performed a simulation for multiple demodulation frequencies. In this case, we considered 16×16 macropixels, each consisting of 4 sub-pixels, and four different demodulation frequencies (10, 20, 30 and 40 MHz) so that all subpixels of a given type are connected to the same frequency. As can be seen in Fig. 6, four well-defined sinusoidal demodulation signals are obtained for the expected L values for each of the resonant frequencies. As the number of connected gates is now 256, the total lumped capacitance, in this case, is $C_A = C_B = 1.79$ pF. Fig. 7 shows the obtained Total Harmonic Distortion (THD) in each case, defined as the ratio

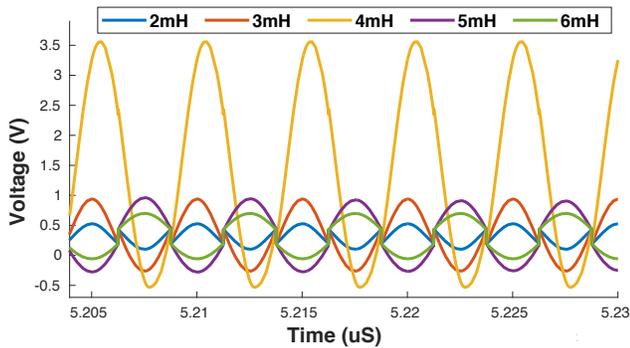


Fig. 5: Resonant demodulation for a frequency of $f = 20$ MHz and different values of the inductance.

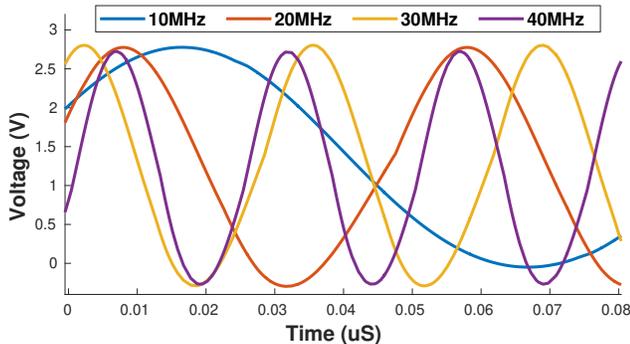


Fig. 6: Multi-frequency resonant demodulation.

of the sum of the powers of all harmonic components to the power of the fundamental frequency values, with respect to the L value. As can be seen, the minimum THD value occurs at the inductance value that corresponds to the corresponding resonant frequency as given by (1). As expected, increasing the number of connected gates increases the lumped gate capacitance, reducing the inductance value needed to achieve resonance. For the 20 MHz demodulating signal, resonance is achieved at $68 \mu\text{H}$ as opposed to the roughly 4.5 mH needed when only four gates were connected in Fig. 5. This means that, a reasonably sized array of 256×256 macro-pixels resonance for a modulating signal at 20 MHz could be achieved for inductance of 270 nH, a value small enough to be implemented on-chip.

V. CONCLUSION

A novel ToF sensor design showing a one-shot multi-frame multi-frequency harmonic canceling modulation has been presented. The demodulation concept is demonstrated by means of electrical simulations for different frequencies.

REFERENCES

- [1] M. Heredia Conde, T. Kerstein, and O. Buxbaum, B. and Loffeld, "Fast multipath estimation for PMD sensors," in *5th International Workshop on Compressed Sensing Theory and its Applications to Radar, Sonar, and Remote Sensing (CoSeRa 2018)*, 2018.
- [2] A. Bhandari and R. Raskar, "Signal processing for time-of-flight imaging sensors: An introduction to inverse problems in computational 3-D imaging," *IEEE Signal Processing Magazine*, vol. 33, no. 5, pp. 45–58, 2016.

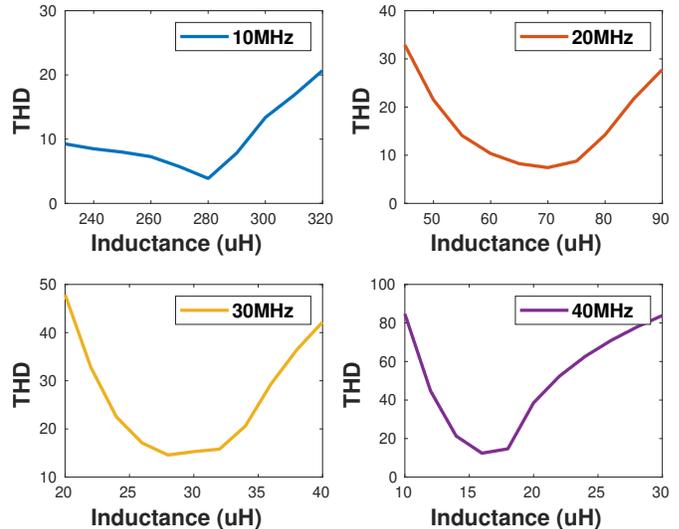


Fig. 7: Total Harmonic Distortion (THD) obtained with 16×16 macropixels and 4 different demodulation frequencies.

- [3] A. Bhandari, M. Heredia Conde, and O. Loffeld, "One-Bit Time-Resolved Imaging," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 42, pp. 1630–1641, 2020.
- [4] A. Kadambi, R. Whyte, A. Bhandari, L. Streeter, C. Barsi, A. Dorrington, and R. Raskar, "Coded time of flight cameras: sparse deconvolution to address multipath interference and recover time profiles," *ACM Trans. Graph*, vol. 32, no. 6, 2013.
- [5] F. Heide, M. Hullin, J. Gregson, and et al., "Low-budget transient imaging using photonic mixer devices," *ACM Trans. Graph*, vol. 32, no. 4, 2013.
- [6] F. Heide, L. Xiao, A. Kolb, M. B. Hullin, and W. Heidrich, "Imaging in scattering media using correlation image sensors and sparse convolutional coding," *Opt. Express*, vol. 22, pp. 26 338–26 350, 2014.
- [7] F. Heide, L. Xiao, W. Heidrich, and M. Hullin, "Diffuse Mirrors: 3D Reconstruction from Diffuse Indirect Illumination Using Inexpensive Time-of-Flight Sensors," in *IEEE Conference on Computer Vision and Pattern Recognition*, 2014, p. 3222–3229.
- [8] M. Heredia Conde, K. Kagawa, T. Kokado, S. Kawahito, and . Loffeld, "Single-Shot Real-Time Multiple-Path Time-of-Flight Depth Imaging for Multi-Aperture and Macro-Pixel Sensors," in *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 2020, pp. 1469–1473.
- [9] S. Kawahito, G. Baek, Z. Li, S.-M. Han, M.-W. Seo, K. Yasutomi, and K. Kagawa, "CMOS lock-in pixel image sensors with lateral electric field control for time-resolved imaging," in *Proceedings of the International Image Sensor Workshop*, vol. 10, 2013, p. 1417–1429.
- [10] T. Kasugai, S.-M. Han, H. Trang, T. Takasawa, S. Aoyama, K. Yasutomi, K. Kagawa, and S. Kawahito, "A Time-of-Flight CMOS range image sensor using 4-tap output pixels with lateral-electric-field control," in *Proceedings of the IS&T International Symposium on Electronic Imaging*, vol. IMS-048, 2016.
- [11] Y. Shirakawa, K. Yasutomi, K. Kagawa, S. Aoyama, and S. Kawahito, "An 8-Tap CMOS Lock-In Pixel Image Sensor for Short-Pulse Time-of-Flight Measurements," *Sensors*, vol. 20, no. 4, 2020.
- [12] K. Kagawa, T. Kokado, Y. Sato, F. Mochizuki, H. Nagahara, T. Takasawa, K. Yasutomi, and S. Kawahito, "Multi-tap macro-pixel based compressive ultra-high-speed cmos image sensor," in *Proceedings of 2019 the International Image Sensor Workshop (IISW)*, 2019, pp. 270–273.
- [13] F. Mochizuki, K. Kagawa, S. Okihara, M. Seo, B. Zhang, T. Takasawa, K. Yasutomi, and S. Kawahito, "Single-shot 200mfps 5×3 -aperture compressive CMOS imager," in *2015 IEEE International Solid-State Circuits Conference (ISSCC) Digest of Technical Papers*, 2015.
- [14] D. Stoppa and et al., "A Range Image Sensor Based on $10\text{-}\mu\text{m}$ Lock-In Pixels in $0.18\text{-}\mu\text{m}$ CMOS Imaging Technology," *IEEE Journal of Solid-State Circuits*, vol. 46, no. 1, 2010.