

39 000-Subexposures/s Dual-ADC CMOS Image Sensor With Dual-Tap Coded-Exposure Pixels for Single-Shot HDR and 3-D Computational Imaging

Rahul Gulve¹, Graduate Student Member, IEEE, Navid Sarhangnejad¹, Member, IEEE, Gairik Dutta, Motasem Sakr¹, Don Nguyen, Roberto Rangel, Graduate Student Member, IEEE, Wenzheng Chen, Zhengfan Xia, Mian Wei¹, Nikita Gusev, Esther Y. H. Lin¹, Xiaonong Sun, Graduate Student Member, IEEE, Leo Hanxu, Nikola Katic¹, Member, IEEE, Ameer M. S. Abdelhadi, Andreas Moshovos, Fellow, IEEE, Kiriakos N. Kutulakos¹, Member, IEEE, and Roman Genov¹, Senior Member, IEEE

Abstract—A dual-tap coded-exposure-pixel (CEP) image sensor is presented and validated in two computational imaging applications. The NMOS-only data-memory pixel (DMP) reduces the transistor count yielding a 7- μm pitch. One frame period can include up to 900 subexposures when operating at 30 frames/s, corresponding to 39 000 coded subexposures/s. The 320 \times 320-pixel sensor features two readout modes using column-parallel analog-to-digital converters (ADCs). ADC1 is a conventional high-accuracy $\Delta\Sigma$ -modulated ADC that digitizes pixel voltage at the end of every frame period, and ADC2 is a fast energy-efficient comparator that compares the pixel voltage with a constant reference voltage during each subexposure. The outputs of the 12-bit frame-rate ADC1 and the 1-bit subexposure-rate ADC2 are adaptively combined to boost the native dynamic range of the uncoded pixel by over 57 dB, demonstrating over 101-dB dynamic range in intensity imaging. In the second demonstrated application, combined with machine-learned projected illumination patterns, the CEP camera enables single-shot structured-light 3-D imaging at the native resolution and the nominal 30 frames/s video rate.

Manuscript received 29 November 2022; revised 26 March 2023; accepted 30 April 2023. This article was approved by Associate Editor David Stoppa. This work was supported in part by the Natural Sciences and Engineering Research Council of Canada through the Discovery Grants Program-Individual (RGPIN), Research Tools and Instruments (RTI) grants program, and Strategic Grant Program (SGP); and in part by CMC Microsystems. (Corresponding authors: Roman Genov; Rahul Gulve.)

Rahul Gulve, Motasem Sakr, Don Nguyen, Roberto Rangel, Xiaonong Sun, Leo Hanxu, Andreas Moshovos, and Roman Genov are with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: rahulgulve@ece.utoronto.ca; roman@eecs.utoronto.ca).

Navid Sarhangnejad, Gairik Dutta, and Nikita Gusev were with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. They are now with Alphawave Semi, Toronto, ON M5J 2M4, Canada.

Wenzheng Chen, Mian Wei, Esther Y. H. Lin, and Kiriakos N. Kutulakos are with the Department of Computer Science, University of Toronto, Toronto, ON M5S 2E4, Canada.

Zhengfan Xia and Nikola Katic were with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. They are now with Stathera Inc., Toronto, QC H3A 1L4, Canada.

Ameer M. S. Abdelhadi was with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada. He is now with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON L8S 4K1, Canada.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/JSSC.2023.3275271>.

Digital Object Identifier 10.1109/JSSC.2023.3275271

Index Terms—3-D imaging, CMOS image sensors (CISs), high dynamic range (HDR) imaging, high-speed imaging systems.

I. INTRODUCTION

COMPUTATIONAL imaging is at the core of most today's high-end consumer cameras, such as those in smartphones. It often involves taking several low-quality shots and combining them into one digitally enhanced high-quality image through software postprocessing. One well-known such example is taking several underexposed and overexposed images of a scene using a low-cost low-dynamic-range (LDR) image sensor and selectively merging them into one high-dynamic-range (HDR) image [1]. When used in conventional standard-frame-rate cameras, computational imaging works well for scenes where light intensity does not change rapidly. However, it typically fails in applications where there is fast-motion or fast-changing illumination in the scene due to motion artifacts, such as motion blur and ghosting.

High-frame-rate image sensors can reduce such motion artifacts and enable fast computational imaging. They operate at frame rates much higher than most conventional cameras and perform one fast readout per short exposure, as illustrated in Fig. 1(a). However, these sensors are often prone to: 1) low signal-to-noise ratio (SNR) due to low photogenerated charge levels; 2) high power consumption due to increased ADC conversion rate; and 3) high output data rate that requires expensive digital hardware to handle.

A. Coded-Exposure Image Sensors

The emerging class of coded-exposure image sensors [2], [3], [4], [5], [6], [7], [8], [9] aims to eliminate these drawbacks and enable novel fast computational imaging applications such as single-shot HDR imaging [2], [6], single-shot compressive sensing for high-speed video capture [9], [10], [11], [12], and single-shot 3-D imaging [2], [4], [13], [14]. The term “single-shot” refers to the standard terminology in computer vision corresponding to the duration of a single frame exposure and readout of a conventional camera. As illustrated in Fig. 1(b)–(e), in these image sensors, the total exposure time of one frame is divided into multiple (N) short programmable subexposures, which are performed within a single

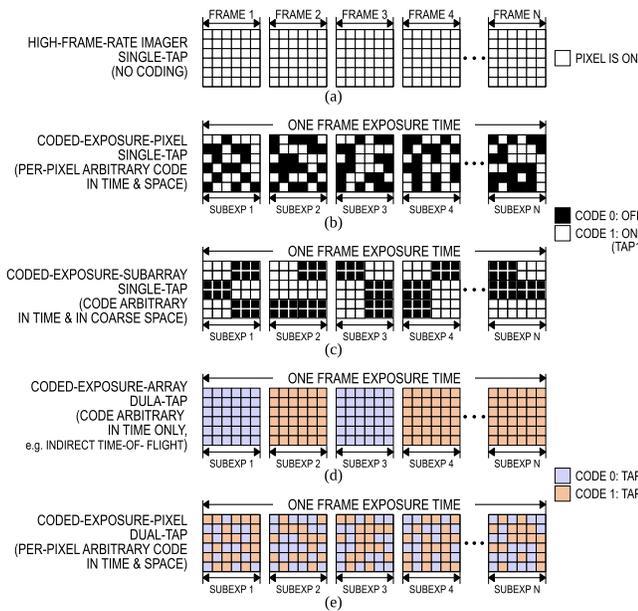


Fig. 1. Overview of different exposure-coding schemes. (a) High-frame-rate cameras. (b) Coded-exposure array with single-tap pixels. (c) Coded-exposure subarray with single-tap pixels. (d) Per-pixel coded-exposure with single-tap. (e) Per-pixel coded-exposure with dual-tap.

frame period and are followed by a single readout. In each subexposure, a pixel selectively accumulates photogenerated charge based on its individual 1-bit binary coefficient, referred to as the “code.” These codes are organized in frame-sized matrices, one per subexposure, referred to as “masks.” This approach attains: 1) a higher SNR as the photogenerated signal is accumulated over the full frame period time before it is read out; 2) a lower ADC sampling rate which keeps the power low; and 3) a lower output data rate yielding lower cost.

1) *Single-Tap Coded-Exposure Image Sensors*: Most of these sensors use a single photogenerated charge collection node, known as a “tap,” to perform the selective accumulation of the photogenerated charge, as shown in Fig. 1(b) and (c) [7], [9], [15]. This integration is temporally controlled based on the binary code assigned to a pixel that turns it on or off in each subexposure. In these image sensors, the on/off exposure-time programmability is implemented either by sharing the same binary code among a subset of pixels (i.e., a subarray of pixels) [7], [15] on a coarse spatial scale, as shown in Fig. 1(b), or using an independent code to control the on/off exposure status of each individual pixel [9], as depicted in Fig. 1(c). The latter approach, referred to as coded-exposure-pixel (CEP) image sensors, yields the highest spatial resolution (i.e., the native resolution) and thus offers the best computational imaging quality and fidelity.

2) *Dual-Tap Coded-Exposure Image Sensors*: Coded-exposure image sensors with two taps have also been recently introduced [2], [3], [4], [5], [16], [17]. In their simplest form, as a point of reference, the well-known indirect time-of-flight (iToF) image sensors [16], [17] can be viewed as two-tap sensors that are limited to performing only full-array spatial coding (i.e., all the pixels use the same binary code) but that offer temporal coding capability (to demodulate the input light phase to measure the distance to the scene) as depicted in Fig. 1(d). This temporal-only coding is sufficient

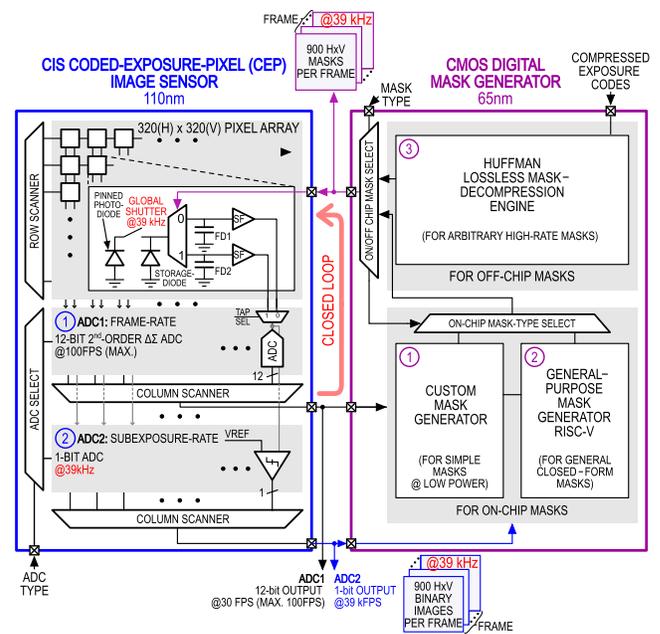


Fig. 2. Single-shot adaptive CEP imaging system block diagram showing the chip architecture of the CEP image sensor IC (left) connected in a closed loop with the digital mask generator IC (right).

for their specific field of use—long-range, fast 3-D imaging, but does not generalize to most other computational imaging applications.

General-purpose two-tap coded-exposure image sensors have also been recently introduced [2], [3], [4], [5] that perform not only temporal coding, as do iToF sensors, but also spatial coding. These two-tap sensors are typically implemented as CEP image sensors, i.e., they use per-pixel arbitrary binary codes that are sent to each pixel individually, in each subexposure, for fine, native-resolution control of exposure, as illustrated in Fig. 1(e). In such two-tap sensors, the photogenerated charge is programmed to be accumulated on one of the two taps in each subexposure, as controlled by an externally supplied code. This further boosts the SNR of computational imaging, as instead of draining the photogenerated charge when a pixel is off (and thus losing that signal), the photogenerated charge is collected on the second tap of that pixel during that subexposure, so no signal is lost. Two taps also offer many additional new capabilities for fast computational imaging, such as single-shot 3-D imaging featured in this work, as well as single-shot depth-gating [13], [14], [18], and single-shot direct-indirect imaging that sorts single-bounce and multibounce photons for robust imaging in the presence of reflection and refraction [19], [20]. To date, however, these sensors have only been implemented using in-pixel PMOS transistors making the pixel large and slow [3], [4], [5].

B. DMP Image Sensor Overview

We present a two-tap CMOS image sensor (CIS) comprising integrated circuits (ICs): a 110-nm CIS image sensor and a 65-nm CMOS mask generator, as depicted in Fig. 2. The image sensor shown in Fig. 2 (left) includes a 320 × 320-pixel array of dual-tap PMOS-free CEP, here referred to as the data-memory pixel (DMP) and a dual-ADC readout. The

pixel achieves a $7\text{-}\mu\text{m}$ pitch and a subexposure rate of 39 000 subexposures/s. The compact NMOS-only implementation eliminates any crosstalk between photo-sensitive pinned photodiode (PPD) and PMOS doping layers. As a result, the DMP is a factor of $3.24\times$ smaller and a factor of $1.7\times$ faster than the best state-of-the-art dual-tap CEP [3] and offers a factor of $2.7\times$ larger pixel array. The two column-parallel ADCs, ADC1 and ADC2, digitize the taps' outputs at the maximum frame rate of 100 frames/s. To reduce the power of wireline communication and external memory, the CIS image sensor can be stacked with a digital-CMOS mask generator, such as the one shown in Fig. 2 (right). The mask generator IC includes: 1) a custom low-power mask generator; 2) an RISC-V processor; and 3) a lossless Huffman-decompression engine, each for different types of masks and power requirements. We have experimentally demonstrated the sensor in a wide range of important fast computational imaging applications, which validate its versatility. For the sake of brevity, we include two such experimental validations in this work: 1) true single-shot adaptive HDR imaging and 2) single-shot structured-light 3-D imaging. We primarily focus on the former as it is a key emerging market driver, but also because, in many cases, it can be combined with other computational imaging paradigms implemented on the same sensor at the same time, to boost their dynamic range (the latter is beyond the scope of the current work). The second application is only briefly discussed to demonstrate the dual-tap sensor's versatility.

Our adaptive-HDR scheme completes exposure and adaptive code generation entirely within a single shot (i.e., within a single frame period) and does not require multiple shots used in most conventional cameras [1], nor does it suffer from one frame period lag needed to generate the adaptive codes in most existing pseudo-single-shot adaptive-HDR image sensors [7], [21]. As a result, artifacts due to the fast-changing intensity of incident light, such as due to fast motion or rapidly changing illumination, are significantly reduced. This HDR scheme extends the native dynamic range of the most conventional photodetectors by $20\log_{10}(N)$ dB (≈ 57 dB for $N = 900$ in this work) and can be implemented in most standard CIS processes without relying on exotic or expensive HDR pixel fabrication technologies.

We also demonstrate this image sensor in another fast CEP imaging application—single-shot structured-light 3-D imaging. This is achieved by simply reprogramming the pixel codes without any other changes to the sensor hardware. This validates the sensor's field-programmable versatility—it can be configured by the end user to perform a wide range of computational imaging tasks by simply reconfiguring its pixel codes (i.e., its “firmware”). The presented image sensor was first reported in [2]. This work expands upon [2] and is organized as follows.

Section I-A provides an in-depth review of various coded-exposure image sensors, including coded pixel-subarray image sensors and CEP sensors. It also includes a detailed comparative analysis of the state-of-the-art coded-exposure sensors with the two-tap CEP image sensors presented in this work.

Section II presents implementation details of different aspects of the work. Section II-A provides a detailed description of the pixel schematic and its layout considerations.

ADC1 and ADC2 circuit design and implementation details are presented in Section II-B. Section II-C describes different types of exposure codes that can be generated ON-chip and their use cases.

Section III presents the experimental results. Section III-A1 explains the characterization methodology and results. The simple coded-exposure results are shown in Section III-B followed by Sections III-B1 and III-B2, showcasing the sensor's abilities in single-shot scene-adaptive HDR imaging and structured-light 3-D imaging, respectively.

Section IV provides an up-to-date comparison to the state-of-the-art and includes a discussion on the advantages, limitations, and future directions.

II. VLSI IMPLEMENTATION

A. Dual-Tap Coded-Exposure DMP

The key challenges in designing CEP image sensors are the pixel area and the time overhead due to the in-pixel exposure control circuits. All the existing CEP image sensor pixels [3], [4], [5], [9] belong to the class of pixels we refer to as code-memory pixels (CMPs). They require in-pixel digital memory with PMOS transistors to store the exposure code at the cost of a large and slow pixel. Here, we introduce an NMOS-only two-tap coded-exposure DMP architecture that eliminates the need for in-pixel storage of the exposure code and yields a smaller pixel pitch.

Fig. 3 shows comparison of the existing pixel architectures with the presented dual-tap DMP. As shown in Fig. 3(a) (top), the conventional iToF pixel has two charge collection nodes controlled by modulation signals MOD and $\overline{\text{MOD}}$ shared by all the pixels in the array. The absence of any additional per-pixel-coding circuit (due to a globally shared modulation signal) leads to a smaller pixel size but does not allow for per-pixel coding. As mentioned earlier in Fig. 1, the iToF pixel is a trivial, temporally but not spatially coded, example of a dual-tap pixel.

In CEPs, some form of per-pixel code memory has been typically necessary to control the transfer gates of taps, as shown in Fig. 3(a) (middle). The code memory may consist of in-pixel pipelined latches [4], static random-access memory (SRAM) [3], or dynamic random-access memory (DRAM) [5], which all require the use of PMOS transistors in the pixel, making them large. In-pixel PMOS devices can also compromise the performance of PPDs.

Compared with pixels with in-pixel code memory, the conventional DMP, as depicted in Fig. 3(a) (bottom), also known as the global-shutter pixel, consists of a data-memory (DM) node that stores the charge before transferring it to a tap. The pipelined nature of the global charge transfer achieves global-shutter operation without the need for extra in-pixel circuits, making the pixels smaller.

The advantages of each of the existing pixel architectures highlighted in green color in Fig. 3(a): 1) dual taps from the iToF pixel; 2) per-pixel coding from the coded-exposure CMP pixel; and 3) compact intermediate-storage node from the DMP, are combined to realize the presented dual-tap coded-exposure DMP, as shown in the schematic in Fig. 3(b). By mirroring the transfer gate TG1 of the conventional noncoded

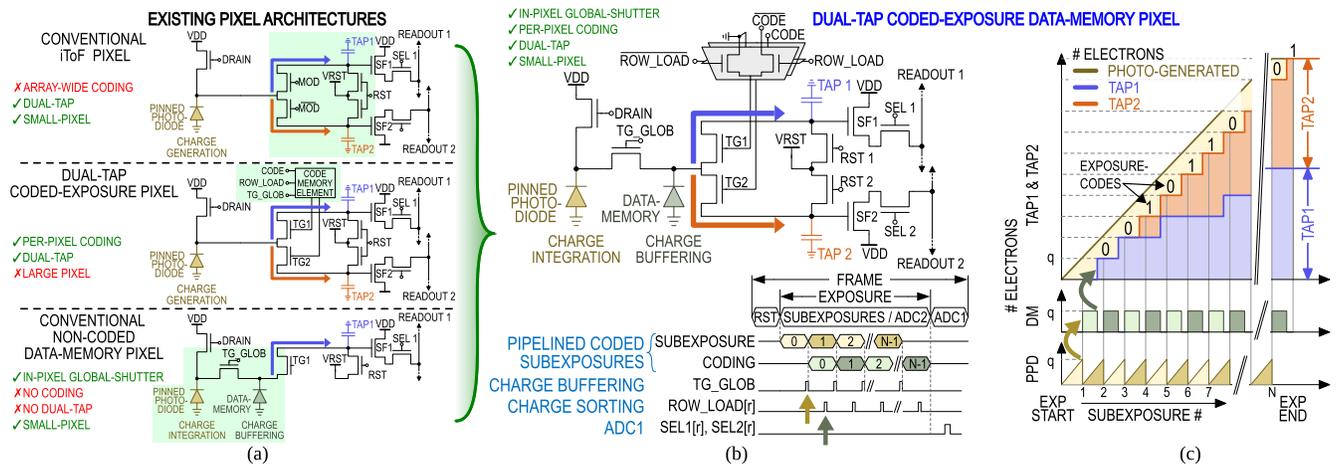


Fig. 3. Comparison between the existing pixel architectures and DMPs. (a) Parts of the existing pixel architectures similar to dual-tap coded-exposure DMP. (b) Schematic and timing diagram of the coded-exposure dual-tap DMP. (c) Amount of charge during exposure at different nodes: TAP1, TAP2, PPD, and DM.

DMP, we add a second tap to realize the dual taps. The pixel now has two charge collection sites, TAP1 and TAP2, accessed by transfer gates TG1 and TG2, respectively. These transfer gates are controlled by a pair of simple NMOS-only 2:1 multiplexers. Rowwise signal, ROW_LOAD, and columnwise signal, CODE, both provided from outside of the pixel. Allow performing for per-pixel coded exposure without the need for in-pixel code memory, as is the case for all the coded-exposure CMPs.

As shown in the timing diagram for coded-exposure cameras in Fig. 3(b) (bottom), the frame time is divided into N coded subexposures. Each coded subexposure has two parts, subexposure and coding, performed in a pipeline fashion. Compared with a conventional global-shutter pixel, the transfer gates are controlled by a combination of ROW_LOAD and externally applied CODE signals for charge sorting. The global signal TG_GLOB marks the end of every subexposure when asserted. It transfers the charge from the photodiode to the DM across all the pixels in the array. This operation allows us to achieve the coded global-shutter exposure. The charge is stored in the DM until it is transferred to one of the taps based on the exposure code. The code is applied to transfer gates when the ROW_LOAD signal is asserted for a given row. While the charge is sorted to respective taps, the photodiode continues to collect light. After the charge sorting is complete for all the rows, the TG_GLOB signal can be asserted to mark the end of the second subexposure. It is then again followed by rowwise coding for the second subexposure. These steps are repeated for all the subexposures. At the end of the frame exposure time, all the photogenerated charge is collected in TAP1 or TAP2, and none of the charge is lost due to coded exposure. As a result, the photogenerated charge across all the subexposures of a frame is selectively integrated on the two taps according to the per-pixel code sequence and is then read out once at the end of the frame as two images. The exposure codes for each row are streamed into the CIS. A bank of 10×200 -MHz dual-data-rate 1:32 deserializers, similar to that in [22], is used to load the exposure codes for each row. The mask upload takes 80 ns per row or $25.6 \mu\text{s}$ per array and is repeated up to $N = 900$ times per frame at

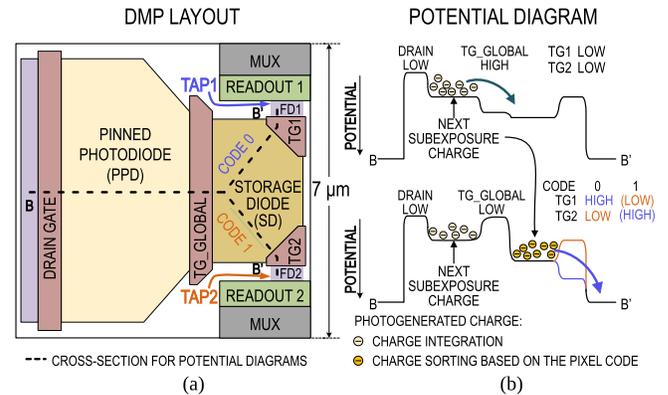


Fig. 4. Dual-tap coded-exposure DMP. (a) Layout and (b) corresponding potential diagrams during the global data sampling and charge sorting phases.

30 frames/s, accounting for 10-ms ADC1 readout time. The total subexposure time of $25.6 \mu\text{s}$ translates to the subexposure rate of more than 39 kHz.

The graph in Fig. 3(c) shows how the charge is transferred from the PPD to the DM, and then to one of the taps based on the exposure code. Fig. 3(c) also shows the amount of electrons at different nodes in the pixel during the exposure period. The combined charge in TAP1 and TAP2 equals all the photogenerated charge during exposure, as no charge is lost due to the dual-tap nature of the pixel.

Fig. 4 shows the abstract layout and potential-well diagram of the DMP pixel. Compared with the global-shutter dual-tap CMP in [3] and [4], DMP eliminates PMOS transistors, reduces the transistor count, and operates at a higher subexposure rate of 39 000 subexposures per second and at a higher pixel-code rate of 4 Gb/s, at 320×320 -pixel sensor resolution. The pixel achieves a 38.5% fill factor (FF). In the coded-exposure DMP, the DM storage diode (SD) must have a comparable area to the PPD for good charge transfer efficiency. In this design, an SD-to-PPD area ratio of approximately 39% is chosen. The two readouts and two multiplexer circuits per pixel moderately reduce the FF. Additional improvement of the effective FF can be achieved using techniques such as incorporating microlenses and light guide structures [23] or backside

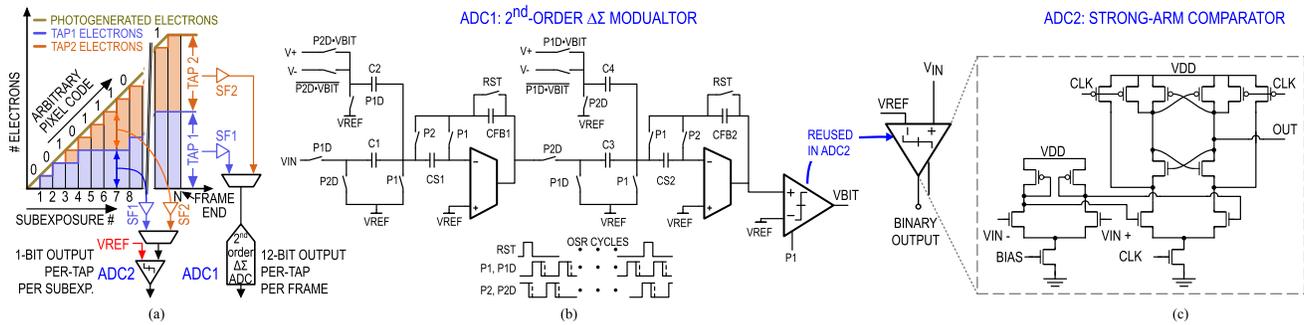


Fig. 5. Operation of two ADCs, ADC1 and ADC2, (a) within a single frame, and their architecture of (b) ADC1—a second-order $\Delta\Sigma$ -modulated ADC and (c) ADC2—a strong-arm comparator with a preamplifier.

illumination. The dual-tap DMP architecture presented here accumulates photogenerated charge in taps during the exposure phase, limiting it to double-sampling and making it unable to perform correlated double-sampling (CDS) during readout. One potential solution to this limitation is the inclusion of in-pixel metal-insulator-metal (MIM) capacitors to sample the reset noise before exposure, allowing for CDS during readout. However, this comes at the cost of a lower FF or increased pixel-pitch—micro-lenses or backside illuminated technologies can be used to, in turn, address these issues. We have also recently developed a technique to perform kTC and other noise compensation using digital regression [24].

B. Dual-Mode ADC Readout

Conventional image sensors typically include a bank of column-parallel analog-to-digital converters (ADCs). The ADCs read the (analog) amount of charge at pixel tap(s) and convert it into a digital number during the readout phase of the operation. Recently, there have been reported sensors that use stacked technology to implement an ADC per pixel [25] or per group of pixels [7] that rely on an expensive fabrication process with per-pixel interconnects.

When compared with conventional sensors, the presented sensor features two readout modes using column-parallel ADC1 and ADC2, as shown in Fig. 5(a). ADC1 is a conventional high-accuracy ADC that converts each pixel-tap voltage into a digital number at the end of every frame. ADC2 is a fast subexposure-rate 1-bit comparator that compares the tap voltage with an external reference voltage during every subexposure.

1) *ADC1 (Frame-Rate $\Delta\Sigma$ ADC)*: The frame-rate ADC1 consists of a second-order $\Delta\Sigma$ modulator, as shown in Fig. 5(b), and a decimation filter, as originally presented in [26]. Each ADC1 in the column-parallel bank digitizes both the taps of all the pixels in its column. The data from decimation filters are transferred using ON-chip serializers. ADC1 bank digitizes the data from both the taps at up to 100 frames/s while consuming 107 mW of power. During the exposure period, ADC1 is idle. This allows us to reuse the strong-arm comparator from ADC1 in ADC2, for area efficiency.

2) *ADC2 (Subexposure-Rate 1-Bit Comparator)*: The ADC2 is a column-parallel 1-bit ADC that compares the

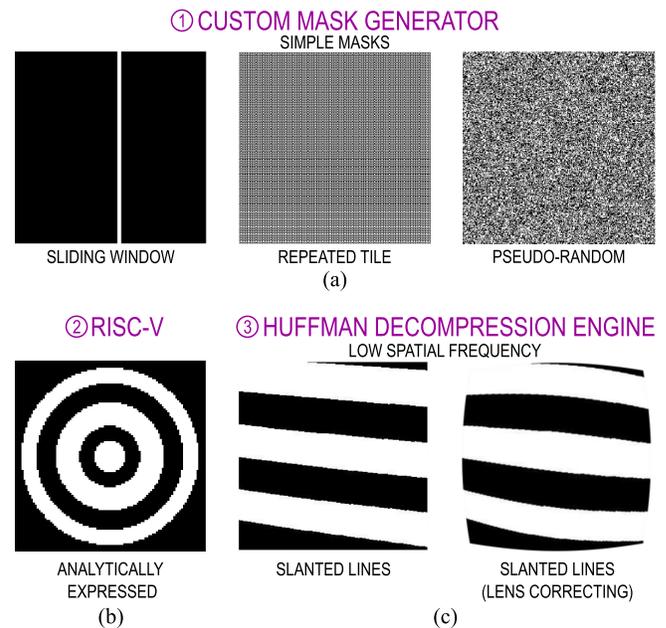


Fig. 6. ON-chip mask generator can produce (a) simple masks from a custom mask generator, (b) analytically expressed, or closed-form, masks from the RISC-V processor, and (c) masks with low spatial frequency decompressed using the Huffman decompression engine.

pixel-tap voltage with a reference voltage during every subexposure. It consists of a strong-arm comparator, as shown in Fig. 5(c). The reference voltage is set to a constant value that is specific to the application. An external voltage regulator can be used to provide a stable voltage. The reference voltage pin in ADC2 consumes a negligible current, as it is directly connected to transistor gates. The comparators generate a thermometer-style bit-stream output for each pixel-tap. When a row is selected for uploading an exposure code to the pixel, the voltage from each tap is buffered on READOUT lines through the pixel's source followers. This allows us to monitor the decrease in the tap voltage during each subexposure and adjust the exposure codes based on the application.

C. Mask Generator IC

The DMP array can receive arbitrary codes at the rate of 4 Gb/s. Conventionally, the flexibility of exposure codes is maintained by generating such codes OFF-chip, stored in

external DRAM, and sent to the sensor over long printed circuit board (PCB) wires. To reduce the power of wireline communication and avoid using energy-costly DRAM, the CIS image sensor can be stacked with a digital-CMOS mask generator, such as the one shown in Fig. 2 (right).

While the sensor is capable of using arbitrary exposure codes, the spatio-temporal complexity of the codes depends on the applications and many applications use simple codes, e.g., code masks with repeated 2×2 tiles [13], [14], rolling window [18], [20], [27]; sparse scene-adaptive [7], [21], and pseudorandom [10] codes. The ON-chip mask generation is realized using three separate digital blocks to offer three different levels of complexities of masks.

- 1) The custom mask generator block is the smallest of the three and generates the simplest set of exposure codes. These codes can have simple scan lines, sliding windows, and repeated tiled patterns, or they can be pseudorandom, as shown in Fig. 6(a). This block consists of simple sequential logic to realize the repeated and sliding patterns and a bank of pseudorandom number generators for random codes.
- 2) The ON-chip RISC-V processor is connected to both the image sensor output and the masking circuit. It can generate closed-form exposure codes based on the sensor output. It can also generate a set of exposure codes that could be efficiently expressed through an algorithm, e.g., concentric circles shown in Fig. 6(b).
- 3) The lossless Huffman-decompression engine is used for all other types of exposure codes, those that are too complex to be generated on the chip, e.g., masks compensating for lens distortion as shown in Fig. 6.

Such code masks are compressed OFF-chip using the Huffman method [28]. The dictionary of the compressed codes is loaded once in the engine's SRAM at the start of the image capture. The Huffman-compressed data stream is transferred to the engine, and the decompressed output from the engine is then fed into the sensor.

III. EXPERIMENTAL RESULTS

A. IC Characterization

Fig. 7 shows the ICs' micrographs and the power breakdown of different blocks. Each IC is 3.3×4.2 mm in dimensions.

Fig. 8 shows the camera system used for experimental characterization. Fig. 8(a) shows the camera PCB that accommodates a CIS IC under the lens and a field-programmable gate array (FPGA) board, which synchronizes subexposures with a digital micromirror device (DMD) projector for active-illumination computational imaging applications, such as single-shot structured-light 3-D imaging demonstrated in Section III-B2. Fig. 8(b) shows the PCB used for mask generation IC characterization. The block diagram of the different components on the CIS and mask generation PCB, the FPGA, and their interconnections is exhibited in Fig. 8(c) and (d). Although both the dies have a compatible pin layout for vertical pad-to-pad connection, we chose to test them separately for ease of experimental characterization. When testing the image sensor die, we used an FPGA to transfer mask data and control signals. Similarly, during testing of the mask generator

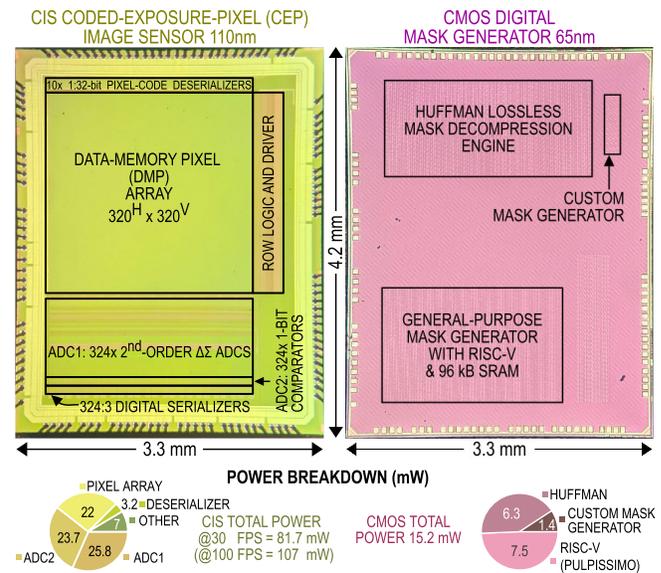


Fig. 7. Micrographs and power consumption of the CEP image sensor IC (left) and the custom mask generator IC (right).

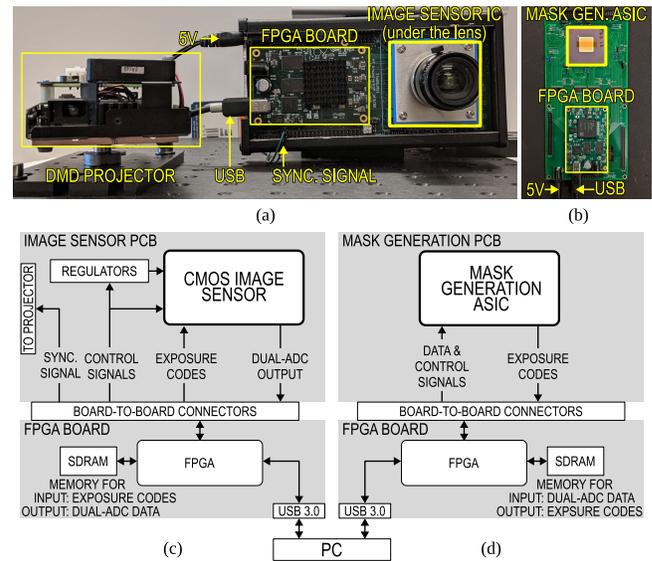


Fig. 8. Experimental setup includes (a) camera with the presented image sensor and a synchronized light-pattern projector and (b) mask generation system. The block diagram of (c) camera and (d) mask generation system is also included.

die, another FPGA was used to transfer image data and control signals. The power consumed during data transfer between the two chips has been simulated. For a maximum data throughput of 4 Gb/s, it is estimated to consume 2.4 mW to drive digital input-output pads, when two dies are connected directly.

1) *Dual-Tap Pixel*: The contrast between two taps is a more important requirement in computational photography sensors compared with iToF image sensors. In iToF sensors, a 60%–70% tap contrast is sufficient in most cases, as the distance is measured using the signal phase [29]. In CEP image sensors, a higher contrast is beneficial as it allows distinction between minute changes in exposure-code sequences, especially when imaging with active illumination, such as using a light-pattern projector.

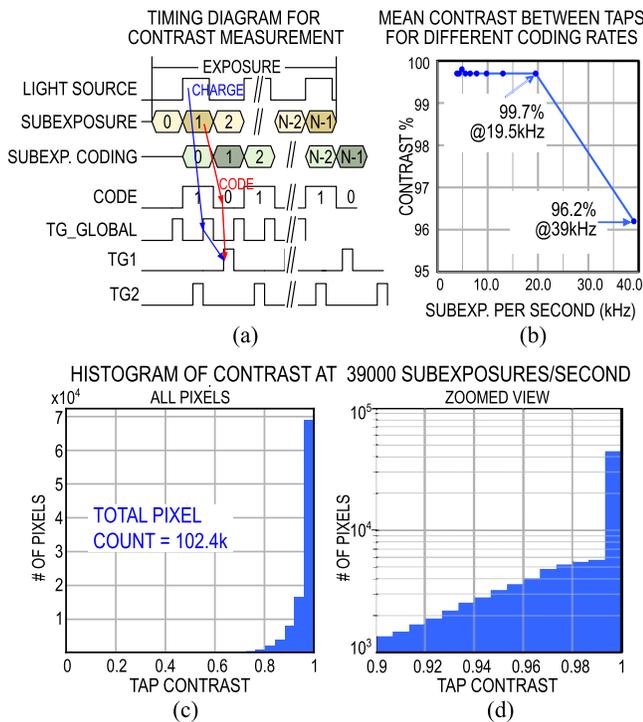


Fig. 9. Experimentally measured tap contrast in coded exposure sensor. (a) Timing diagram of the experimental setup. (b) Mean contrast at different subexposure speeds. (c) Histogram of contrast of all the pixels. (d) Zoomed in x-axis view at the highest subexposure speed of 39 kHz.

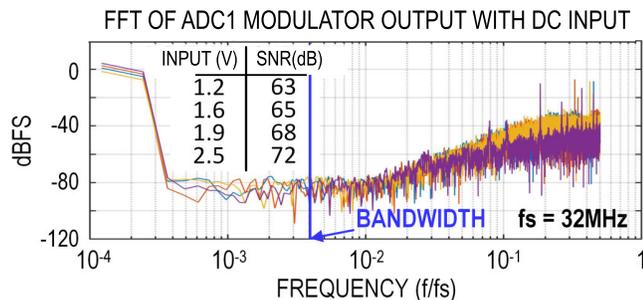


Fig. 10. Experimentally measured SNR of ADC1 output for several dc inputs.

Fig. 9(a) shows the timing diagram used to measure the tap contrast. During the measurement, all the pixels receive codes 0 and 1 in alternating subexposures. A uniform light source (wavelength 465 nm) is also turned on and off during every other subexposure. In the ideal case, all the photogenerated electrons are collected in TAP1. The contrast of the sensor is calculated as follows:

$$\text{CONTRAST} = \frac{Q1 - Q2}{Q1 + Q2} \times 100\% \quad (1)$$

where $Q1$ and $Q2$ are the amount of charge collected in TAP1 and TAP2 at the end of the exposure, respectively. Fig. 9(b) shows that the DMP pixel array can achieve an average tap contrast of more than 96% for a subexposure speed of 39 kHz. The pixel array has more than 99% mean tap contrast at half of that subexposure speed. A histogram of the tap contrast of all the pixels in the array at 39 kHz subexposure rate is shown in Fig. 9(c). Fig. 9(d) shows a zoomed-in view of the

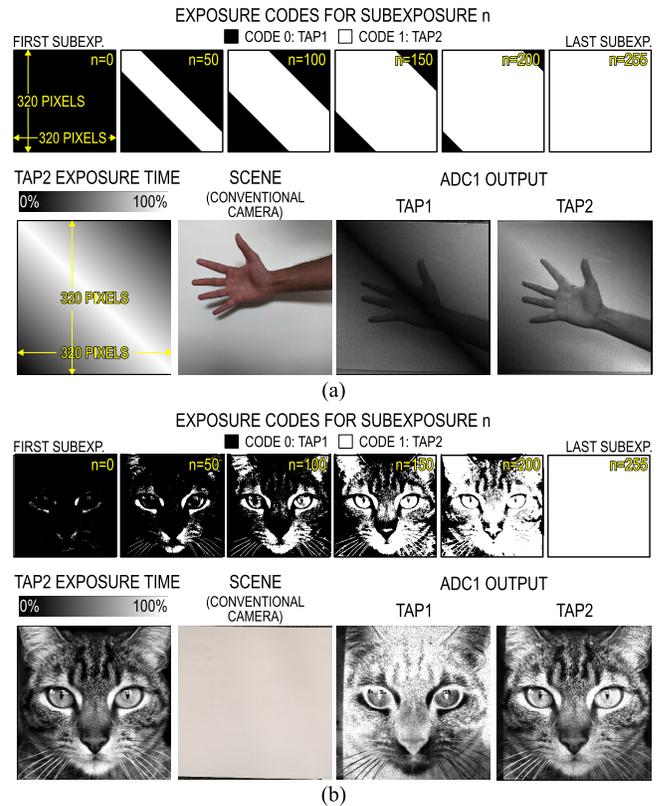


Fig. 11. Coded-exposure imaging experimental results captured with different exposure codes. (a) Analytically generated codes. (b) Codes generated from an arbitrary image.

contrast distribution near the value of 1, with the y-axis scaled logarithmically. This distribution of contrast may be attributed to a small amount of photogenerated charge getting trapped under some of the transfer gates due to process variations.

2) $\Delta\Sigma$ -Modulated ADC1: Fig. 10 shows the fast Fourier transform (FFT) of the ADC1 output with dc input signals measured at a sampling frequency of 32 MHz. The pixel-tap output voltage ranges from 1.2 to 2.5 V, and over this input range, ADC1 maintains the minimum SNR of 63 dB corresponding to 10.1 effective number of bits (ENOB) in the digital conversion.

The mean output-referred read noise of the readout path in the sensor was 23 DN, and the mean full well capacity (FWC) of 3642 DN was measured for each tap across the entire pixel array. As a result, the native dynamic range of the sensor is 44 dB per tap per pixel.

B. Validation in Applications

First, we demonstrate the coded-exposure imaging capability as a generic functionality useful for various applications, such as those requiring analytically expressed codes and codes derived from the existing images, as depicted in Fig. 11(a) and (b), respectively. Fig. 11(a) and (b), bottom row, shows examples of two simple uniformly lit scenes, one with a hand in front of a white board, and the other with a white board without any objects, for these two applications, respectively. In this experiment, we use $N = 256$ coded subexposures at 30 frames/s, where each subexposure corresponds to a

different gray level of the 8-bit 320×320 -pixel resolution pictures shown in Fig. 11(a), bottom-left, and (b), bottom-left. Each 8-bit pixel value in these pictures denotes the number of subexposures when the corresponding pixel of the presented CEP sensor receives an exposure code of 1 and collects photogenerated charge in TAP2. The binary images in Fig. 11(a), top row, and (b), top row, show exposure-code masks for the subexposures $n \in \{0, 50, 100, 150, 200, 255\}$. The resulting TAP1 and TAP2 outputs digitized using ADC1 are shown on the right side in the bottom row of Fig. 11(a) and (b) each when the scene is uniformly illuminated. Due to the dual-tap nature of the DMP, no photogenerated charge is lost during coded exposure. These results experimentally validate coded-exposure imaging with both the arbitrary and analytically derived masks. These two types of exposure codes are chosen for the following two application examples, which are discussed next: 1) single-shot scene-adaptive HDR imaging, where the exposure codes depend on the scene and cannot be analytically expressed and 2) single-shot 3-D structured-light imaging, where the exposure codes are analytically generated.

1) *Single-Shot Scene-Adaptive HDR Imaging*: In this application, the goal is fast HDR imaging. Fast HDR imaging is emerging as a key market driver, not only in the consumer segment but also in security, robotics, automotive, and other segments where light intensity in the scene changes rapidly.

There exist several conventional HDR techniques [1], [6], [7], [21], [30], [31], [32], each with its own disadvantages. As mentioned in Section I, one such technique merges multiple shots [1] taken by an LDR camera, each exposed for a different time, but this results in significant image quality degradation due to artifacts from motion or time-varying illumination. Higher end HDR image sensors exist that can perform single-shot HDR but require large HDR pixels or expensive exotic HDR pixel fabrication technologies [31], [32]. Single-photon avalanche diode (SPAD) array image sensors can also perform HDR imaging but have the disadvantages of high power, large pixels, low spatial resolution, and, for high incident light intensities, a high output data rate [6], [30].

Coded-exposure image sensors are uniquely positioned to offer fast, low-cost, low-power, and low-output-data-rate HDR imaging capabilities in well-established main-stream CIS processes. Coded-exposure image sensors can perform HDR imaging adaptively, by adapting the pixel exposure code based on the incident light intensity of that pixel, for example, to avoid its saturation. Such adaptive HDR imaging can be implemented as either a stand-alone functionality or as a means of extending the dynamic range of other coded imaging modalities. CIS implementations of adaptive coded-exposure HDR imaging have been recently reported [7], [21], but they use the previous frame's intensity to determine the current frame's exposure codes (here referred to as pseudo-single-shot HDR) and either have a non-native resolution (e.g., 16×16 -pixel subarrays per single code in [7]) or a large and slow pixel due to a large number of in-pixel transistors including PMOS devices [4], [21].

The presented coded-exposure DMP image sensor overcomes these problems. Fig. 12 shows the scene-adaptive single-shot HDR imaging flow, requiring only a single tap,

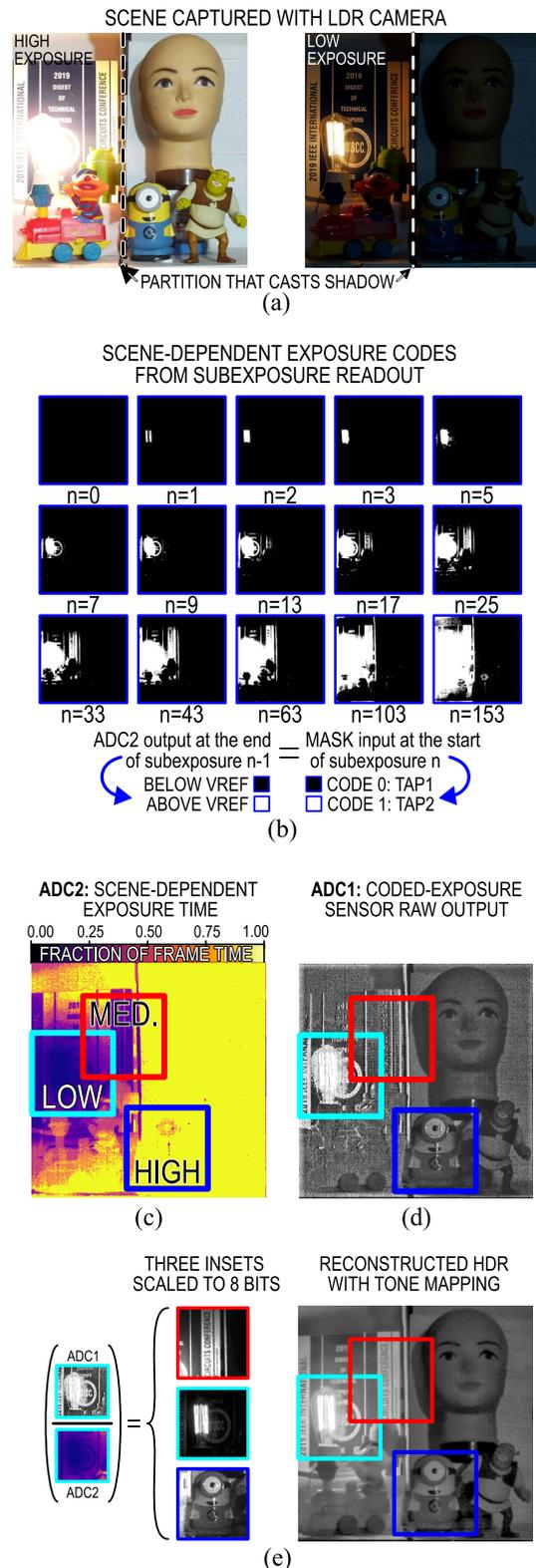


Fig. 12. Single-shot adaptive HDR imaging. (a) Scene. (b) ADC2 outputs which are used as the codes in the next subexposure. (c) Resulting per-pixel exposure time. (d) ADC1 output. (e) HDR image reconstructed by normalizing the ADC1 output by the ADC2 output, the latter comprising the per-pixel exposure time (left), and then tone-mapped for easier viewing on an LDR medium (right).

TAP1, and results captured using the combination of ADC1 and ADC2 outputs. The HDR scene captured with an LDR camera under high- and low-exposure settings is shown in

Fig. 12(a). The scene contains a partition in the middle with a bright lamp onto the left side that casts a shadow onto the right side. An LDR camera either overexposes (Fig. 12(a), left) or underexposes (Fig. 12(a), right) bright or dark elements of the scene, respectively.

As opposed to conventional HDR image sensors, the CEP sensor captures the scene in each subexposure and generates a 1-bit output image per subexposure. This output is fed back to the sensor as a code mask for the next subexposure. Fig. 12(b) shows the masks for 15 different subexposures within the frame exposure time. The mask for the subexposure $[n]$ is equal to the output of ADC2 in subexposure $[n - 1]$. To collect photogenerated charge close to the pixel-tap's FWC while also allowing for pixel-to-pixel variation, the reference voltage in ADC2 is set to 90% of the saturation level. For later subexposures, i.e., as the exposure progresses, more and more pixels' TAP1 outputs cross the reference voltage and stop integrating light any further to avoid saturating TAP1. This is done using the corresponding exposure codes to switch charge integration from TAP1 to TAP2.

Fig. 12(c) shows the per-pixel exposure time realized using the ADC2 output and the described adaptive mask control. At the end of the frame exposure time, ADC1 digitizes the raw output from the sensor, as shown in Fig. 12(d).

The HDR image is calculated by dividing ADC1 output by the per-pixel exposure time. The HDR image, tone-mapped and scaled to 8-bits to visualize it on an LDR medium, is shown in Fig. 12(e). The three insets with pixels having mostly low (cyan), medium (red), and high (blue) integration times scaled to the respective 8-bit ranges are also shown. Coded exposure, along with a combination of ADC1 and ADC2, allows capturing HDR videos at 30 frames/s.

By design, different pixels in the pixel array can have different exposure durations. As a result, it is worth pointing out that nonuniformity of motion artifacts among some or all the pixels can increase. For example, bright pixels are more motion-tolerant than dark pixels, as they have shorter exposure times. However, exposure codes for each subexposure are updated within $25.6 \mu\text{s}$ which is several orders of magnitude faster than the total exposure time (30 ms). This means that none of these exposure intervals is greater than the exposure time of a conventional pixel, so all the motion artifacts in the proposed pixel are inherently reduced when compared with conventional pixels. In fact, bright objects are of most interest in many applications, such as headlights, brake lights, and light emitting diode (LED) road signs in the case of automotive cameras, so the ability to better tolerate motion of bright objects is a clear advantage. In addition, it may be possible to correct for the pixel-to-pixel nonuniformity of motion artifacts, if needed in some special cases, using the codes used for each pixel exposure.

Compared with high-frame-rate image sensors, the power dissipation is maintained low, as only single-bit (fast) quantization is performed on each subexposure output, and one (slow) full-resolution readout is performed per frame period. High SNR is maintained, as the photogenerated charge is collected for the entire frame exposure time and is only read out once at the end of it, maintaining low read noise.

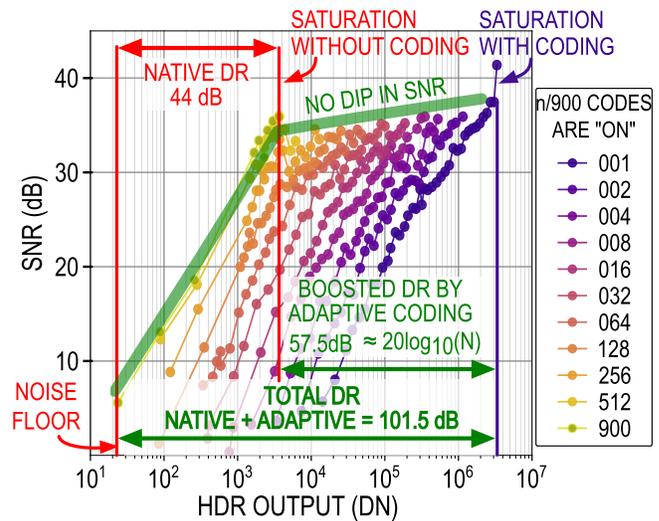


Fig. 13. Experimentally measured dynamic range and SNR of pixel output for different exposure codes.

Fig. 13 shows an experimentally measured SNR plot of pixel intensities for different exposure codes. Without coding, the sensor has a (native) dynamic range of 44 dB. With adaptive coding, the dynamic range is boosted by up to 57 dB to achieve the total dynamic range of around 101 dB. Due to the high granularity of adaptive exposure codes, we do not observe a significant SNR dip when switching between adjacent exposure codes.

2) *Single-Shot Structured-Light 3-D Imaging*: To demonstrate the versatility of the presented image sensor, we have also validated it in single-shot 3-D imaging, an application that requires two taps. Live 3-D imaging techniques and applications (e.g., bio-metric face unlock in smartphones and autonomous driving) have seen tremendous growth in the past few years due to more powerful computing resources and cheaper imaging hardware. Some of the popular methods of single-camera 3-D imaging are structured-light imaging and time-of-flight (ToF) imaging, the latter with either iToF pixels, or SPADs. Depending on the depth and accuracy requirements of an application, different 3-D imaging techniques are used. iToF cameras suffer from limited depth accuracy in short-range imaging. SPAD cameras consume higher power and can be expensive to manufacture. Therefore, structured-light imaging systems have been the method of choice for accurate short-range 3-D imaging [33]. In structured-light 3-D imaging, a projector illuminates a structured pattern of light onto the scene, and the scene's geometry distorts the pattern. In a mutually calibrated camera-projector system, the captured image of the scene with the distorted structured pattern can be reconstructed to estimate a 3-D map of the scene. The accuracy of 3-D maps can be improved when the scene is captured multiple times while illuminated with different structured-light patterns. Conventional implementations combine multiple frame readouts to generate one 3-D depth map and require the use of high-frame-rate cameras to reduce motion blur, incurring significant penalties in terms of performance and cost as described in Section I.

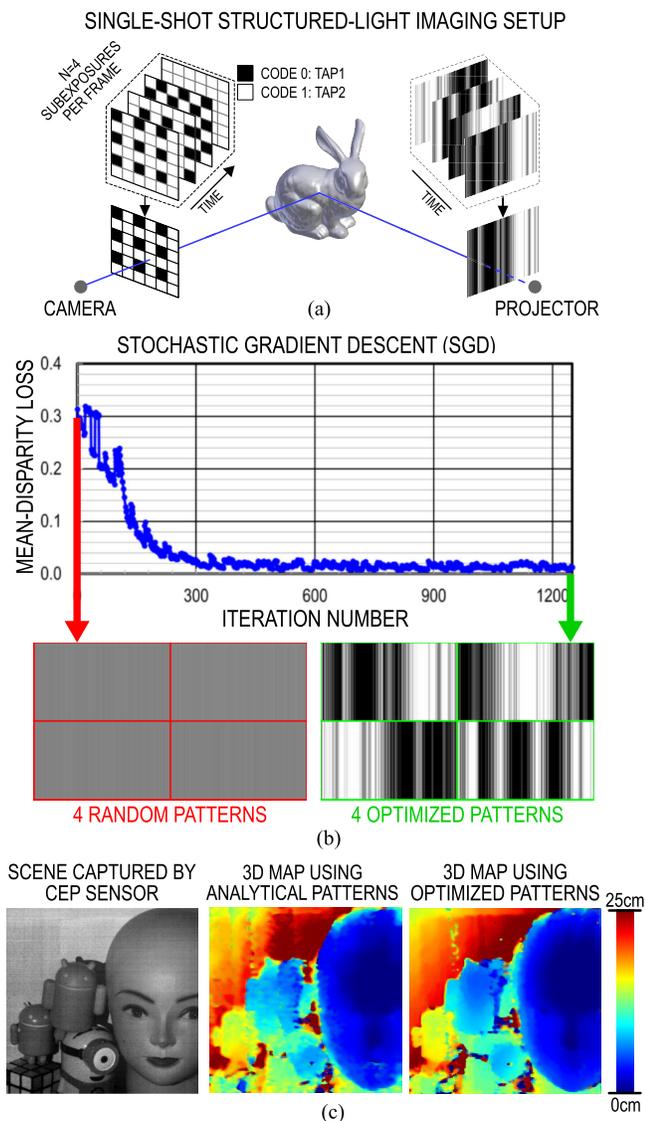


Fig. 14. Single-shot optimal structured-light 3-D imaging. (a) Experimental setup. (b) SGD results for projected illumination pattern optimization. (c) Three-dimensional imaging results without (middle) and with (right) SGD-optimized projected illumination patterns demonstrating a significant improvement in fidelity.

We demonstrate 3-D imaging performed in a single shot (i.e., within one frame period), with four illumination patterns using the presented CEP sensor. The single-shot approach generates one 3-D depth map per frame and reduces the motion blur similar to high-frame-rate cameras, but without the penalties associated with them. Fig. 14(a) depicts the principle of operation. We program our sensor to have four Bayer-like mosaic pattern exposure codes in four subexposures in a single frame. The projector is synchronized with the camera and, over the same four subexposures, projects four illumination patterns, which are optimized using optical stochastic gradient descent (SGD) [14]. In this application, the camera operation has two phases. In the initial calibration phase, optical SGD is performed only once to optimize illumination patterns. In the second phase, the optimized illumination patterns are used to perform the single-shot 3-D imaging at native video rate. As long as the relative position between the camera and the

projector is undisturbed, the first phase can be skipped, and the same illumination patterns can be used.

In the optical SGD method, to find the optimal illumination patterns, we start by projecting a random set of four illumination patterns. The scene is captured at a video rate of 30 frames/s, and the two coded-exposure images, one for each tap, readout at the end of a single frame are demosaiced and demultiplexed [13] to generate four images each corresponding to the same scene illuminated with a different structured-light illumination pattern. These four images are used to find the disparity map (which includes depth information) of the scene and compute the mean-disparity error with respect to the ground truth. Minor variations are introduced in these patterns to minimize the error and obtain the optimal set of patterns, as shown in Fig. 14(b).

The set of patterns optimized using this method is scene-agnostic, and the patterns are optimized considering fixed noise sources in the projector (e.g., nonuniform projection patterns) and the camera (e.g., columnwise fixed-pattern-noise, lens distortion) system. Fig. 14(c) (right) shows the improved single-shot 3-D map captured using the four learned optimal illumination patterns compared with the 3-D map captured with the four analytically generated illumination patterns [4] in Fig. 14(c) (middle).

IV. DISCUSSION

A comparison to the state-of-the-art coded-exposure image sensors is given in Table I. This table compares this work with the most recent sensors, which offer spatio-temporal [3], [4], [5], [6], [7] or temporal-only [8] coded exposure. Compared with the existing coded-exposure sensors that offer per-pixel coded exposure [3], [4], [5], our image sensor's DMP achieves the smallest pixel pitch of $7 \mu\text{m}$. In the presented sensor, the pixel pitch was mainly constrained by the lack of micro-lenses and the need for a reasonable FF. The pixel pitch can be further improved by any combination of micro-lenses, lower technology node, dense pixel-level 3-D interconnect, smaller photodiodes, and backside-illuminated technology. The dual-tap pixel architecture ensures no light is lost while maintaining high tap contrast—96.8%, at the highest reported subexposure speed—39 000 kHz, and with the highest spatial resolution— 320×320 pixels. The small pixel pitch and the high resolution are enabled by an all-NMOS implementation without large in-pixel PMOS circuits. The sensor can receive up to 4-Gigabit pixel codes per second. The sensor yields arbitrary global-shutter coded exposure across the whole array or within a region of interest and offers dual-ADC readout that offers both high-speed and high output resolution.

The table also shows a comparison with coded-exposure sensors that offer application-specific (not arbitrary) per-pixel coding [6], coding per larger, 16×16 pixels subarrays [7], or array-wide coding [8]. Even in this broader group of sensors, the presented sensor outperforms others in terms of subexposure rate and coding rate. The high subexposure rate and coding rate allow scene interrogation at a faster rate and with more patterns reducing artifacts due to rapidly changing incident light. The presented architecture relies on row-by-row scanning to update exposure codes, resulting in a subexposure

TABLE I
COMPARISON WITH STATE-OF-THE-ART CODED-EXPOSURE SENSORS

	THIS WORK	[3] UBC JSSC 22	[4] Toronto ISSCC 19	[5] UBC OE 19	[6] Canon ISSCC 22	[7] Nikon ISSCC 21	[8] Stanford JSSC 12
CODED-EXPOSURE MODE	ARBITRARY PER-PIXEL CODING, i.e. PIXELWISE SPATIAL AND TEMPORAL CODING				PER-PIXEL HDR-ONLY	PER-SUBARRAY 16x16 PIXELS	PER-FULL- ARRAY
PIXEL							
TECHNOLOGY [nm]	110 CIS / 65 CMOS	130 CIS	110CIS	130 CIS	90 CIS / 40 CMOS	65 CIS / 65 CMOS	130 CIS
PIXEL PITCH [μm]	7 (PPD)	12.6 (PG)	11.2 (PPD)	10.2 (PG)	11.1 (SPAD)	2.7	5
FILL FACTOR [%]	38.5	38.7	45.3	41.5	100(BSI)	100(BSI)	42
NUMBER OF TAPS	2	1	2	2	1	1	2
TAP CONTRAST [%]	96.8 @ 39k sfps ²	–	99 @ 180 sfps ²	–	N/A	N/A	–
SYSTEM							
PIXEL COUNT [HxV]	320 × 320 ≈ 100k	192 × 192 ≈ 37k	244 × 162 ≈ 40k	128 × 128 ≈ 16k	960 × 960 ≈ 0.9M	4.2k × 4.2k ≈ 17.8M	640 × 576 ≈ 369k
MAX READOUT RATE [fps]	ADC1:100 / ADC2:39k	30	25	10	90	1000	–
POWER [mW]	107 CIS / 15.2 CMOS	31.5	34.5 ⁴	1.4	370	–	–
POWER FOM [nJ/pixel] ¹	11	28.5	34	8.5	4.5	–	–
CODING							
IN-PIXEL DIG. CODE MEMORY (REQUIRES PMOS)	NO	YES (1 SRAM)	YES (2 LATCHES)	YES (DRAM-LIKE)	YES (STACKED)	NO (STACKED, PER SUBARRAY)	N/A
IN-PIXEL ANA. DATA MEMORY	YES (CHARGE)	YES (CHARGE)	NO	NO	NO	NO	
SUBFRAME RATE [sfps ²]	39000	23000	180 @ 40k pixels	1.28k @ 16k pixels	370	1k @ 69k blocks	
CODE RATE [Mbps]	4000	850	7.1	21	340	69.7	
ARBITRARY CODE/ROI ³	YES/YES	YES/YES	YES/YES	YES/–	NO/–	NO/–	
FRAME-CODE SHUTTER	GLOBAL	GLOBAL	GLOBAL	ROLLING	GLOBAL	GLOBAL	

BOLD font denotes the best performance among per-pixel coded sensors 1: FoM = Power/Number of Pixels×Frame Rate 3: ROI: region of interest
Underline denotes the overall best performance 2: sfps: subframes per second 4: no on-chip ADC

rate that is inversely proportional to the number of rows in the pixel array, assuming the need for a full-frame code update. Many applications require only codes for a subset of rows to be updated relaxing this constraint on the subexposure rate [24]. Coded-exposure sensors with silicon stacking technologies with dense pixel-level 3-D interconnect [6], [7] have demonstrated scalable architectures that can update exposure codes across the entire array without the need for a rowwise access. However, the approaches in [6] and [7] offer only coding specific to a certain application, such as HDR imaging [6], or sharing an exposure code among multiple pixels in the subarray [7]. In contrast, the presented architecture offers arbitrary per-pixel coding, providing greater flexibility, and can also benefit from dense pixel-level 3-D interconnect to maintain high subexposure rate for full-frame code updates at high array pixel counts.

The sensor is showcased using two single-shot applications: adaptive HDR imaging and structured-light 3-D imaging. The adaptive single-shot HDR imaging application shows synergistic use of a combination of ADC1, ADC2, and coded exposure. It boosts the dynamic range of the sensor by 57 dB without a significant dip in the SNR when compared with [6] and [7] due to high temporal resolution of exposure codes. Compared with [7] and [21], which have a higher dynamic range, the latency of HDR imaging is limited to one subexposure time rather than one frame period. In this work, the constant VREF is used with ADC2. However, different VREF waveforms [34], [35] may lead to even better performance with respect to power, dynamic range, and SNR. In the second demonstrated application—of single-shot structured-light 3-D imaging—the learned optimal projected patterns improve the results compared with analytical/random patterns.

V. CONCLUSION

A dual-tap CEP image sensor is presented. The pipelined NMOS-only DMP reduces the transistor count to achieve

a pixel pitch of 7 μm and yields 39 000 subexposures/s at 320 × 320 sensor resolution. This work also introduces a method for ON-chip exposure code generation or decompression. The sensor is showcased using two single-shot computational imaging applications. The outputs of a 12-bit frame-rate ADC1 and a 1-bit subexposure-rate ADC2 are adaptively combined to boost the native dynamic range by over 57 dB, demonstrating an over 101-dB dynamic range in intensity imaging. The single-shot structured-light 3-D imaging with optimal patterns reduces artifacts due to rapidly changing incident light and improves the depth map accuracy.

REFERENCES

- [1] P. E. Debevec and J. Malik, "Recovering high dynamic range radiance maps from photographs," in *Proc. ACM SIGGRAPH Classes*. New York, NY, USA: Association for Computing Machinery, Aug. 2008, pp. 1–10.
- [2] R. Gulve et al., "A 39,000 Subexposures/s CMOS image sensor with dual-tap coded-exposure data-memory pixel for adaptive single-shot computational imaging," in *Proc. IEEE Symp. VLSI Technol. Circuits (VLSI Technol. Circuits)*, Jun. 2022, pp. 78–79.
- [3] Y. Luo and S. Mirabbasi, "A 30-fps 192 × 192 CMOS image sensor with per-frame spatial-temporal coded exposure for compressive focal-stack depth sensing," *IEEE J. Solid-State Circuits*, vol. 57, no. 6, pp. 1661–1672, Jun. 2022.
- [4] N. Sarhangnejad et al., "Dual-tap pipelined-code-memory coded-exposure-pixel CMOS image sensor for multi-exposure single-frame computational imaging," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019, pp. 102–104.
- [5] Y. Luo, Y. Luo, J. Jiang, M. Cai, M. Cai, and S. Mirabbasi, "CMOS computational camera with a two-tap coded exposure image sensor for single-shot spatial-temporal compressive sensing," *Opt. Exp.*, vol. 27, no. 22, pp. 31475–31489, 2019.
- [6] Y. Ota et al., "A 0.37 W 143 dB-dynamic-range 1 Mpixel backside-illuminated charge-focusing SPAD image sensor with pixel-wise exposure control and adaptive clocked recharging," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 65, Feb. 2022, pp. 94–96.
- [7] T. Hirata, H. Murata, H. Matsuda, Y. Tezuka, and S. Tsunai, "A 1-inch 17 Mpixel 1000 fps block-controlled coded-exposure back-illuminated stacked CMOS image sensor for computational imaging and adaptive dynamic range control," in *Proc. IEEE Int. Solid-State Circuits Conf. (ISSCC)*, vol. 64, Feb. 2021, pp. 120–122.

- [8] G. Wan, X. Li, G. Agranov, M. Levoy, and M. Horowitz, "CMOS image sensors with multi-bucket pixels for computational photography," *IEEE J. Solid-State Circuits*, vol. 47, no. 4, pp. 1031–1042, Apr. 2012.
- [9] J. Zhang, T. Xiong, T. Tran, S. Chin, and R. Etienne-Cummings, "Compact all-CMOS spatiotemporal compressive sensing video camera with pixel-wise coded exposure," *Opt. Exp.*, vol. 24, no. 8, pp. 9013–9024, 2016.
- [10] Y. Li et al., "End-to-end video compressive sensing using anderson-accelerated unrolled networks," in *Proc. IEEE Int. Conf. Comput. Photography (ICCP)*, Apr. 2020, pp. 1–12.
- [11] E. Vargas, J. N. P. Martel, G. Wetzstein, and H. Arguello, "Time-multiplexed coded aperture imaging: Learned coded aperture and pixel exposures for compressive imaging systems," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2021, pp. 2672–2682.
- [12] C. M. Nguyen, J. N. P. Martel, and G. Wetzstein, "Learning spatially varying pixel exposures for motion deblurring," in *Proc. IEEE Int. Conf. Comput. Photography (ICCP)*, Aug. 2022, pp. 1–11.
- [13] M. Wei et al., "Coded two-bucket cameras for computer vision," in *Computer Vision—ECCV (Lecture Notes in Computer Science)*, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham, Switzerland: Springer, 2018, pp. 55–73.
- [14] W. Chen, P. Mirdehghan, S. Fidler, and K. N. Kutulakos, "Auto-tuning structured light by optical stochastic gradient descent," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2020, pp. 5969–5979.
- [15] F. Mochizuki et al., "Single-shot 200 Mfps 5×3 -aperture compressive CMOS imager," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2015, pp. 1–3.
- [16] D. Kim et al., "A dynamic pseudo 4-tap CMOS time-of-flight image sensor with motion artifact suppression and background light cancelling over 120 klux," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 100–102.
- [17] C. S. Bamji et al., "IMpixel 65nm BSI 320MHz demodulated TOF image sensor with $3 \mu\text{m}$ global shutter pixels and analog binning," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2018, pp. 94–96.
- [18] J. Bartels, J. Wang, W. Whittaker, and S. Narasimhan, "Agile depth sensing using triangulation light curtains," in *Proc. IEEE/CVF Int. Conf. Comput. Vis. (ICCV)*, Oct. 2019, pp. 7899–7907.
- [19] H. Kubo, S. Jayasuriya, T. Iwaguchi, T. Funatomi, Y. Mukaigawa, and S. G. Narasimhan, "Programmable non-epipolar indirect light transport: Capture and analysis," *IEEE Trans. Vis. Comput. Graphics*, vol. 27, no. 4, pp. 2421–2436, Apr. 2021.
- [20] M. O'Toole, S. Achar, S. G. Narasimhan, and K. N. Kutulakos, "Homogeneous codes for energy-efficient illumination and imaging," *ACM Trans. Graph.*, vol. 34, no. 4, p. 35, 2015.
- [21] H. Ke et al., "Extending image sensor dynamic range by scene-aware pixelwise-adaptive coded exposure," in *Proc. Int. Image Sensor Workshop*, 2019, pp. 1–11.
- [22] N. Sarhangnejad et al., "Dual-tap computational photography image sensor with per-pixel pipelined digital memory for intra-frame coded multi-exposure," *IEEE J. Solid-State Circuits*, vol. 54, no. 11, pp. 3191–3202, Nov. 2019.
- [23] M. Kobayashi et al., "A $1.8e^{-6}$ temporal noise over 110-dB-dynamic range $3.4 \mu\text{m}$ pixel pitch global-shutter CMOS image sensor with dual-gain amplifiers SS-ADC, light guide structure, and multiple-accumulation shutter," *IEEE J. Solid-State Circuits*, vol. 53, no. 1, pp. 219–228, Jan. 2018.
- [24] R. Gulve et al., "Dual-port CMOS image sensor with regression-based HDR flux-to-digital conversion and 80 ns rapid-update pixel-wise exposure coding," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2023, pp. 104–106.
- [25] M. Seo et al., "2.45 e-RMS low-random-noise, 598.5 mW low-power, and 1.2 kfps high-speed 2-Mp global shutter CMOS image sensor with pixel-level ADC and memory," *IEEE J. Solid-State Circuits*, vol. 57, no. 4, pp. 1125–1137, Apr. 2022.
- [26] G. Dutta, "Column-parallel $7 \mu\text{m}$ -pitch 2^{nd} -order $\Delta - \Sigma$ ADCs for computational image sensors," thesis, Dept. Elect. Comput. Eng., Univ. Toronto, Toronto, ON, Canada, Jun. 2019.
- [27] N. Antipa, P. Oare, E. Bostan, R. Ng, and L. Waller, "Video from stills: Lensless imaging with rolling shutter," in *Proc. IEEE Int. Conf. Comput. Photography (ICCP)*, May 2019, pp. 1–8.
- [28] A. Moffat, "Huffman coding," *ACM Comput. Surveys*, vol. 52, no. 4, p. 85, Aug. 2019.
- [29] C. Bamji et al., "A review of indirect time-of-flight technologies," *IEEE Trans. Electron Devices*, vol. 69, no. 6, pp. 2779–2793, Jun. 2022.
- [30] J. Ogi et al., "A 250 fps 124 dB dynamic-range SPAD image sensor stacked with pixel-parallel photon counter employing sub-frame extrapolating architecture for motion artifact suppression," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 64, Feb. 2021, pp. 113–115.
- [31] C. Xu et al., "A stacked global-shutter CMOS imager with SC-type hybrid-GS pixel and self-knee point calibration single frame HDR and on-chip binarization algorithm for smart vision applications," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2019, pp. 94–96.
- [32] Y. Sakano et al., "A 132 dB single-exposure-dynamic-range CMOS image sensor with high temperature tolerance," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, Feb. 2020, pp. 106–108.
- [33] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," *Opt. Lasers Eng.*, vol. 106, pp. 119–131, Jul. 2018.
- [34] R. Ikeno et al., "A $4.6\text{-}\mu\text{m}$, 127-dB dynamic range, ultra-low power stacked digital pixel sensor with overlapped triple quantization," *IEEE Trans. Electron Devices*, vol. 69, no. 6, pp. 2943–2950, Jun. 2022.
- [35] S. Kim, T. Kim, K. Seo, and G. Han, "A fully digital time-mode CMOS image sensor with 22.9 pJ/frame.pixel and 92 dB dynamic range," in *IEEE Int. Solid-State Circuits Conf. (ISSCC) Dig. Tech. Papers*, vol. 65, Feb. 2022, pp. 1–3.



Rahul Gulve (Graduate Student Member, IEEE) received the B.Tech. and M.Tech. degrees in electrical engineering from IIT Madras, Chennai, India, in 2017. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Toronto, Toronto, ON, Canada.

His research interests include design and development of mixed signal systems and pixel architecture in transport aware 3-D image sensors and cameras for computational photography.



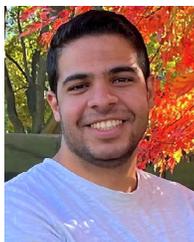
Navid Sarhangnejad (Member, IEEE) received the B.Sc. degree in electrical and computer engineering from the University of Tehran, Tehran, Iran, in 2008, the M.S. degree in electrical and computer engineering from the Delft University of Technology, Delft, The Netherlands, in 2010, and the Ph.D. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2021.

From 2011 to 2014, he was an Analog and Mixed-Signal Design Engineer with CMOSIS, Antwerp, Belgium. During this time, he worked on readout and peripheral circuits for CMOS image sensors. In summer 2016, he was a Visiting Researcher with the Integrated Radiation and Image Sensors (IRIS) Research Unit, Fondazione Bruno Kessler (FBK), Trento, Italy, where he was involved in pixel simulation and design. From 2019 to 2021, he was an Analog Design Engineer at Huawei Technologies, Toronto, working on high-speed wireline circuits for electrical and optical links. In 2021, he joined Alphawave Semi, Toronto, where he currently holds a Staff Engineer position working on clocking and high-speed circuits for wireline applications.



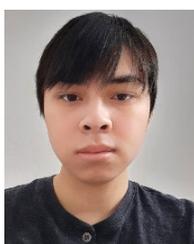
Gairik Dutta received the B.Tech. degree in instrumentation engineering from IIT Kharagpur, Kharagpur, India, in 2016, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2019.

He is currently a Senior Analog Design Engineer with Alphawave IP, Toronto. His research interests include analog and mixed-signal integrated circuits, data converters, and high-speed digital communication circuits.



Motasem Sakr received the B.Sc. degree in electronics and communication engineering from The American University in Cairo, New Cairo, Egypt, in 2019, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2021.

He is currently an ASIC Engineer with NVIDIA, Toronto. His research interests include deep learning acceleration, chip design, computer architecture, and computer vision applications.



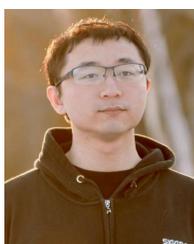
Don Nguyen received the B.A.Sc. degree in engineering physics from The University of British Columbia, Vancouver, BC, Canada, in 2019, and the M.A.Sc. degree in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2023.

His research interests include computational image sensors, semiconductor modeling/simulation, and signal and image processing.



Roberto Rangel (Graduate Student Member, IEEE) received the B.Sc. degree in electrical engineering from Rio de Janeiro State University, Rio de Janeiro, Brazil, in 2014, and the M.Sc. degree in electrical engineering from the University of São Paulo, São Paulo, Brazil, in 2019. He is currently pursuing the Ph.D. degree in electrical and computer engineering with the University of Toronto, Toronto, ON, Canada.

From 2015 to 2019, he joined the National IC Design Training Program in Brazil (IC Brazil), where he was an Instructor for analog and mixed-signal design. His research interests include low-power CMOS data communication and data conversion integrated circuits, digital systems, and CMOS image sensors.



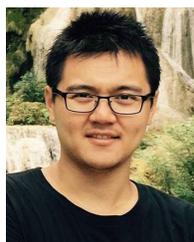
Wenzheng Chen received the bachelor's and master's degrees from Shandong University, Jinan, China, in 2014 and 2017, respectively. He is currently pursuing the Ph.D. degree with the Department of Computer Science, University of Toronto, Toronto, ON, Canada.

His research focuses on computational photography and 3-D vision.



Zhengfan Xia received the B.E. degree in electronic and information engineering from the China University of Geosciences, Beijing, China, in 2008, and the M.S. and Ph.D. degrees in information sciences from Tohoku University, Sendai, Japan, in 2011 and 2014, respectively.

From 2014 to 2017, he was a Research Engineer with the Toshiba Research and Development Center, Kawasaki, Japan, where he was involved in hardware and embedded system security. From 2017 to 2019, he was a Post-Doctoral Fellow with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada. From 2019 to 2021, he was a Hardware Engineer with Tradetone Research Laboratories, Toronto. Since 2021, he has been a Senior Digital and Systems Engineer with Stathera Inc., Toronto.



Mian Wei received the B.Sc. degree in computer science and mathematics and the M.S. degree in computer science from the University of Toronto, Toronto, ON, Canada, in 2015 and 2017, respectively, where he is currently pursuing the Ph.D. degree in computer science.



Nikita Gusev received the B.A.Sc. and M.A.Sc. degrees in electrical and computer engineering from the University of Toronto, Toronto, ON, Canada, in 2017 and 2019, respectively.

In 2020, he joined Rambus, Toronto, as an Application Engineer. Since 2021, he has been a Hardware Engineer with Alphawave IP Group, Toronto.



Esther Y. H. Lin received the B.A.Sc. degree in engineering physics from The University of British Columbia, Vancouver, BC, Canada, in 2020, and the M.S. degree in computer science from the University of Toronto, Toronto, ON, Canada, in 2022, where she is currently pursuing the Ph.D. degree in computer science.

Her research interests include computational photography, computational imaging, and optics.



Xiaonong Sun (Graduate Student Member, IEEE) received the B.A.Sc. degree in electrical engineering from the University of Toronto, Toronto, ON, Canada, in 2022, where he is currently pursuing the M.A.Sc. degree in electrical and computer engineering.

From 2020 to 2021, he was a Mixed Signal Analog Designer with Advanced Micro Devices, Markham, ON, Canada, where he was involved with IC layout. His research area includes analog and digital mixed-signal circuits, embedded and hardware systems, and computer vision.



Leo Hanxu is currently pursuing the B.Sc. degree in electrical and computer engineering with the University of Toronto, Toronto, ON, Canada.

He is particularly interested in computer graphics animation and game development, and is seeking the M.A.Sc degree in these fields.



Nikola Katic (Member, IEEE) received the B.Sc. degree in electrical engineering from the School of Electrical Engineering, University of Belgrade, Belgrade, Serbia, in 2008, and the M.Sc. and Ph.D. degrees in electrical and electronic engineering from the Swiss Federal Institute of Technology Lausanne (EPFL), Lausanne, Switzerland, in 2010 and 2014, respectively.

He joined the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, in 2016, where he was a Post-Doctoral

Fellow until 2017. From 2014 to 2021, he also held a senior analog IC design and research and development positions at Samsung Electronics, Seoul, Korea, Synopsys Inc., Toronto, and Intel Corporation. He is currently a Director of ASIC design and development with Stathera Inc, Montreal, QC, Canada, where he works on clock generation, frequency synthesis, and overall timing solutions for the new generation of communication products. His research interests include CMOS image sensor design, analog-to-digital converters (ADCs), phase locked loops (PLLs), high-speed analog and mixed-signal integrated circuit design, and signal and image processing.



Ameer M. S. Abdelhadi received the Ph.D. degree in computer engineering from The University of British Columbia, Vancouver, BC, Canada, in 2016.

Before pursuing his graduate studies, he held multiple design and research positions in the semiconductor industry. He is currently an Assistant Professor of computer engineering with the Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada. Prior to joining McMaster, he held various research and teaching fellowships at the University of Toronto,

Toronto, ON, Canada; Imperial College London, London, U.K.; and Simon Fraser University, Burnaby, BC, Canada. His research interests span multiple areas, including application-specific custom-tailored computer architecture and hardware acceleration, hardware-efficient deep learning, neurotechnology, reconfigurable computing, and asynchronous circuits.



Andreas Moshovos (Fellow, IEEE) received the Ptyhio and master's degrees in computer science from the University of Crete, Rethymno, Greece, in 1990 and 1992, respectively, and the Ph.D. degree in computer science from the University of Wisconsin–Madison, Madison, WI, USA, in 1998.

He is currently a Professor with the Electrical and Computer Engineering Department, University of Toronto, Toronto, ON, Canada. He has taught computer design at Northwestern University, Evanston, IL, USA, as an Assistant Professor from 1998 to 2000. In 2011, he was an Invited Professor with the Ecole Polytechnique de Lausanne, Lausanne, Switzerland. Since 2000, he was with the Electrical and Computer Engineering Department, University of Toronto, where he is currently a Professor. His research interests include architecting highly efficient and high-performance computing hardware.

Dr. Moshovos has served as the Program Chair for the ACM/IEEE International Symposium on Microarchitecture in 2011 and on numerous technical program committees in the area of computer architecture.



Kiriakos N. Kutulakos (Member, IEEE) received the B.S. degree from the University of Crete, Rethymno, Greece, in 1988, and the Ph.D. degree from the University of Wisconsin–Madison, Madison, WI, USA, in 1994, both in computer science.

He is currently a Professor of computer science with the University of Toronto, Toronto, ON, Canada. His research interests include computer vision, computational imaging, and 3-D sensing.

He has also been a pioneer in the area of computational light transport, developing theoretical tools, and computational cameras to analyze light propagation in real-world environments.

Dr. Kutulakos was a recipient of an Alfred P. Sloan Fellowship, a Marr Prize in 1999, a Marr Prize Honorable Mention in 2005, and five more paper awards (CVPR 1994, ECCV 2006, CVPR 2014, CVPR 2017, and CVPR 2019). He was the Program Co-Chair of CVPR 2003 and ICCV 2013, and also served as the Program Co-Chair for the Second International Conference on Computational Photography in 2010. He also served as an Associate Editor for the IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE from 2005 to 2010.



Roman Genov (Senior Member, IEEE) received the B.S. degree in electrical engineering from the Rochester Institute of Technology, Rochester, NY, USA, in 1996, and the M.S.E. and Ph.D. degrees in electrical and computer engineering from Johns Hopkins University, Baltimore, MD, USA, in 1998 and 2003, respectively.

He held engineering positions at Atmel Corporation, Columbia, MD, USA, in 1995, and Xerox Corporation, Rochester, in 1996. He was a Visiting Researcher with the Laboratory of Intelligent Sys-

tems, Swiss Federal Institute of Technology (EPFL), Lausanne, Switzerland, in 1998, and the Center for Biological and Computational Learning, Massachusetts Institute of Technology, Cambridge, MA, USA, in 1999. He is currently a Full Professor with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON, Canada, where he is a member of the Electronics Group and the Biomedical Engineering Group, and the Director of the Intelligent Sensory Microsystems Laboratory. His research interests are primarily in analog/digital integrated circuits and systems for energy-constrained biological, medical, and consumer sensory applications, such as implantable, wearable and disposable sensory microsystems, sensory-edge machine learning accelerators, and wireless sensors; and applications include brain–chip interfaces, neuro-stimulators, image sensors, and molecular biosensors.

Dr. Genov was a member of the IEEE International Solid-State Circuits Conference International Program Committee and the IEEE European Solid-State Circuits Conference Technical Program Committee. He was a co-recipient of the Jack Kilby Award for Outstanding Student Paper from the IEEE International Solid-State Circuits Conference, the Best Paper Award of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS, the Best Paper Award of the IEEE Biomedical Circuits and Systems Conference, the Best Student Paper Award of the IEEE International Symposium on Circuits and Systems, the Best Paper Award of the IEEE Circuits and Systems Society Sensory Systems Technical Committee, the Brian L. Barge Award for Excellence in Microsystems Integration, the MEMSCAP Microsystems Design Award, the DALSA Corporation Award for Excellence in Microsystems Innovation, and the Canadian Institutes of Health Research Next Generation Award. He was the Technical Program Co-Chair of the IEEE Biomedical Circuits and Systems Conference. He was also an Associate Editor of IEEE TRANSACTIONS ON CIRCUITS AND SYSTEMS—II: EXPRESS BRIEFS and IEEE SIGNAL PROCESSING LETTERS, and a Guest Editor of IEEE JOURNAL OF SOLID-STATE CIRCUITS. He is currently an Associate Editor of IEEE TRANSACTIONS ON BIOMEDICAL CIRCUITS AND SYSTEMS.