39000-Subexposures/s Dual-ADC CMOS Image Sensor with Dual-Tap Coded-Exposure Pixels for Single-Shot HDR and 3D Computational Imaging

Rahul Gulve, Navid Sarhangnejad, Gairik Dutta, Motasem Sakr, Don Nguyen, Roberto Rangel, Wenzheng Chen, Zhengfan Xia, Mian Wei, Nikita Gusev, Esther Y. Lin, Xiaonong Sun, Leo Hanxu, Nikola Katic, Ameer Abdelhadi, Andreas Moshovos, Kiriakos N. Kutulakos, and Roman Genov

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 $7 \mu m$ pitch. One frame period can include up to 900 subexposures when operating at 30 frames per second (fps), corresponding to 39,000 coded subexposures/s. The 320×320-pixel sensor features two readout modes using column-parallel analog-to-digital converters (ADC). 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 57dB, demonstrating over 101dB dynamic range in intensity imaging. In the second demonstrated application, combined with machine-learnt projected illumination patterns, the CEP camera enables single-shot structured-light 3D imaging at the native resolution and the nominal 30fps video rate.

Index Terms—CMOS image sensors, high-speed imaging systems, 3D imaging, high dynamic range imaging

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 post-processing. One well-known such example is taking several under-exposed and over-exposed images of a scene using a low-cost low-dynamic-range 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
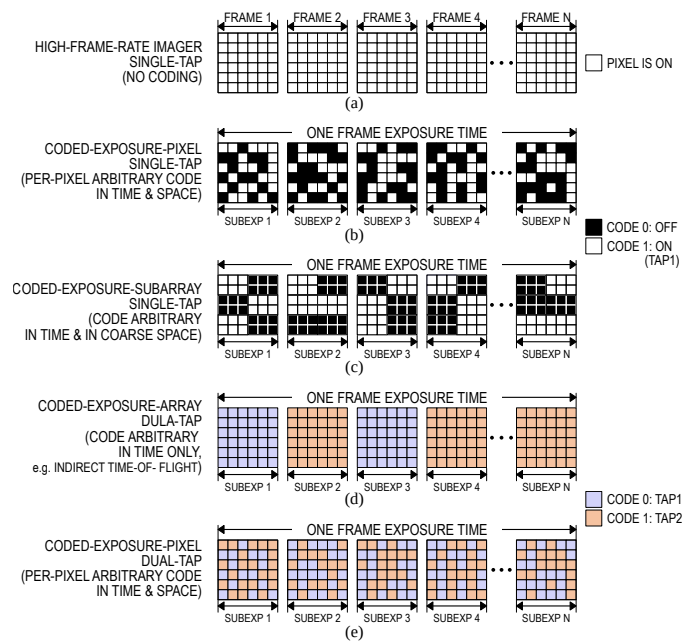
Figure 1. Overview of different exposure-coding schemes: (a) high-frame-rate cameras, (b) coded-exposure array with single-tap pixels, (c) coded-exposure subarray with single-tap pixels, (d) per-pixel coded-exposure with single-tap, and (e) per-pixel coded-exposure with dual-tap.

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]–[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]–[12], and single-shot 3D 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 Figs. 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 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 lower, 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 Figs. 1 (b, 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 by 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]–[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 pixels use the same binary code) but that offer temporal coding capability (to demodulate the input light phase in order to measure the distance to the scene) as depicted in Fig. 1 (d). This temporal-only coding is sufficient for their specific field of use - long-range, fast 3D 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]–[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
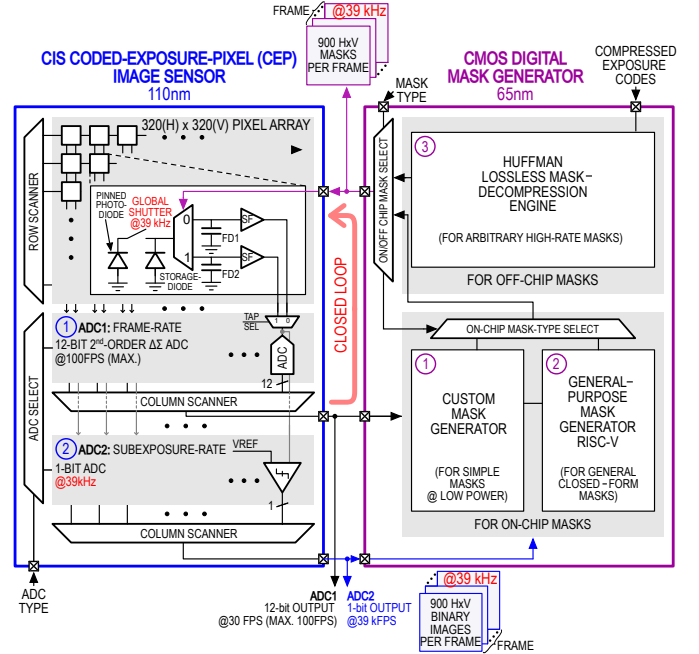

Figure 2. Single-shot adaptive coded-exposure-pixel (CEP) imaging system block diagram showing the chip architecture of the CEP image sensor IC (left) connected in a closed loop with the digital mask generator IC (right).

subexposure, so no signal is lost. Two taps also offer many additional new capabilities for fast computational imaging, such as single-shot 3D imaging featured in this work, as well as single-shot depth-gating [18] [13], [14], and single-shot direct-indirect imaging that sorts single-bounce and multi-bounce 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]–[5].

B. Data-memory pixel (DMP) Image Sensor Overview

We present a two-tap CIS image sensor comprised of ICs: a 110nm-CIS image sensor and a 65nm-CMOS mask generator, as depicted in Fig. 2. The image sensor shown in Fig. 2 (left) includes a 320x320-pixel array of dual-tap PMOS-free coded-exposure pixel, here referred to as the data-memory pixel (DMP) and a dual-ADC readout. The pixel achieves a 7 μm pitch and a subexposure rate of 39,000 subexposures per second. The compact NMOS-only implementation eliminates any cross-talk 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 coded-exposure pixel [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 100fps. 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) a 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 3D 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 in order 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 most conventional photodetectors by $20log_{10}(N)$ dB ($\approx$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 3D 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 re-configuring 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 coded-exposure-pixel (CEP) sensors. It also includes a detailed comparative analysis of 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 Sec. II-B. Section II-C describes different types of exposure codes that can be generated on-chip and their use cases.

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

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

II. VLSI Implementation

A. Dual-Tap Coded-Exposure Data-Memory Pixel

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 existing CEP image sensor pixels [3]–[5], [9] belong to the class of pixels we refer to as code-memory pixels (CMP). 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 data-memory pixel (DMP) architecture that eliminates the need for in-pixel storage of the exposure code and yields a smaller pixel pitch.

Figure 3 compares the existing pixel architectures with the presented dual-tap data-memory pixel. As shown in Fig. 3 (a, top), the conventional indirect time-of-flight (iToF) pixel has two charge collection nodes controlled by modulation signals MOD and $\overline{MOD}$ 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 coded-exposure pixels, 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], SRAM [3], or 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 pinned photodiodes.

Compared to pixels with in-pixel code memory, the conventional data-memory pixel, depicted in Fig. 3 (a, bottom), also known as the global-shutter pixel, consists of a data-memory 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 data-memory pixel, are combined to realize the presented dual-tap coded-exposure data-memory pixel, as shown in the schematic in Fig. 3 (b). By mirroring the transfer gate *TG1* of the conventional non-coded data-memory pixel, 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. Row-wise signal, *ROW_LOAD*, and column-wise 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 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 to a conventional global-shutter pixel, the transfer
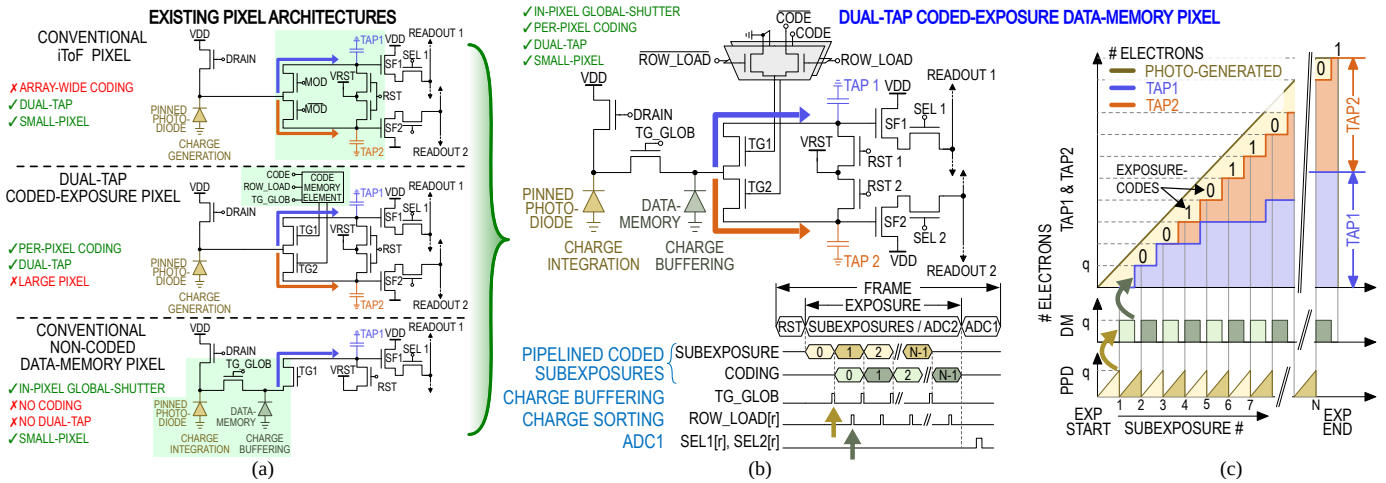
Figure 3. Comparison between existing pixel architectures and data-memory pixels. (a) Parts of existing pixel architectures similar to dual-tap coded-exposure data-memory pixel. (b) Schematic and timing diagram of the coded-exposure dual-tap data-memory pixel. (c) Amount of charge during exposure at different nodes: *TAP1*, *TAP2*, pinned photodiode *PPD*, data memory *DM*.

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 data memory across all the pixels in the array. This operation allows us to achieve the coded global-shutter exposure. The charge is stored in the data memory 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 row-wise 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 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 $10\times$ 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 $80ns$ per row or 25.6 μs per array and is repeated up to $N = 900$ times per frame at 30 fps, accounting for 10 ms ADC1 readout time. The total subexposure time of 25.6 μs translates to the subexposure rate of more than 39 kHz.

The graph in Fig. 3 (c) shows how the charge is transferred from the pinned-photodiode (PPD) to the data memory (DM), and then to one of the taps based on the exposure code. Figure 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.
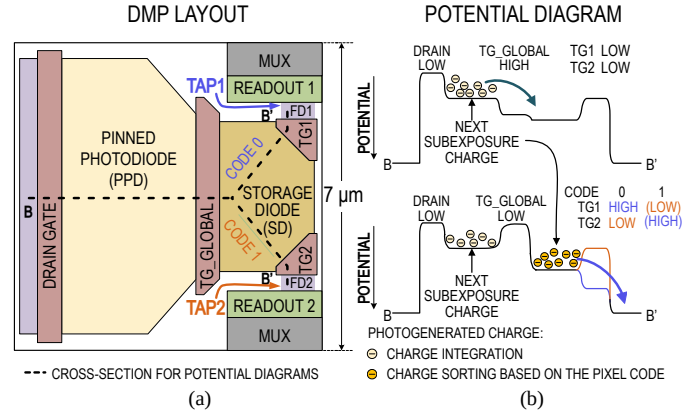

Figure 4. The dual-tap coded-exposure data-memory pixel: (a) layout, and (b) the corresponding potential diagrams during the global data sampling and charge sorting phases.

Figure 4 shows the abstract layout and potential-well diagram of the DMP pixel. Compared to the global-shutter dual-tap CMP in [3], [4], DMP eliminates PMOS transistors, reduces the transistor count, operates at a higher subexposure rate of 39,000 subexposures per second and at a higher pixel-code rate of 4 Gbps, at 320×320-pixel sensor resolution. The pixel achieves a 38.5% fill factor (FF). In the coded-exposure DMP, the data-memory 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 readout 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 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
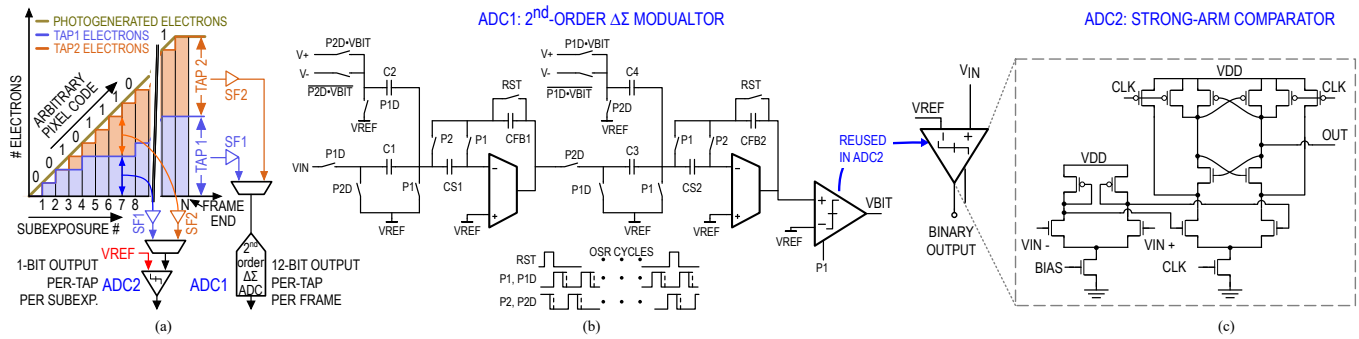
Figure 5. The operation of two ADCs, ADC1 and ADC2, within a single frame (a), and their architecture: (b) ADC1 - a 2nd-order $\Delta\Sigma$-modulated ADC, and (c) ADC2 - a strong-arm comparator with a pre-amplifier.

in-pixel MIM capacitors to sample the reset noise before exposure, allowing for CDS during readout. However, this comes at the cost of a lower fill-factor or increased pixel-pitch - micro-lenses or backside illuminated technologies can be utilized 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 (ADC). The ADCs read the (analog) amount of charge at pixel tap(s) and convert it to 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.

As compared to 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 to a digital number at the end of every frame. ADC2 is a fast subexposure-rate 1-bit comparator that compares 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 originially presented in [26]. Each ADC1 in the column-parallel bank digitizes both taps of all the pixels in its column. The data from decimation filters is transferred using on-chip serializers. ADC1 bank digitizes the data from both taps at up to 100fps while consuming 107mW of power. During the exposure period, the 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 pixel-tap voltage with a reference voltage during every subexposure. It consists of a strong-arm comparator, as shown in 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
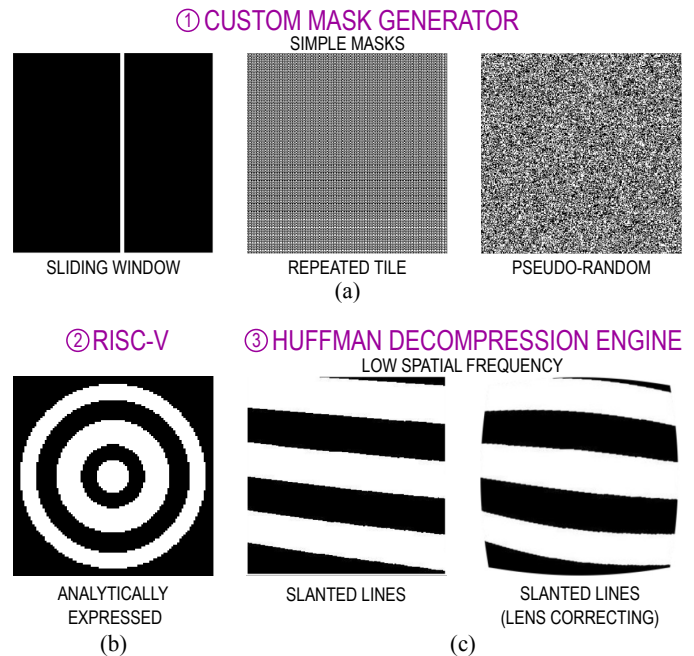

Figure 6. The on-chip mask generator can produce: (a) simple masks from a custom mask generator, (b) analytically expressed, or closed-form, masks from the RISC-V processor, and (c) masks with low spatial frequency decompressed using the Huffman decompression engine.

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 of the tap voltage during each subexposure and adjust the exposure codes based on the application.

C. Mask Generator IC

The data-memory pixel (DMP) array can receive arbitrary codes at the rate of 4 Gbps. 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 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).
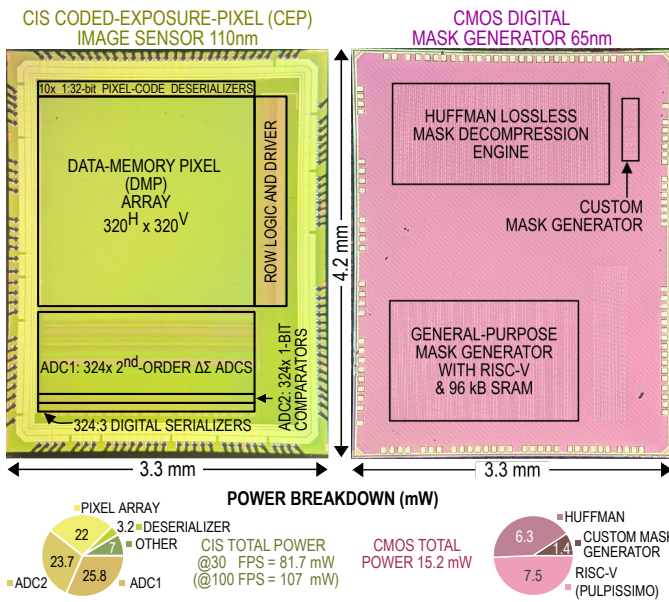
Figure 7. The micrographs and power consumption of the CEP image sensor IC (left) and the custom mask generator IC (right).
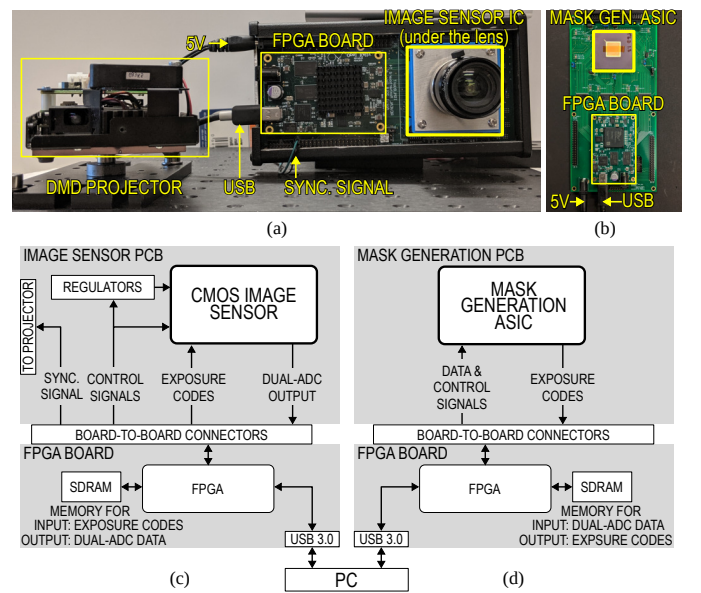

Figure 8. The experimental setup includes (a) the camera with the presented image sensor and a synchronized light-pattern projector, and (b) the mask generation system. The block diagram of (c) the camera and (d) the mask generation system is also included.

III. EXPERIMENTAL RESULTS

A. IC Characterization

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 2x2 tiles [13], [14], rolling window [18], [20], [27]; sparse scene-adaptive [7], [21], pseudo-random [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, repeated tiled patterns, or they can be pseudo-random, as shown in Fig. 6 (a). This block consists of simple sequential logic to realize the repeated and sliding patterns and a bank of pseudo-random 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.

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

Figure 8 shows the camera system used for experimental characterization. Figure 8 (a) shows the camera PCB that accommodates a CIS IC under the lens and a FPGA board, which synchronizes subexposures with a DMD projector for active-illumination computational imaging applications, such as single-shot structured-light 3D imaging demonstrated in Sec. III-B2. Figure 8 (b) shows the PCB utilized for the 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) & (d). Although both 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 utilized an FPGA to transfer mask data and control signals. Similarly, during testing of the mask generator 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 Gbps, 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 to 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.

Figure 9 (a) shows the timing diagram used to measure the tap contrast. During the measurement, all pixels receive codes
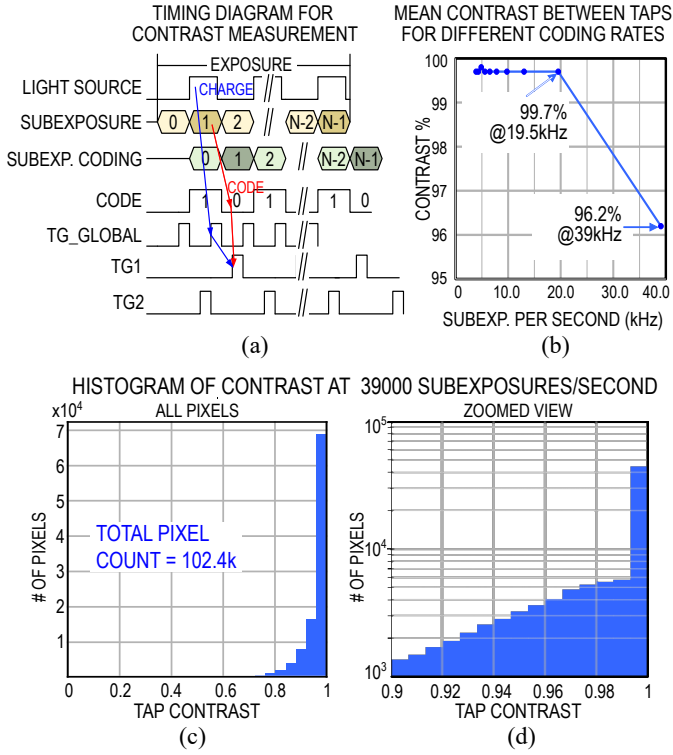
Figure 9. Experimentally measured tap contrast in coded exposure sensor: (a) timing diagram of the experimental setup, (b) mean contrast at different subexposure speeds, and (c) histogram of contrast of all the pixels and (d) zoomed in x-axis view at the highest subexposure speed of 39 kHz.
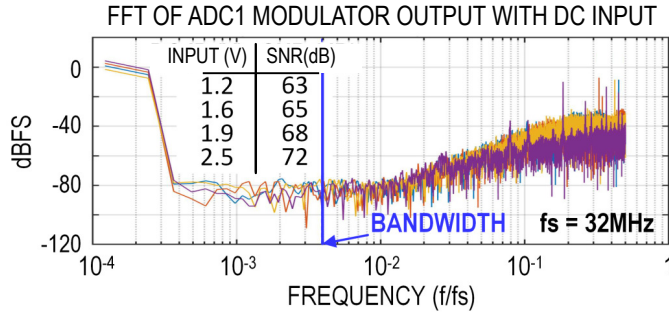

Figure 10. Experimentally measured SNR of ADC1 output for several DC inputs.

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:

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

where Q1 and Q2 are the amount of charge collected in *TAP1* and *TAP2* at the end of the exposure, respectively. Figure 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 over 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
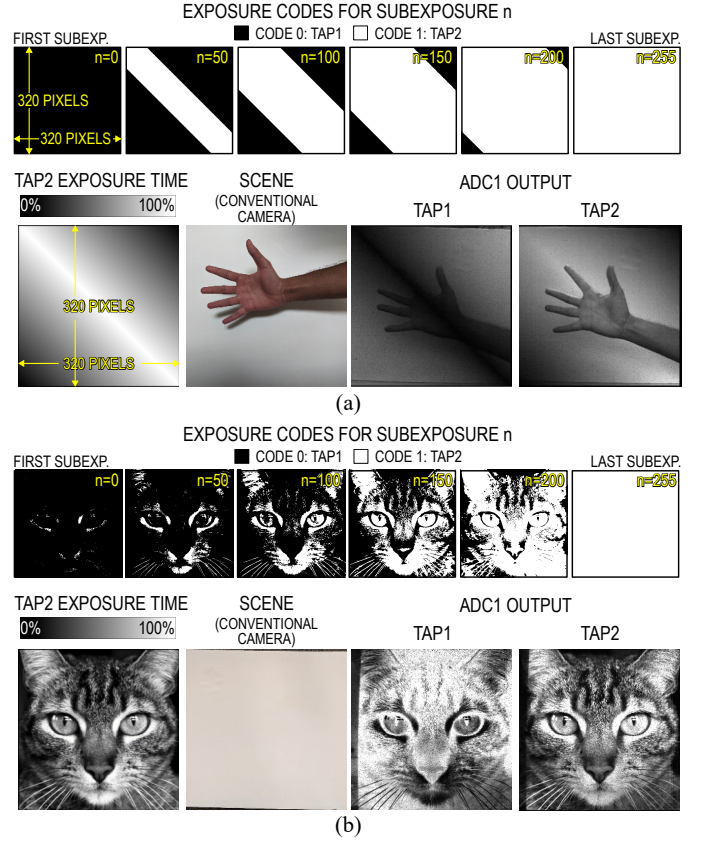

Figure 11. Coded-exposure imaging experimental results captured with different exposure codes: (a) analytically generated codes, and (b) codes generated from an arbitrary image.

in Fig. 9 (c). Figure 9 (d) shows a zoomed-in view of the 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) ΔΣ-Modulated ADC1: Figure 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 1.2V to 2.5V, and over this input range, ADC1 maintains the minimum SNR of 63dB 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 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 existing images, as depicted in Fig. 11 (a) & (b), respectively. Figures 11 (a) & (b), bottom row, show examples of two simple uniformly lit scenes, one with a hand in front of a white board, and the other with a white board

SCENE CAPTURED WITH LDR CAMERA

(a)

SCENE-DEPENDENT EXPOSURE CODES
FROM SUBEXPOSURE READOUT

n=0 n=1 n=2 n=3 n=5

n=7 n=9 n=13 n=17 n=25

n=33 n=43 n=63 n=103 n=153

ADC2 output at the end — MASK input at the start
of subexposure n-1 — of subexposure n

BELOW VREF ■ — CODE 0: TAP1 ■
ABOVE VREF □ — CODE 1: TAP2 □

(b)

ADC2: SCENE-DEPENDENT
EXPOSURE TIME

0.00 0.25 0.50 0.75 1.00
FRACTION OF FRAME TIME

MED.
LOW
HIGH

(c)

ADC1: CODED-EXPOSURE
SENSOR RAW OUTPUT

(d)

THREE INSETS
SCALED TO 8 BITS

ADC1
ADC2

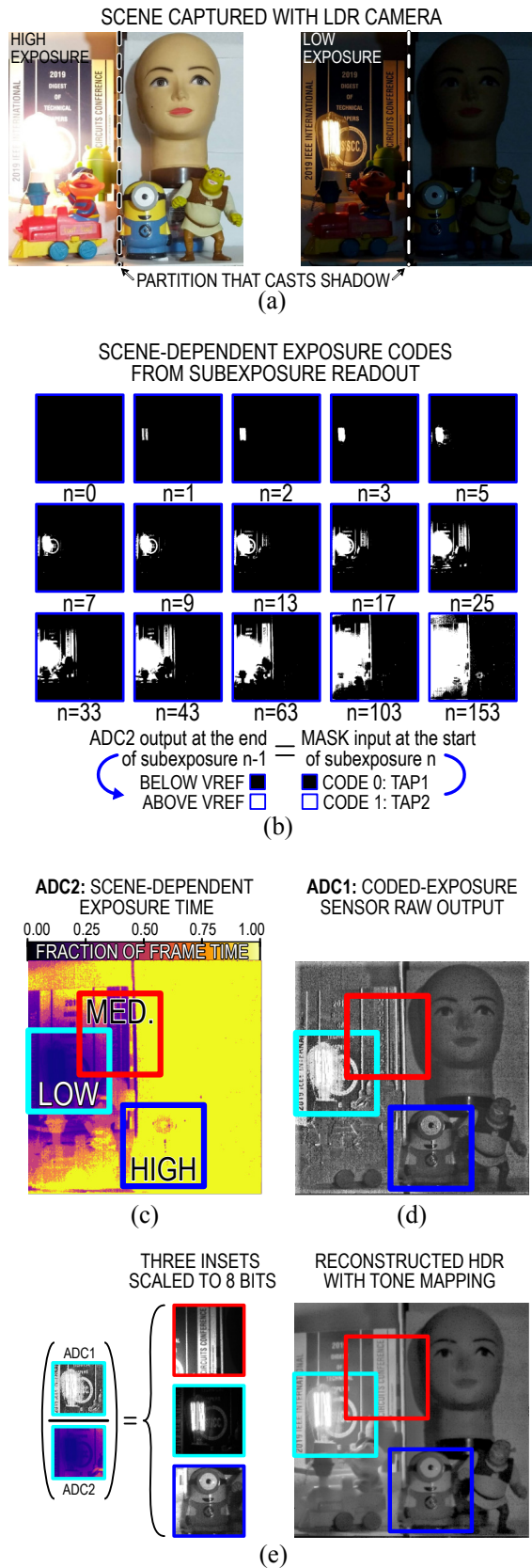RECONSTRUCTED HDR
WITH TONE MAPPING

(e)

Figure 12. Single-shot adaptive HDR imaging: (a) the scene, (b) ADC2 outputs which are utilized as the codes in the next subexposure, (c) the resulting per-pixel exposure time, (d) ADC1 output, and (e) the 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).

without any objects, for these two applications, respectively. In this experiment, we use $N = 256$ coded subexposures at 30 fps, 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 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 arbitrary and analytically derived masks. These two types of exposure codes are chosen for the following two application examples, that 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 3D structured-light imaging, where the exposure codes are analytically generated.

1) Single-Shot Scene-Adaptive HDR Imaging: In this application the goal is fast high-dynamic-range (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]–[32], each with its own disadvantages. As mentioned in Sec. I, one such technique merges multiple shots [1] taken by a low-dynamic-range 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
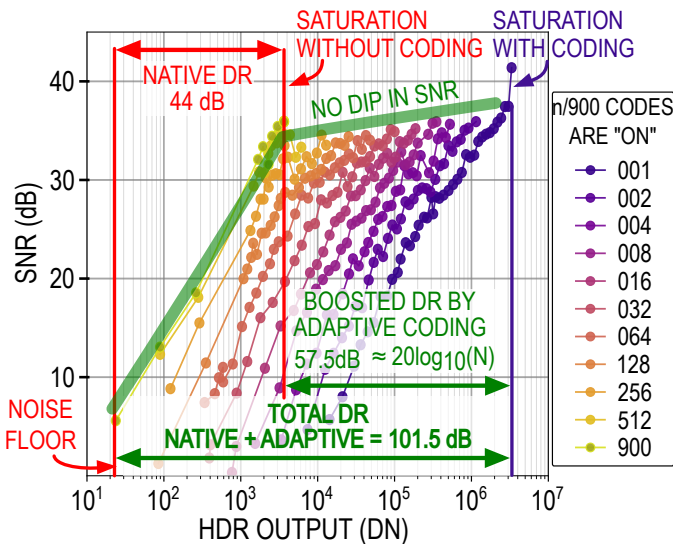
Figure 13. Experimentally measured dynamic range and SNR of pixel output for different exposure codes.

PMOS devices [4], [21].

The presented coded-exposure DMP image sensor overcomes these problems. Figure 12 shows the scene-adaptive single-shot HDR imaging flow, requiring only a single tap, *TAP1*, and results captured using the combination of ADC1 and ADC2 outputs. The HDR scene captured with a low dynamic range (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. Figure 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 full-well capacity (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 by using the corresponding exposure codes to switch charge integration from *TAP1* to *TAP2*.

Figure 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 fps.

By design, different pixels in the pixel array can have different exposure duration. As a result, it is worth pointing out that non-uniformity of motion artifacts among some or all 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 μs which is several orders of magnitude faster than the total exposure time (30 ms). This means that none of these exposure intervals are greater than the exposure time of a conventional pixel, so all motion artifacts in the proposed pixel are inherently reduced when compared to conventional pixels. In fact, bright objects are of most interest in many applications, such as headlights, brake lights and LED road signs in the case of automotive cameras, so the ability to better tolerate motion of bright objects is a clear advantage. Additionally, it may be possible to correct for the pixel-to-pixel non-uniformity of motion artifacts, if needed in some special cases, using the codes utilized for each pixel exposure.

Compared to 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.

Figure 13 shows an experimentally measured SNR plot of pixel intensities for different exposure codes. Without coding, the sensor has a (native) dynamic range of 44dB. With adaptive coding, the dynamic range is boosted by up to 57dB to achieve the total dynamic range of around 101dB. 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 3D Imaging: To demonstrate the versatility of the presented image sensor, we have also validated it in single-shot 3D imaging, an application that requires two taps. Live 3D 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 3D 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 3D imaging techniques are employed. 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 3D imaging [33]. In structured-light 3D 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 3D map of the scene. The accuracy
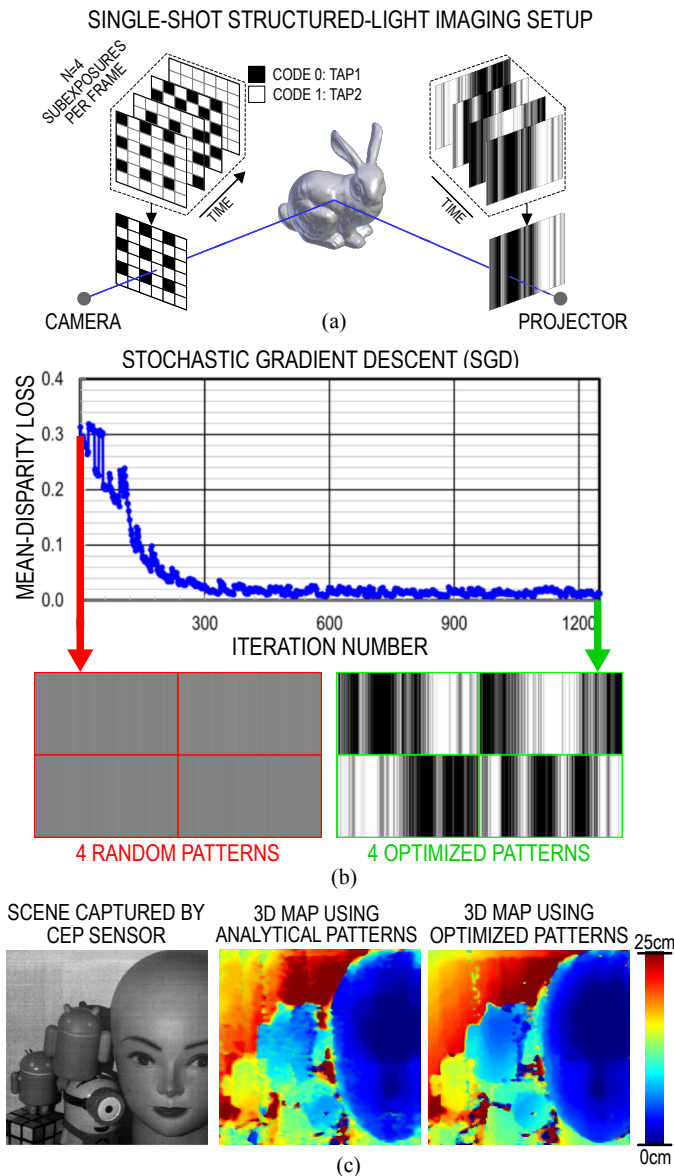
Figure 14. Single-shot optimal structured-light 3D imaging: (a) experimental setup, (b) stochastic gradient descent (SGD) results for projected illumination pattern optimization, and (c) 3D imaging results without (middle) and with (right) SGD-optimized projected illumination patterns demonstrating a significant improvement in fidelity.

of 3D 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 3D depth map, and require the use of high-frame-rate cameras in order to reduce motion blur, incurring significant penalties in terms of performance and cost as described in Sec. I.

We demonstrate 3D imaging performed in a single shot (i.e. within one frame period), with 4 illumination patterns using the presented CEP sensor. The single-shot approach generates one 3D depth map per frame and reduces the motion blur similar to high-frame-rate cameras, but without the penalties associated with them. Figure 14 (a) depicts the principle of operation. We program our sensor to have 4 Bayer-like mosaic pattern exposure codes in 4 subexposures in a single frame. The projector is synchronized with the camera and, over the same 4 subexposures, projects 4 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 3D imaging at native video rate. As long as the relative position between camera and 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 30fps, 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 4 images each corresponding to the same scene illuminated with a different structured-light illumination pattern. These 4 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., non-uniform projection patterns) and the camera (e.g., column-wise fixed-pattern-noise, lens distortion) system. Figure 14 (c, right) shows the improved single-shot 3D map captured using the four learned optimal illumination patterns compared to the 3D map captured with the four analytically generated illumination patterns [4] 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]–[7] or temporal-only [8] coded exposure. Compared with the existing coded-exposure sensors that offer per-pixel coded exposure [3]–[5], our image sensor's DMP achieves the smallest pixel pitch of 7 μm. In the presented sensor, the pixel pitch was mainly constrained by the lack of micro-lenses and the need for a reasonable fill factor. The pixel pitch can be further improved by any combination of micro-lenses, lower technology node, dense pixel-level 3D 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-39000 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.

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	<u>65 CIS / 65 CMOS</u>	130 CIS
PIXEL PITCH [μm]	**7 (PPD)**	12.6 (PG)	11.2 (PPD)	10.2 (PG)	11.1 (SPAD)	<u>2.7</u>	5
FILL FACTOR [%]	38.5	38.7	**45.3**	41.5	<u>100(BSI)</u>	<u>100(BSI)</u>	42
NUMBER OF TAPS	<u>**2**</u>	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	<u>4.2k × 4.2k ≈ 17.8M</u>	640 × 576 ≈ 369k
MAX READOUT RATE [fps]	ADC1:100 / ADC2:39k	30	25	10	90	<u>1000</u>	
POWER [mW]	107 CIS / 15.2 CMOS	31.5	34.5^4	**<u>1.4</u>**	370	–	–
POWER FOM [nJ/pixel]^1	11	28.5	34	**8.5**	<u>4.5</u>	–	–
CODING							
IN-PIXEL DIG. CODE MEMORY (REQUIRES PMOS)	<u>**NO**</u>	YES (1 SRAM)	YES (2 LATCHES)	YES (DRAM-LIKE)	YES (STACKED)	NO (STACKED, PER SUBARRAY)	
IN-PIXEL ANA. DATA MEMORY	**YES (CHARGE)**	**YES (CHARGE)**	NO	NO	NO	NO	N/A
SUBFRAME RATE [sfps^2]	<u>**39000**</u>	23000	180 @40k pixels	1.28k @16k pixels	370	1k @ 69k blocks	
CODE RATE [Mbps]	<u>**4000**</u>	850	7.1	21	340	69.7	
ARBITRARY CODE/ROI^3	**YES/YES**	**YES/YES**	**<u>YES/YES</u>**	YES/–	NO/–	NO/–	
FRAME-CODE SHUTTER	<u>**GLOBAL**</u>	**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
<u>Underline</u> denotes the overall best performance   2: sfps: subframes per second   4: no on-chip ADC

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 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 3D interconnect [6], [7] have demonstrated scalable architectures that can update exposure codes across the entire array without the need for a row-wise access. However, the approaches in [6], [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 3D 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 3D 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 as compared to [6], [7] due to high temporal resolution of exposure codes. Compared to [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 3D imaging - the learned optimal projected patterns improve the results compared to analytical/random patterns.

V. CONCLUSION

A dual-tap coded-exposure-pixel (CEP) image sensor is presented The pipelined NMOS-only data-memory pixel (DMP) reduces 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 3D 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 *ACM SIGGRAPH 2008 Classes*, ser. SIGGRAPH '08.   New York, NY, USA: Association for Computing Machinery, Aug. 2008, pp. 1–10.

[2] R. Gulve, N. Sarhangnejad, G. Dutta, M. Sakr, D. Nguyen, R. Rangel, W. Chen, Z. Xia, M. Wei, N. Gusev, E. Y. H. Lin, X. Sun, L. Hanxu, N. Katic, A. Abdelhadi, A. Moshovos, K. N. Kutulakos, and R. Genov, "A 39,000 Subexposures/s CMOS Image Sensor with Dual-tap Coded-exposure Data-memory Pixel for Adaptive Single-shot Computational Imaging," in *2022 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and 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 Journal of Solid-State Circuits*, vol. 57, no. 6, pp. 1661–1672, Jun. 2022.

[4] N. Sarhangnejad, N. Katic, Z. Xia, M. Wei, N. Gusev, G. Dutta, R. Gulve, H. Haim, M. M. Garcia, D. Stoppa, K. N. Kutulakos, and R. Genov, "5.5 Dual-Tap Pipelined-Code-Memory Coded-Exposure-Pixel CMOS Image Sensor for Multi-Exposure Single-Frame Computational Imaging," in *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, 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," *Optics Express*, vol. 27, no. 22, pp. 31475–31489, Oct. 2019.

[6] Y. Ota, K. Morimoto, T. Sasago, M. Shinohara, Y. Kuroda, W. Endo, Y. Maehashi, S. Maekawa, H. Tsuchiya, A. Abdelahafar, S. Hikosaka, M. Motoyama, K. Tojima, K. Uehira, J. Iwata, F. Inui, Y. Matsuno, K. Sakurai, and T. Ichikawa, "A 0.37W 143dB-Dynamic-Range 1Mpixel Backside-Illuminated Charge-Focusing SPAD Image Sensor with Pixel-Wise Exposure Control and Adaptive Clocked Recharging," in *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, vol. 65, Feb. 2022, pp. 94–96.

[7] T. Hirata, H. Murata, H. Matsuda, Y. Tezuka, and S. Tsunai, "7.8 A 1-inch 17Mpixel 1000fps Block-Controlled Coded-Exposure Back-Illuminated Stacked CMOS Image Sensor for Computational Imaging and Adaptive Dynamic Range Control," in *2021 IEEE International Solid- State Circuits Conference (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 Journal of 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," *Optics Express*, vol. 24, no. 8, pp. 9013–9024, Apr. 2016.

[10] Y. Li, M. Qi, R. Gulve, M. Wei, R. Genov, K. N. Kutulakos, and W. Heidrich, "End-to-End Video Compressive Sensing Using Anderson-Accelerated Unrolled Networks," in *2020 IEEE International Conference on Computational Photography (ICCP)*, Apr. 2020, pp. 1–12.

[11] E. Vargas, J. N. Martel, G. Wetzstein, and H. Arguello, "Time-Multiplexed Coded Aperture Imaging: Learned Coded Aperture and Pixel Exposures for Compressive Imaging Systems," in *2021 IEEE/CVF International Conference on Computer Vision (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 *2022 IEEE International Conference on Computational Photography (ICCP)*, Aug. 2022, pp. 1–11.

[13] M. Wei, N. Sarhangnejad, Z. Xia, N. Gusev, N. Katic, R. Genov, and K. N. Kutulakos, "Coded Two-Bucket Cameras for Computer Vision," in *Computer Vision – ECCV 2018*, ser. Lecture Notes in Computer Science, V. Ferrari, M. Hebert, C. Sminchisescu, and Y. Weiss, Eds. Cham: Springer International Publishing, 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 *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun. 2020, pp. 5969–5979.

[15] F. Mochizuki, K. Kagawa, S.-i. Okihara, M.-W. Seo, B. Zhang, T. Takasawa, K. Yasutomi, and S. Kawahito, "6.4 Single-shot 200Mfps 5×3-aperture compressive CMOS imager," in *2015 IEEE International Solid-State Circuits Conference - (ISSCC) Digest of Technical Papers*, Feb. 2015, pp. 1–3.

[16] D. Kim, S. Lee, D. Park, C. Piao, J. Park, Y. Ahn, K. Cho, J. Shin, S. M. Song, S.-J. Kim, J.-H. Chun, and J. Choi, "5.4 A Dynamic Pseudo 4-Tap CMOS Time-of-Flight Image Sensor with Motion Artifact Suppression and Background Light Cancelling Over 120klux," in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, Feb. 2020, pp. 100–102.

[17] C. S. Bamji, S. Mehta, B. Thompson, T. Elkhatib, S. Wurster, O. Akkaya, A. Payne, J. Godbaz, M. Fenton, V. Rajasekaran, L. Prather, S. Nagaraja, V. Mogallapu, D. Snow, R. McCauley, M. Mukadam, I. Agi, S. McCarthy, Z. Xu, T. Perry, W. Qian, V.-H. Chan, P. Adepu, G. Ali, M. Ahmed, A. Mukherjee, S. Nayak, D. Gampell, S. Acharya, L. Kordus, and P. O'Connor, "IMpixel 65nm BSI 320MHz demodulated TOF Image sensor with 3μm global shutter pixels and analog binning," in

[18] J. Bartels, J. Wang, W. Whittaker, and S. Narasimhan, "Agile Depth Sensing Using Triangulation Light Curtains," in *2019 IEEE/CVF International Conference on Computer Vision (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 Transactions on Visualization and Computer 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 Transactions on Graphics*, vol. 34, no. 4, pp. 35:1–35:13, Jul. 2015.

[21] H. Ke, N. Sarhangnejad, R. Gulve, Z. Xia, N. Gusev, N. Katic, K. N. Kutulakos, and R. Genov, "Extending image sensor dynamic range by scene-aware pixelwise-adaptive coded exposure," in *Proc. International Image Sensor Workshop*, 2019.

[22] N. Sarhangnejad, N. Katic, Z. Xia, M. Wei, N. Gusev, G. Dutta, R. Gulve, P. Z. X. Li, H. F. Ke, H. Haim, M. Moreno-García, D. Stoppa, K. N. Kutulakos, and R. Genov, "Dual-Tap Computational Photography Image Sensor With Per-Pixel Pipelined Digital Memory for Intra-Frame Coded Multi-Exposure," *IEEE Journal of Solid-State Circuits*, vol. 54, no. 11, pp. 3191–3202, Nov. 2019.

[23] M. Kobayashi, Y. Onuki, K. Kawabata, H. Sekine, T. Tsuboi, T. Muto, T. Akiyama, Y. Matsuno, H. Takahashi, T. Koizumi, K. Sakurai, H. Yuzurihara, S. Inoue, and T. Ichikawa, "A 1.8e rms Temporal Noise Over 110-dB-Dynamic Range 3.4 $\mu \textrm{m}$ Pixel Pitch Global-Shutter CMOS Image Sensor With Dual-Gain Amplifiers SS-ADC, Light Guide Structure, and Multiple-Accumulation Shutter," *IEEE Journal of Solid-State Circuits*, vol. 53, no. 1, pp. 219–228, Jan. 2018.

[24] R. Gulve, R. Rangel, A. Barman, D. Nguyen, M. Wei, M. Sakr, X. Sun, D. B. Lindell, K. N. Kutulakos, and R. Genov, "5.8 Dual-Port CMOS Image Sensor with Regression-Based HDR Flux-to-Digital Conversion and 80ns Rapid-Update Pixel-Wise Exposure Coding," in *2023 IEEE International Solid- State Circuits Conference - (ISSCC)*, vol. 66, Feb. 2023.

[25] M.-W. Seo, M. Chu, H.-Y. Jung, S. Kim, J. Song, D. Bae, S. Lee, J. Lee, S.-Y. Kim, J. Lee, M. Kim, G.-D. Lee, H. Shim, C. Um, C. Kim, I.-G. Baek, D. Kwon, H. Kim, H. Choi, J. Go, J. Ahn, J.-K. Lee, C.-R. Moon, K. Lee, and H.-S. Kim, "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 Journal of Solid-State Circuits*, vol. 57, no. 4, pp. 1125–1137, Apr. 2022.

[26] G. Dutta, "Column-parallel 7um-pitch 2nd-order Delta-Sigma ADCs for Computational Image Sensors," Thesis, University of Toronto, Jun. 2019.

[27] N. Antipa, P. Oare, E. Bostan, R. Ng, and L. Waller, "Video from Stills: Lensless Imaging with Rolling Shutter," in *2019 IEEE International Conference on Computational Photography (ICCP)*, May 2019, pp. 1–8.

[28] A. Moffat, "Huffman Coding," *ACM Computing Surveys*, vol. 52, no. 4, pp. 85:1–85:35, Aug. 2019.

[29] C. Bamji, J. Godbaz, M. Oh, S. Mehta, A. Payne, S. Ortiz, S. Nagaraja, T. Perry, and B. Thompson, "A Review of Indirect Time-of-Flight Technologies," *IEEE Transactions on Electron Devices*, vol. 69, no. 6, pp. 2779–2793, Jun. 2022.

[30] J. Ogi, T. Takatsuka, K. Hizu, Y. Inaoka, H. Zhu, Y. Tochigi, Y. Tashiro, F. Sano, Y. Murakawa, M. Nakamura, and Y. Oike, "7.5 A 250fps 124dB Dynamic-Range SPAD Image Sensor Stacked with Pixel-Parallel Photon Counter Employing Sub-Frame Extrapolating Architecture for Motion Artifact Suppression," in *2021 IEEE International Solid- State Circuits Conference (ISSCC)*, vol. 64, Feb. 2021, pp. 113–115.

[31] C. Xu, Y. Mo, G. Ren, W. Ma, X. Wang, W. Shi, J. Hou, K. Shao, H. Wang, P. Xiao, Z. Shao, X. Xie, X. Wang, and C. Yiu, "5.1 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 *2019 IEEE International Solid- State Circuits Conference - (ISSCC)*, Feb. 2019, pp. 94–96.

[32] Y. Sakano, T. Toyoshima, R. Nakamura, T. Asatsuma, Y. Hattori, T. Yamanaka, R. Yoshikawa, N. Kawazu, T. Matsuura, T. Iinuma, T. Toya, T. Watanabe, A. Suzuki, Y. Motohashi, J. Azami, Y. Tateshita, and T. Haruta, "5.7 A 132dB Single-Exposure-Dynamic-Range CMOS Image Sensor with High Temperature Tolerance," in *2020 IEEE International Solid- State Circuits Conference - (ISSCC)*, Feb. 2020, pp. 106–108.

2018 IEEE International Solid - State Circuits Conference - (ISSCC), Feb. 2018, pp. 94–96.

[33] S. Zhang, "High-speed 3D shape measurement with structured light methods: A review," *Optics and Lasers in Engineering*, vol. 106, pp. 119–131, Jul. 2018.

[34] R. Ikeno, K. Mori, M. Uno, K. Miyauchi, T. Isozaki, I. Takayanagi, J. Nakamura, S.-G. Wuu, L. Bainbridge, A. Berkovich, S. Chen, R. Chilukuri, W. Gao, T.-H. Tsai, and C. Liu, "A 4.6-Mm, 127-dB Dynamic Range, Ultra-Low Power Stacked Digital Pixel Sensor With Overlapped Triple Quantization," *IEEE Transactions on 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 22.9pJ/frame.pixel and 92dB Dynamic Range," in *2022 IEEE International Solid- State Circuits Conference (ISSCC)*, vol. 65, Feb. 2022, pp. 1–3.