# Flexible Spectrally-Scanning Snapshot Multispectral Imaging On Dual-Tap Coded-Exposure-Pixel CMOS Image Sensors

Roberto Rangel, Navid Sarhangnejad, Mian Wei, Rahul Gulve, Gairik Dutta, Zhengfan Xia, Nikita Gusev, Nikola Katic,
Harel Haim, Kiriakos N. Kutulakos, and Roman Genov

*Abstract*— We present a method of spectrally-scanning snapshot multispectral imaging (MSI) that employs a dual-tap coded-exposure-pixel (CEP) CMOS image sensor. A frame exposure time is divided into $N$ subexposures. During each subexposure, an arbitrarily programmable exposure code is sent to each pixel to control the integration of the photogenerated charge into one of the two taps. We employ the data-memory pixel (DMP) architecture for the CEP, which achieves the smallest pixel size or all CEP sensors. Five unique-wavelength LEDs are sequentially turned on, synchronously with five unique 2x2-pixel code tiles, and submitted to the sensor over five subexposures. The sorted photogenerated charges are read out, and five images at the five wavelengths are subsequently extracted by demultiplexing. The number of wavelengths is flexible and can be easily extended using a larger pixel tile. As a result, spectra for a scene are captured at 5 wavelengths in the visible light and NIR spectrum in a single frame, at 30 frames per second, without using a color filter array.

## I. INTRODUCTION

Image classification by modern inference methods such as deep neural networks (DNNs) has surpassed human capabilities in several applications, such as face recognition and skin cancer detection. Spectral sensors, including multispectral imaging (MSI) cameras, yield additional visual information that can further boost the image classification performance in a wide range of applications where spectral information, beyond RGB, is present.

MSI cameras operate by sensing light in a small number of spectral bands, with common applications such as aerial surveillance and crop/food inspection. Two general classes of 2-D MSI cameras exist: (1) spectrally scanning cameras and (2) non-scanning, or snapshot, cameras. The former has slow operation speed that leads to motion artifacts when the incident light is changing rapidly; and the latter suffer from high computational demands and cost. For example, a typical spectrally scanning MSI camera is realized by operating a monochrome sensor in conjunction with several optical filters or illumination sources [1], each with a different wavelength, which are used one per frame, requiring multiple frames in order to acquire one image.

We present a method of spectrally-scanning snapshot MSI that offers the best of both worlds, eliminating the disadvantages of each. It employs an image sensor with a dual-tap coded-exposure pixel (CEP) [2-6] which enables single-shot operation with low computational complexity and cost. The emerging class of CEP image sensors has already been demonstrated to offer superior performance, particularly in the presence of rapidly changing incident light, with a wide range of novel capabilities, such as single-shot compressive sensing [3], and single-shot HDR and 3D imaging [5]. In the presented method, a dual-tap CEP image sensor is utilized to perform single-shot multispectral imaging, with a programmable number of arbitrary wavelengths spanning the visible and near-infrared light spectrum.

High-speed cameras can be employed for spectrally-scanning snapshot MSI, but they have limited output frame rate and suffer from high power, high read noise and high output data rate making them very expensive and thus of limited utility [7]. To avoid some of these drawbacks, general-purpose non-integrated CEP imaging systems that could be suitable for spectrally-scanning snapshot MSI have been developed that employed digital micromirror devices (DMDs) or liquid crystals on silicon (LCoS) to either pass or block light coming to each single-tap pixel of a camera depending on a digital “code” for that pixel [8]. Such systems offer lower readout power, lower read noise, and lower output data rate, as the photogenerated charge is accumulated over multiple coded intra-frame subexposures and is read out only once per frame, as compared to high-speed cameras where a readout takes place for each exposure and thus contributes to higher

---

All authors are/were with the University of Toronto, Canada.

R. Rangel, N. Sarhangnejad, R. Gulve, G. Dutta, Z. Xia, N. Gusev, N. Katic, and R. Genov are/were with the Department of Electrical and Computer Engineering (email: roman@eecg.utoronto.ca).

M. Wei, H. Haim, and K. Kutulakos are/were with the Department of Computer Science (e-mail: kyros@cs.toronto.edu).

power dissipation, read noise and output data rate [5]. However, such non-integrated systems require bulky, expensive and distortive optical components.

CMOS-based multi-tap CEP image sensors [9] offer the additional key advantages of not only smaller form factor and lower cost due to pixel programmability directly in a CMOS Image Sensor (CIS) technology, but also better optical signal fidelity, since no external optical devices or moving parts are needed; and better light efficiency, as the photogenerated charge is sorted among multiple taps instead of being discarded when the one-tap pixel in non-integrated CEP imaging systems is "off".

## II. SYSTEM IMPLEMENTATION

The CEP image sensor employs the state-of-the-art dual-tap coded-exposure pixel architecture we refer to as the data-memory pixel (DMP) architecture. Figure 1 (left) highlights the detailed principle of operation of the DMP sensor. During the sensor's image acquisition operation, the exposure period of each frame is divided into *N* subexposures which are performed before a single readout is made. While photogeneration for the current subexposure takes place, the photogenerated charge from the previous subexposure, temporarily stored on a light-shielded pinned storage diode, referred to as the data memory, is being sorted between two taps row by row, based on each pixel's binary coefficient referred to as the exposure code, or simply the code.

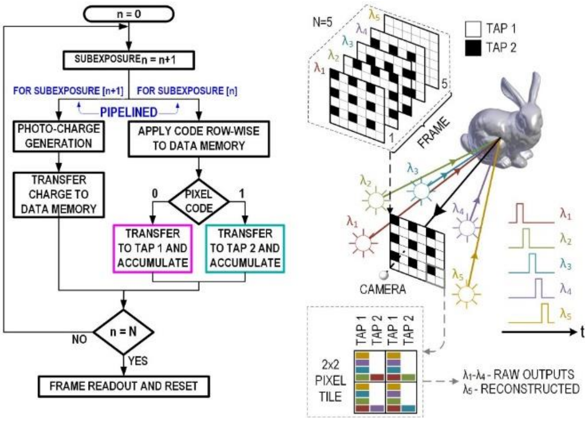

Fig. 1: The flow chart for the data-memory pixel (DMP) (left) and multispectral imaging setup (right).

Figure 1 (right) depicts the MSI camera setup which uses 5 LED lights sources, each with a unique wavelength. The 5 LEDs are sequentially turned on during the 5 respective subexposures. Synchronously with the LEDs, 5 code matrices, organized in 2x2-pixel tiles repeated over the entire pixel array, are submitted to the sensor. The sorted photogenerated charges are accumulated during the 5 subexposures and are read out once at the end of the frame as two images. From the 8 taps of each 2x2-pixel tile, 5 images at the 5 different wavelengths are then extracted by demultiplexing [9]. The number of wavelengths is flexible and can be easily extended using a larger pixel tile, at the cost of a modest reduction in the spatial resolution.

## III. SENSOR ARCHITECTURE

The DMP schematic and timing diagrams are shown in Figures 2 and 3, respectively. For each subexposure, the photogenerated light is first globally buffered on the pinned storage diode (SD), which acts as the data memory, before being sorted between taps 1 and 2, row-by-row. The exposure on the pinned photodiode (PPD) and the sorting of the SD charge are done in a pipelined fashion as desrcribed in section II.

Figure 4 (top) depicts the global exposure period, where the photogenerated charge collected at the PPD is transferred to the SD, by turning on the transfer gate by the global signal TG_GLOBAL .

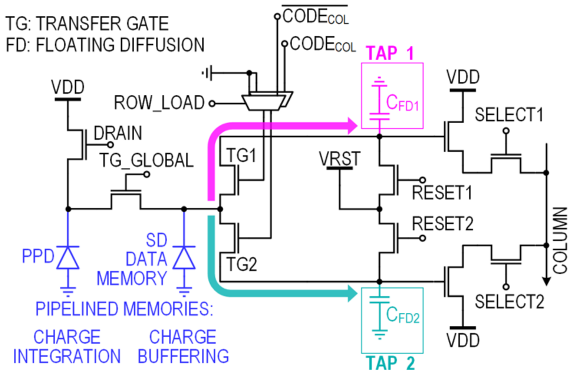

Fig. 2: Schematic of the data-memory pixel (DMP).

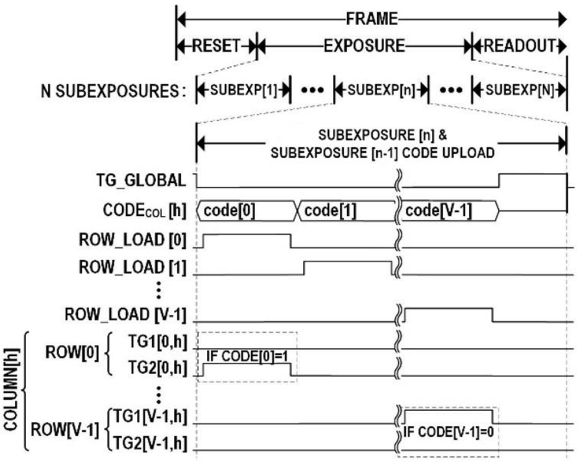

Fig. 3: Timing diagram for the data-memory pixel (DMP).

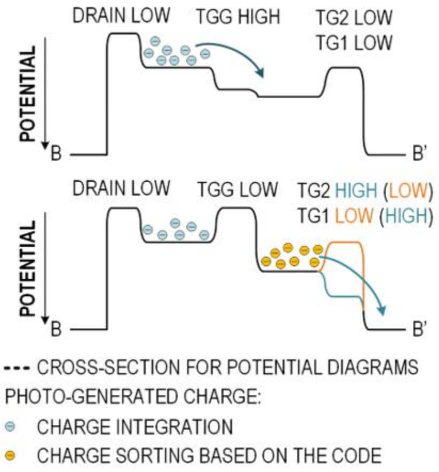

Fig. 4: Potential diagrams for the global (top) and coded (bottom) transfer of the photogenerated charge in data-memory pixel (DMP).

Figure 4 (bottom) shows the coded integration of photogenerated charges. During each subexposure, rows are accessed one after another by activating their respective ROW_LOAD terminals, then the exposure code is applied. The coded-exposure operation for that subexposure is completed with the charges from SD being transferred to taps 1 or 2 for codes 0 and 1, respectively. As a result, the photogenerated charges across all subexposures of a frame are selectively integrated over the two taps according to the per-pixel code sequence and are read out once at the end of the frame as two images. An aggregated code transfer rate of 4Gbps and charge sorting transfer time of 1µs yield a pixel code rate of 270MHz, corresponding to approximately 2700 exposures per second at 312 x 320 sensor resolution.

The compact 7µm pixel layout is shown in Figure 5. A key challenge in pixel design for CEP image sensors is the area and time overhead due to the in-pixel exposure control circuits. Most existing CEP image sensors [3-6] employ a pixel architecture we refer to as the code-memory pixel (CMP) where an in-pixel memory to store the exposure code is used, which has significant area penalty resulting in a larger pixel. In this work, a two-tap pixel 2.5x smaller than the state of the art 2-tap CMP pixel [5] is achieved. The DMP architecture eliminates the need for in-pixel storage of the exposure code or, in fact, for any PMOS transistors.

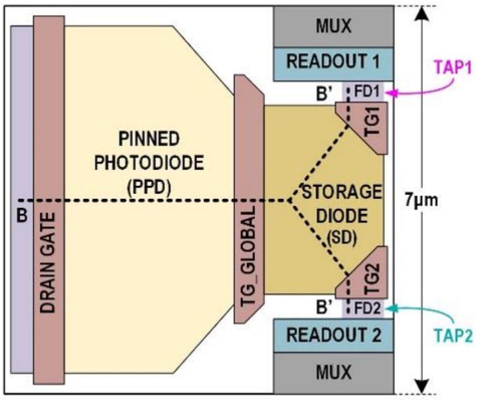

Fig. 5: Data-memory pixel (DMP) simplified layout visualization.

The top-level architecture of the sensor is depicted in Figure 6. The codes are sent to the sensor at a 4 Gbps rate and deserialized before being sent to 312 columns of pixels. This operation is done in a row-by-row fashion, taking advantage of the pixel's high-speed code transfer. The outputs of the pixels are made available by the horizontal readout scanner and serialized using 3 output channels. Figure 7 depicts the chip micrograph (left) and the camera prototype board (right). A compact FPGA board featuring Xilinx Artix 7 is used to interface with the sensor chip, providing input control signals and reading out the image data.

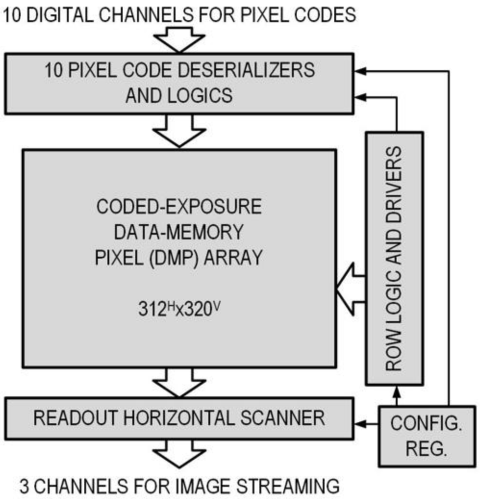

Fig. 6: Top-level architecture of the image sensor.

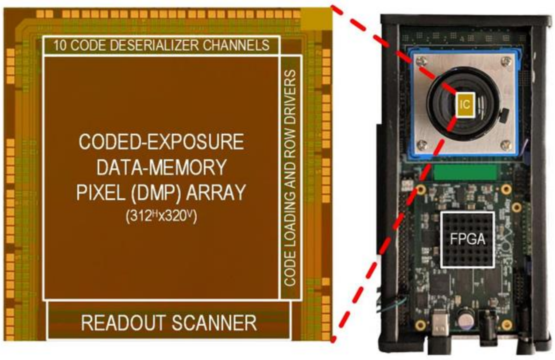

Fig. 7: Chip micrograph (left) and camera prototype (right).

## IV. EXPERIMENTAL RESULTS

The presented CEP image sensor has been employed to produce snapshot multispectral images, without using on-chip or off-chip color filters. The experimental procedure utilized the method depicted in Figure 1 (right) and described in Section II. Figure 8 depicts the imaged object - a colorful blanket (left) and the experimentally obtained output images of the camera for 5 different wavelengths (640nm to 940nm) demultiplexed from the two tap images of a single frame, at 30fps. A reconstructed RGB color image is also depicted (bottom, right). The images for the blue, green, and red wavelengths contain information from the visible light spectrum used to reconstruct the RGB color image. The NIR wavelengths are useful in making it possible to understand material surface texture and other properties.

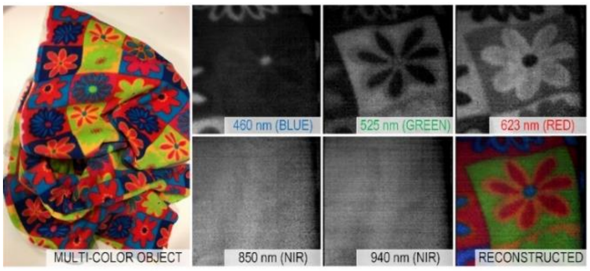

Fig. 8: Multispectral imaging experimental results at 30fps video rate. The scene on the left is captured by a conventional camera.

Table 1 summarizes the advantages of this work as compared to other CEP image sensors. The sensor achieves a 7µm pixel pitch and the maximum rate of 2700 coded subexposures per second. This subexposure rate can be used to increase the sensor frame rate when used with a small number of bands, or to trade the frame rate for a larger number of spectral bands by using larger tile sizes.

| | | THIS WORK | [3] UBC JSSC 2022 | [4] Canon ISSCC 2022 | [5] Toronto ISSCC 2019 |
|---|---|---|---|---|---|
| PIXEL | TECHNOLOGY [nm] | **110 CIS** | 130 CIS | 90 SPAD/<br>40 CMOS | **110 CIS** |
| | PINNED PHOTODIODE | **YES** | NO (PG) | NO (SPAD) | **YES** |
| | PIXEL PITCH [µm] | **7** | 12.6 | 9.5 | 11.2 |
| | FILL FACTOR [%] | 38.5 | 38.7 | **~100** | 45.3 |
| | NUMBER OF TAPS | **2** | 1 | 1 | **2** |
| | TAP CONTRAST [1] | 90 | N/A | N/A | **99.5** |
| ARCHITECTURE | PIXEL COUNT [HxV] | 312 x 320 | 192x192 | **960 x 960** | 244 x 162 |
| | FRAME RATE [fps] | 30 | 30 | **90** | 25 |
| | POWER [mW] | 54 [3] | 31.5 | 330 | 34.4 |
| | POWER FoM [nJ/frame·pixel] | 18 | 28.5 | **4** | 34 |
| SYSTEM | IN-PIXEL CODE MEMORY | **NO** | YES (SRAM) | YES (SRAM) | YES (2 LATCHES) |
| | IN-PIXEL DATA MEMORY | **YES (CHARGE)** | NO | **YES (SRAM)** | NO |
| | SUBEXPOSURE RATE [kHz] | 2.7 | 23 | **370** | .18 |
| | PIXEL CODE-RATE [MHz] | 270 | **850** | 340 | 7.1 |
| | ARBITRARY CODE / ROI [2] | **YES/YES** | **YES/YES** | NO/-- | **YES/YES** |
| | FRAME-CODE SHUTTER | **GLOBAL/ROLLING** | **GLOBAL/ROLLING** | **GLOBAL** | **GLOBAL** |
| | IMAGING APPLICATIONS | Multispectral imaging | Compressive depth sensing | HDR imaging | 1 Structured-light<br>2 Photometric stereo |

1: Also known as Extinction Ratio
2: ROI: region of interest
3: without ADC power

N/A: Not Applicable; --: Not Available
Bold font denotes the best performance

Table 1: Comparison Table.

## V. CONCLUSIONS

We have demonstrated flexible spectrally-scanning snapshot multispectral imaging on a dual-tap coded-exposure-pixel CMOS image sensor. Using the data-memory coded-exposure pixel, scene images for 5 wavelengths in the visible light and NIR spectrum obtained at are obtained within a single frame, at 30 frames per second, without using a color filter array.

## REFERENCES

[1] J.-I. Park, et al., “Multispectral Imaging Using Multiplexed Illumination,” ICCV, 2007.

[2] G. Wan, et al., “CMOS Image Sensors with Multi-Bucket Pixels for Computational Photography,” JSSC, 2012.

[3] Y. Luo, et al., “A 30-fps 192x 192 CMOS Image Sensor with Per-Frame Spatial-Temporal Coded Exposure for Compressive Focal-Stack Depth Sensing,” JSSC, 2022.

[4] Y. Ota, et al., “A 0.37W 143dB-Dynamic-Range 1Mpixel Backside-Illuminated Charge-Focusing SPAD Image Sensor with Pixel-Wise Exposure Control and Adaptive Clocked Recharging,” ISSCC, 2022.

[5] N. Sarhangnejad, et al., "Dual-Tap Pipelined-Code-Memory Coded-Exposure-Pixel CMOS Image Sensor for Multi-Exposure Single-Frame Computational Imaging," ISSCC, 2019.

[6] H. Ke, et al, “Extending Image Sensor Dynamic Range by Scene-aware Pixelwise-adaptive Coded Exposure,” IISW, 2019.

[7] L. Wang, et al., “High-speed hyperspectral video acquisition with a dual-camera architecture,” CVPR, 2015

[8] C. Xun, et al., “Computational Snapshot Multispectral Cameras: Toward dynamic capture of the spectral world,” IEEE Signal Processing Magazine, 2016-09

[9] M. Wei, et al., “Coded two-bucket cameras for computer vision,” ECCV, 2018.