A 2.5 kfps 320×320 Lite-Lens Computational Optical Imager with a 101dB Dynamic-Range SWIR FPA and Opto-Computing NPU Chipset

Chenhua Lin[1], Jie Ma[1], Tiantian Guan[1], Huosheng Wen[1], Lei Wang[2], Hongyu Che[1], Yifan Wu[1], Miao Sun[3], Rui Bai[2], Xuquan Wang[1], Yujie Xing[1], Bingbing Yao[1], Lei Qiu[1], Shenglong Zhuo[1]
[1]Tongji University, Shanghai, China; [2]LYNE Technologies, Shanghai, China; [3]University of Wisconsin-Madison, Madison, USA
Email: slzhuo@tongji.edu.cn

***Abstract*—This work presents a low-SWaP lite-lens SWIR computational optical imager for remote sensing in unmanned aerial vehicles (UAVs). Three key innovations are introduced to improve system performance. First, a complementary dual-mode (CDM) pixel and readout scheme is proposed, by interleaving the two complementary modes within one frame, the proposed imager achieves a high dynamic range (HDR) of 101 dB. Second, a pipeline column CDS readout scheme is introduced to support high-throughput acquisition, enabling a maximum raw readout rate of 2.5 kfps at 135 mW. Third, a task-specific Opto-Computing Neural Processing Unit (OC-NPU) is implemented for real-time reconstruction, achieving 7.15 TOPS/W and 1.92 TOPS/mm², and delivering 60 fps reconstructed output throughput. The tightly coupled optical-sensor-computing system achieves a PSNR of 37.12 dB and an SSIM of 0.9516.**

Keywords*—SWIR image sensor, computational optical imaging, multi-mode image processing, high dynamic range, near-sensor processing, neural reconstruction.

I. Introduction

Short-wave infrared (SWIR) imaging has been widely used in UAV-based remote sensing due to its capability for material identification and its robustness under degraded visibility conditions [1-2]. These systems are typically battery-powered and payload-limited, imposing strict constraints on size, weight, and power (SWaP), therefore requiring low-SWaP and high dynamic range (HDR) imaging solutions. To meet SWaP constraints, lightweight single-lens systems are often adopted, which introduce optical degradation and result in spatial information loss. Techniques such as DCG HDR are already being used in high-end CIS products [3-5]. In [5], readouts from HCG mode and LCG mode are merged by a huge computation power HDR merger in an ISP, such ISP-based fusion is less effective in recovering spatial information loss caused by defocus blur, limiting its application in lightweight single-lens imaging systems. To compensate for information loss introduced by front-end, image reconstruction techniques have been explored [6-8]. However, these approaches typically rely on general-purpose processors such as FPGAs, CPUs, and GPUs, leading to increased latency, high power consumption, and substantial data movement, which limit real-time operation.

To address these limitations, we present a low-SWaP HDR SWIR computational optical imaging system (Fig.1 and Fig. 2). Three key innovations are introduced. First, to achieve HDR under optical degradation, a CDM pixel and readout scheme is proposed, where a high-sensitivity mode is used for low-light conditions and a wide-dynamic-range (WDR) mode is used for bright conditions. In addition, a reconfigurable AC-coupled inverter-based CTIA is introduced, which suppresses noise in the high-sensitivity mode, while reusing the coupling capacitor for charge integration in WDR mode. Second, to mitigate frame rate degradation and large fixed-pattern noise (FPN) and reset noise caused by CDM operation, an in-pixel super source follower (SSF) is adopted to improve column bus settling, and a pipeline column CDS readout is employed to overlap acquisition and readout. Third, to enable power and area efficient real-time reconstruction and fusion, a task-specific OC-NPU is introduced, which is trained on CDM dataset to compensate for optical blur and perform HDR fusion.

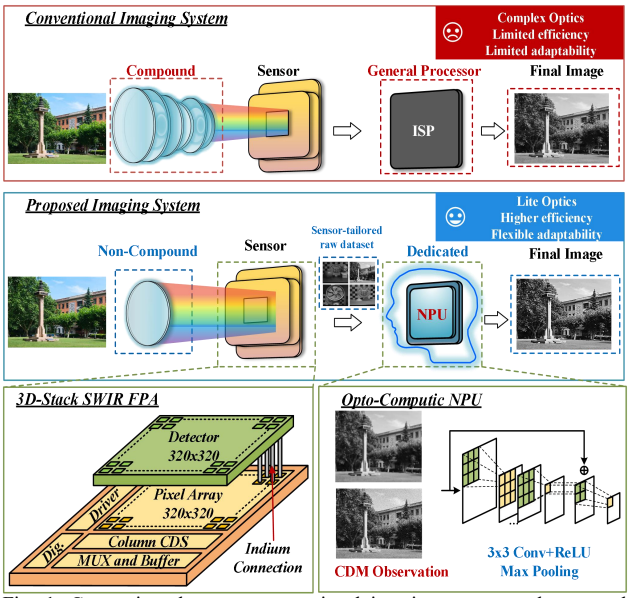

Fig. 1. Comparison between conventional imaging system and proposed low-SWaP computational optical imager.

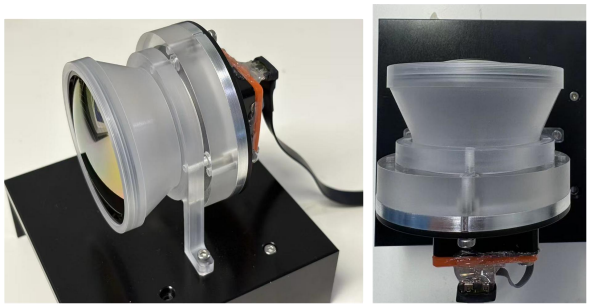

Fig. 2. Proof of the lite-lens SWIR computational optical imager.

II. Architecture and Operation

A. Complementary Dual-Mode Sensor Architecture

Fig. 3 illustrates the SWIR FPA architecture. Under lightweight lens imaging conditions, optical degradation and limited photon flux make it difficult to achieve high dynamic range. To address this limitation, a CDM pixel architecture with a reconfigurable AC-coupled inverter-based CTIA is proposed. In high sensitivity mode, the pixel operates in a

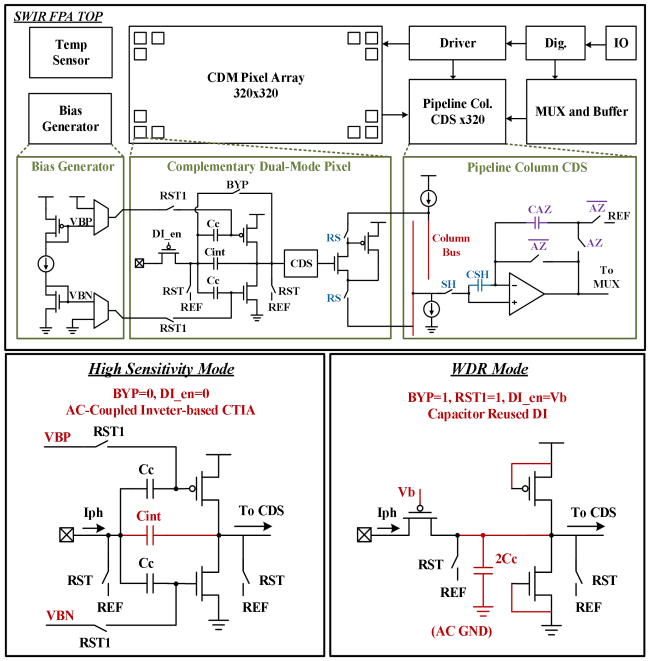

Fig. 3. Block diagram of the SWIR FPA with CDM pixel structure and pipeline column CDS.

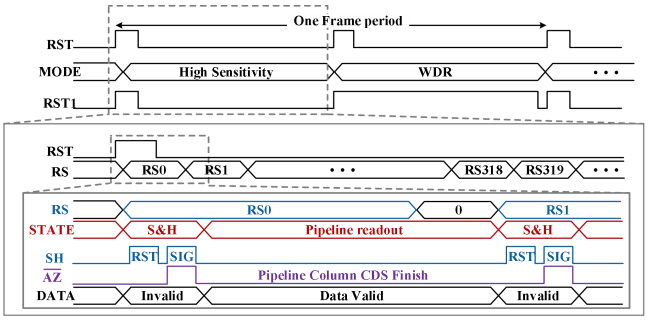

Fig. 4. Timing diagram of the SWIR FPA.

CTIA configuration, where the input node and the gates of the NMOS and PMOS in the amplifier are reset to proper bias, stabilizing the quiescent current at 25 nA. During integration, AC coupling enables both NMOS and PMOS to contribute to signal amplification, effectively reducing the read noise by $\sqrt{2}$ compared to a conventional common-source amplifier, achieving a read noise of 14 e^-_{rms}. In WDR mode, the CTIA amplifier is turned off by disabling both NMOS and PMOS, and the coupling capacitors Cc are reused as integration capacitors, forming a direct-integration (DI) configuration. With an effective capacitance of 150 fF, a full well capacity (FWC) of 1.9 Me^- is achieved, enabling low-power operation under bright conditions. The two modes are interleaved within each frame to capture complementary scene information for subsequent reconstruction and fusion, achieving a dynamic range of 101 dB.

Fig. 4 depicts the SWIR FPA timing diagram. Due to the large column-bus capacitance, row-to-row switching introduces a non-negligible settling overhead. Under multi-channel readout, the effective readout time per row becomes comparable to this overhead, reducing readout efficiency. To mitigate this issue, an in-pixel super source follower (SSF) is adopted to accelerate column bus settling, and a pipeline column CDS scheme is employed to overlap sampling and readout. During each row readout, sampling and readout are arranged in a pipelined manner across adjacent rows. While

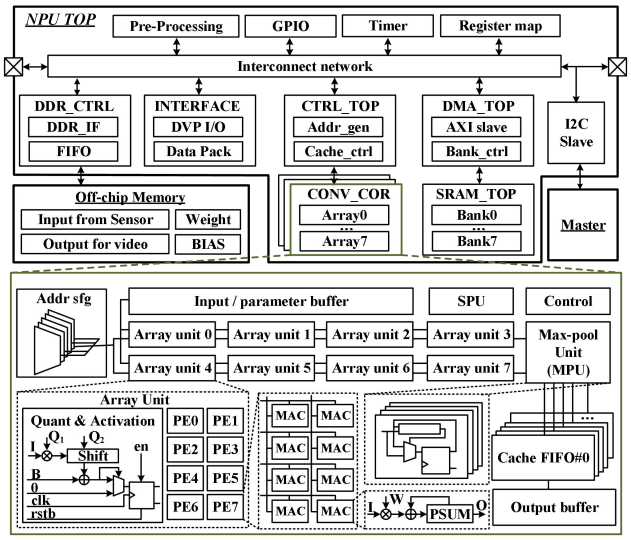

Fig. 5. Overall architecture of proposed OC-NPU with highly flexible and configurable convolution kernels.

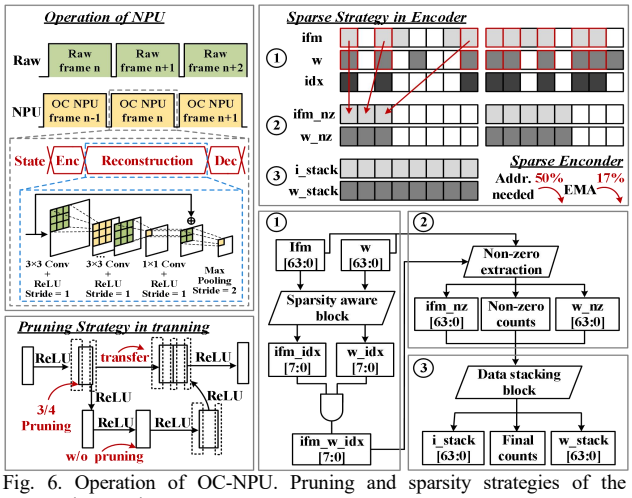

Fig. 6. Operation of OC-NPU. Pruning and sparsity strategies of the proposed network.

the current row is being sampled onto C_{SH}, the signal from the previous row is read out through the column bus. The column amplifier is configured for CDS, with its positive input connected to the top plate of C_{SH}, which suppresses reset noise and FPN from both the pixel source follower and the column circuitry. This pipelined operation reduces the effective overhead within each row period and improves readout efficiency, enabling a maximum raw readout rate of 2.5 kfps at 135 mW.

B. Opto-Computing NPU Architecture

Figs. 5 and 6 show the architecture and operation of the proposed OC-NPU. To support efficient reconstruction from interleaved dual-mode raw data and avoid excessive computation and memory overhead, the OC-NPU employs a 512-MAC convolution core with fused activation and pooling to support streaming raw-domain processing, achieving 60 fps reconstruction at 0.2 W. To further reduce redundant computation and on-chip memory traffic, a structure-aware pruning and sparsity scheme is proposed, as shown in Fig. 6. A variance-based cross-layer channel pruning method removes redundant channels while propagating the pruned topology across layers to maintain regular dataflow and avoid pipeline stalls. In addition, an OC-NPU-specific sparsity engine skips zero-valued operands

by gating MAC operations and memory accesses, reducing switching activity and memory bandwidth. As a result, the proposed scheme reduces parameter volume by 44% and energy-memory cost by 17%, with only 3% PSNR and 2% SSIM degradation.

III. Measurement Results

Fig. 7 characterizes the SWIR FPA noise performance, achieving 14 e^-_{rms} read noise, a dynamic range of 101 dB, 0.02% DSNU, and 0.16% PRNU after OC-NPU-based reconstruction and fusion, indicating that the proposed system maintains both high dynamic range and low noise characteristics. Fig. 8 shows captured sample images under different operating conditions. In the WDR mode, the system operates at a minimum power of 1.2 mW at 30 fps, while the high-sensitivity mode supports high-speed acquisition up to 2.5 kfps at 135 mW. Figs. 9 evaluate the impact of model optimization on reconstruction performance. With the proposed structure-aware pruning and sparsity scheme, parameter volume is reduced by 44% and EMA by 17%, while PSNR and SSIM degrade by only 3% and 2%, respectively. This indicates that high reconstruction fidelity can be maintained under significantly

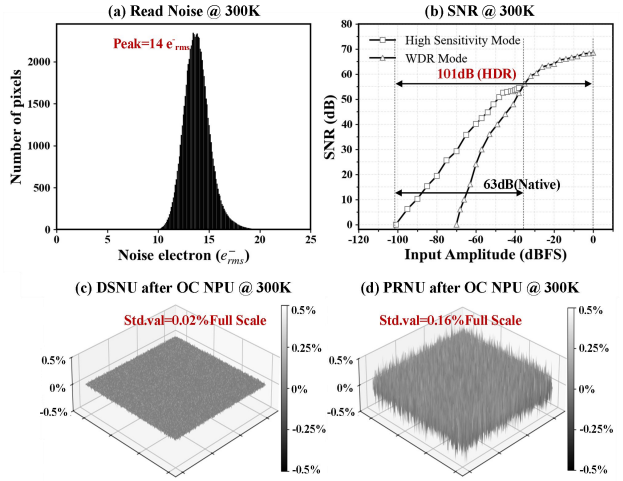

Fig. 7. Measured noise performance SWIR computational optical imager: (a) read noise. (b) HDR after OC-NPU. (c) DSNU. (d) PRNU.

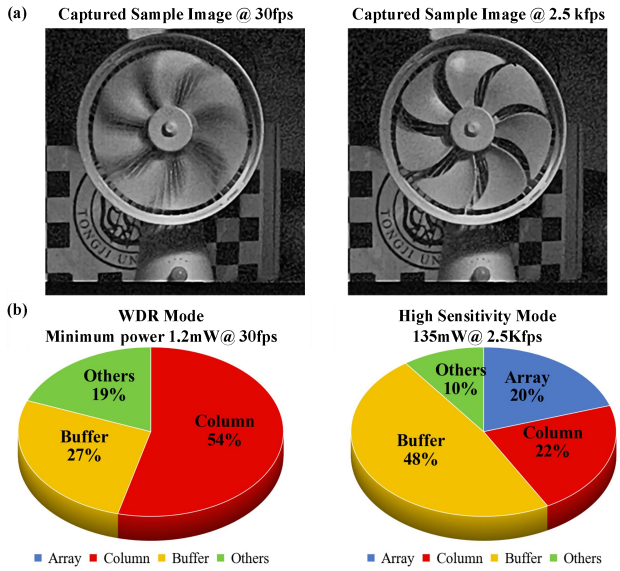

Fig. 8. (a) High frame rate of this SWIR FPA. (b) Power breakdown.

reduced model complexity. Fig. 10 demonstrates the HDR fusion results of the proposed system. Unlike conventional DCG-HDR imaging, where dual readouts are merged by a high-power ISP-based HDR merger [5], the proposed OC-NPU operates at 60 fps with a power consumption of 200 mW, enabling efficient HDR fusion suitable for low-SWaP applications. Fig. 11 shows the reconstruction results of the proposed system. The reconstructed image improves PSNR from 29.72 dB to 37.12 dB and SSIM from 0.8821 to 0.9516, approaching the image quality of compound-lens systems. Fig. 12 and Fig. 13 show the chip micrograph of SWIR sensor and OC-NPU. Table I summarizes the high-efficiency performance of the OC-NPU.

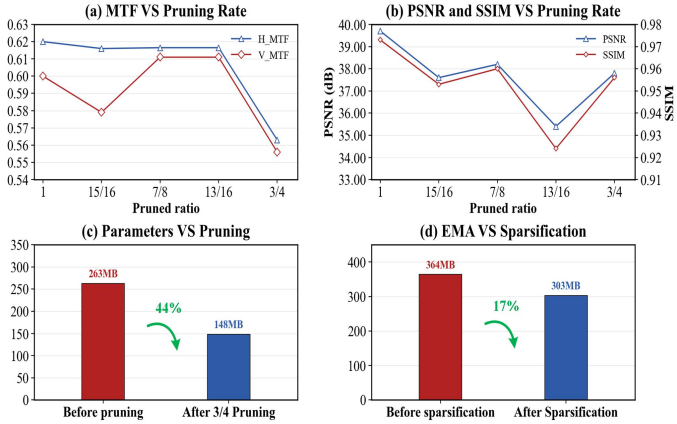

Fig. 9. (a) Horizontal and vertical MTF at different pruned ratio. (b) PSNR and SSIM at different pruned ratio. (c) Parameters before and after pruning. (d) EMA before and after sparsification.

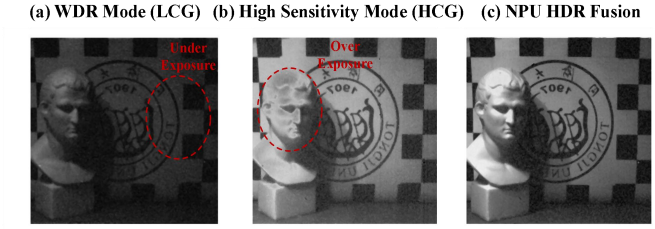

Fig. 10. HDR fusion results: (a) WDR mode, (b) high-sensitivity mode, and (c) NPU HDR fusion.

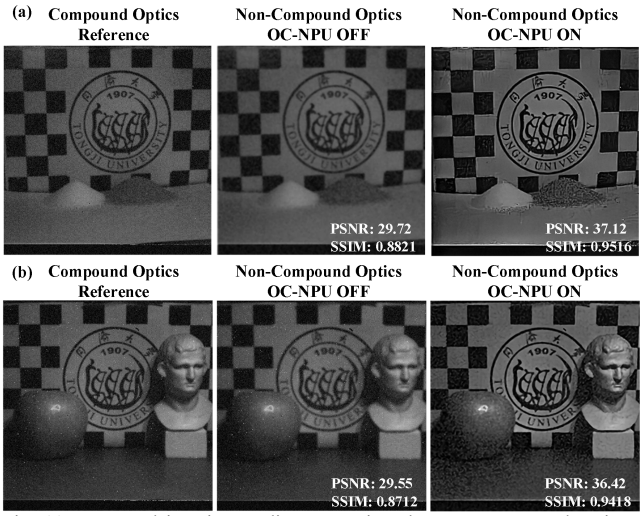

Fig. 11. Measured imaging quality comparison between compound optics and non-compound optics with OC-NPU on and off. (a) Material identification, salt and sugar. (b) Sculpture and apple scene.

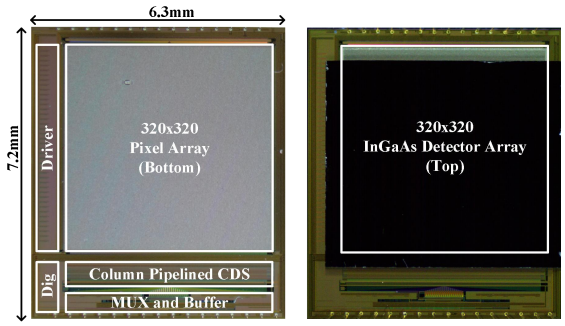

Fig. 12. Chip micrograph of CDM SWIR FPA.

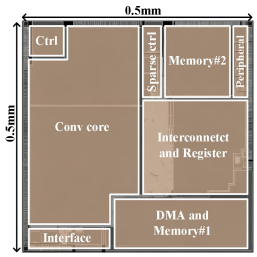

TABLE I. Performance of OC-NPU

Technology	22nm
Core Area	0.5mm x 0.5mm
Voltage	0.6-0.8V
Frequency	50-250MHz
Weight bit	Signed int8
Memory size	512Kb SRAM
System Power	89mW@50MHz 203mW@250MHz
Performance	0.48 TOPS
Power efficiency	1.98-7.15 TOPS/W
Area efficiency	1.92 TOPS/mm^2

Fig. 13. Chip micrograph and summary table of OC-NPU.

Table II presents a system-level comparison with state-of-the-art imaging systems. Compared to [5-7], the proposed system achieves the highest dynamic range of 101 dB together with chip-level real-time reconstruction and fusion, while existing approaches typically support either reconstruction or HDR merging, but not both simultaneously. This demonstrates the capability of the proposed system to jointly address dynamic range enhancement and optical degradation within a unified framework. Compared to [9], camera power and volume are reduced by 66% and 38%, respectively, enabling a more compact and energy-efficient low-SWaP imaging solution.

IV. Conclusion

In this work, a low-SWaP lite-lens SWIR computational imaging system is presented, integrating a dual-mode HDR sensor and a task-specific OC-NPU. The proposed system enables high dynamic range imaging through interleaved dual-mode acquisition, while the OC-NPU performs efficient reconstruction and fusion to compensate for optical degradation introduced by lightweight optics. Compared to prior works, it achieves a unique combination of HDR imaging and chip-level real-time reconstruction with improved energy efficiency and reduced latency. In addition, the system supports high-speed operation and significantly reduces camera power and volume, enabling a compact and efficient imaging solution under strict SWaP constraints.

Acknowledgment

This work was supported by the “National Natural Science Foundation of China” (Grant No. 62574147) and the “Natural Science Foundation of Shanghai” (Grant No. 25ZR1401365).

TABLE II. System-Level Comparison with Prior Imager

Parameter	**This Work**	**VLSI 23 [5]**	**JSSC 21 [6]**	**TIE 21 [7]**	**SPIE 14 [9]**
Sensor Front-End					
Sensor Type	IRFPA	CIS	CIS	CIS	IRFPA
DR(dB)	**63 (Native) 101(HDR)**	63(Native) 90.5(HDR)	61(Native)	68(Native)	70(Native)
Frame Rate(fps)	**30~2500 (Native) 60(HDR)**	60(HDR)	15~45 (Native)	3124 (Native)	200 (Native)
Resolution	320x320	512x320	640x480	256x256	640x512
Pitch(μm)	15	7.2	4	6.5	15
Sensor Power(mW)	1.2~ 135	97	0.258~ 0.706	56	150
Reconstruction Back-End					
Processor	**Dedicated NPU**	Off-chip ISP	Off-chip PC	Off-chip PC	N/R
Operation	**Inference reconstruct + HDR merge**	HDR merge	Iterative CS reconstruct	Iterative CS reconstruct	N/R
Throughput (fps)	60	60	N/A	N/A	N/R
Imaging System					
End-to-End Imaging System	**YES**	YES	NO	NO	YES
Peak PSNR(dB)	**37.2**	N/A	37.2	32.92	N/A
Peak SSIM	**0.956**	N/A	0.946	0.931	N/A
Real Time Processing	**YES**	YES	NO	NO	N/R
Processing Latency(ms)	**16.7**	N/R	>10400	N/R	N/A
Process(nm)	180/22	350	110	110	180
Camera Power	**0.6W @60fps**	N/A	N/A	N/A	1.8W @30fps
Camera Size w/o lens(mm^3)	**35x35x18**	N/A	N/A	N/A	32x32x35
Optics	**Single-lens**	N/R	N/R	N/R	Multi-lens

References

[1] Y. Zhong et al., "Mini-UAV-Borne Hyperspectral Remote Sensing: From Observation and Processing to Applications," in IEEE Geoscience and Remote Sensing Magazine, vol. 6, no. 4, pp. 46-62, Dec. 2018.

[2] Z. Zhang et al., "UAV Hyperspectral Remote Sensing Image Classification: A Systematic Review," in IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, vol. 18, pp. 3099-3124, 2025.

[3] Y. Sakano et al., "5.7 A 132dB Single-Exposure-Dynamic-Range CMOS Image Sensor with High Temperature Tolerance," 2020 IEEE International Solid-State Circuits Conference - (ISSCC), San Francisco, CA, USA, 2020, pp. 106-108.

[4] T. -H. Tsai et al., "6.4 A 400×400 3.24μm 117dB-Dynamic-Range 3-Layer Stacked Digital Pixel Sensor," 2025 IEEE International Solid-State Circuits Conference (ISSCC), San Francisco, CA, USA, 2025, pp. 120-122.

[5] Y. Luo et al., "A 60fps 9.9nJ/frame·pixel CMOS Image Sensor with On-Chip Pixel-wise Conversion Gain Modulation for Per-frame Adaptive DCG-HDR Imaging," 2023 IEEE Symposium on VLSI Technology and Circuits (VLSI Technology and Circuits), Kyoto, Japan, 2023, pp. 1-2.

[6] C. Park et al., "A 51-pJ/Pixel 33.7-dB PSNR 4× Compressive CMOS Image Sensor With Column-Parallel Single-Shot Compressive Sensing," in IEEE Journal of Solid-State Circuits, vol. 56, no. 8, pp. 2503-2515, Aug. 2021.

[7] H. Lee et al., "A Compressive Sensing CMOS Image Sensor With Partition Sampling Technique," in IEEE Transactions on Industrial Electronics, vol. 68, no. 9, pp. 8874-8884, Sept. 2021.

[8] Wang C et al.. "Do: A Differentiable Engine for Deep Lens Design of Computational Imaging Systems," IEEE Transactions on Computational Imaging, 2022, 8: 905-916.

[9] Selim Eminoglu et al., "A miniature VGA SWIR camera using MT6415CA ROIC," Proc. SPIE 9070, Infrared Technology and Applications XL, 90700A, June. 2014