Progress Report (2025)

Analog-Mixed-Signal Computing Microcontrollers for Artificial Intelligence of Things: Toward Real-Time Carbon Monitoring and Management

1. Personnel:

- *Mingoo Seok, Associate Professor, Columbia University*
- *Yannis Tsividis, Professor, Columbia University*
- *Rentao Wan, Graduate student, Columbia University*
- *Mosom Jana, Graduate student, Columbia University*

2. Project Period:

- Total period: Sep/1/2022 - Aug/30/2026 (Four years)
- Current: Sep/1/2024 – Aug/30/2025 (Third year)

3. Summary of the Progress in the Current Period

OASIS: Computer vision systems are essential for autonomous vehicles, supporting a wide range of tasks such as object detection, lane following, and 3D scene reconstruction with high accuracy. However, these systems face a critical bottleneck: the massive data traffic between image sensors and onboard processors. This excessive data transfer not only consumes significant energy but also imposes stringent requirements on the communication hardware, leading to high complexity and reduced efficiency.

To address this challenge, we propose OASIS, a novel computer vision framework that integrates three key components: (1) a custom image sensor with an embedded analog compression processor, (2) a state-of-the-art deep neural network (DNN) for object tracking, and (3) an adaptive closed-loop feedback controller. The core idea of OASIS is to monitor the output of the task model in real time and dynamically adjust the compression ratio of the sensor. It optimizes data reduction and task-level accuracy, thus enabling efficient yet robust system performance.

The contribution of OASIS can be summarized as follows. First, we design an online adaptive video compression scheme that observes the DNN's tracking outputs and adaptively tunes the compression level of the sensor. This ensures that compression is aggressive when task accuracy is less sensitive, while relaxing compression when task accuracy is at risk.

Second, we develop the hardware model of a 2.1-megapixel image sensor integrated with an analog compression processor, building on the AJPEG hardware prototype. The model provides detailed estimates of energy consumption, latency, and image quality under varying compression levels.

Finally, we validate the complete OASIS system on object tracking (MOT) tasks, adopting the latest YOLOX detector and ByteTrack tracker. Extensive experiments on MOT17, MOT20, and DanceTrack datasets demonstrate that OASIS achieves an average 6× reduction in data bandwidth, while incurring less than 1.5% degradation in real-time tracking performance compared with a no-compression baseline.

AnaRefresh: Analog computing has emerged as a promising solution for artificial intelligence (AI) hardware. However, prior works require frequent data conversions between analog and digital domains, which quickly undermines the efficiency benefits of analog computing. To mitigate this challenge, recent works have proposed employing capacitor-based analog memory circuits, which can store intermediate data in an analog form. Unfortunately, the capacitor-based analog memory cells leak and lose information over time. As a result, AI tasks employing analog dataflow must finish computing before their analog memory loses significant information, severely limiting the size of tasks that can be performed. Simple read-&-rewrite refresh scheme fails to compensate for accumulated cell voltage drift.

To address these issues, we propose AnaRefresh, a capacitor-based analog memory macro featuring a periodic digital-sample-analog-refresh (DSAR) technique, which is the first reliable in-situ refresh technique for capacitor-based analog memory to the best of our knowledge. Thanks to the proposed scheme, the retention time of the analog memory is extended to support any workload.

We prototype AnaRefresh in 28-nm CMOS. It consists of 64-by-64 cells, each of which can store 6-bit information. AnaRefresh can successfully refresh all of its 4096 cells, with an energy overhead of 24.6 pJ and a refresh time overhead of 2.82%. AnaRefresh takes 8 ns and consumes 3.25 pJ to read a row and 18 ns and 16.62 pJ to write. AnaRefresh attains a state-of-the-art refresh FoM even when compared against eDRAM-based digital memory. It also achieves a cell density of 0.207 μm^2/bit, which is amongst the highest for 28-nm CMOS memory, and a low macro standby power of 16.95 μW/Mb.

The operation of AnaRefresh can be summarized as below.

(1) READ: Voltages stored in C_{MEM} are read out onto the RBL from the source follower M1, by turning on the RWL while WWL is kept off.
(2) WRITE: Since the readout voltage has the source follower M1's Vth drop, we use a PVT-tolerant write-with-feedback (WFB) scheme. The WFB scheme applies the write voltage to the positive terminal of the operational trans amplifier (OTA). The OTA's negative terminal is connected to the negative feedback loop by turning on RWL and WWL, while its output is connected to WBL. This ensures the readout voltage is the same as the write voltage.
(3) REFRESH: The refresh is performed row-by-row across both 64-by-32 banks in parallel. For a given row, the DSAR modules' successive approximation register-based (SAR) binary search logic approximates the readout voltages (using (1)) from every column to uniformly quantized reference voltages and writes (using (2)) the approximated voltage back to the memory cell. The DSAR employs a shared global SAR controller, Vcm-based switching schemes, and a capacitive digital-to-analog converter (DAC) overlaid on top of the transistors to reduce the power and area overhead by about 68% and 26% respectively, compared to the conventional SAR ADC.

References

[1] Rentao Wan, Yannis Tsividis, Mingoo Seok, "AACIM: a 2785-TOPS/W, 161-TOP/mm2, <1.17%-RMSE, Analog-in Analog-Out Computing-In-Memory Macro in 28nm," IEEE European Solid-State Circuits Conference (ESSCIRC), 2024

[2] R. Wan, Y. Xu, D. -W. Jee and M. Seok, "AJPEG: A 26.4-pJ/pixel, 252-fps, 128×128 Image Sensor with an In-Sensor Analog DCT Processor for Data Compression," *2025 IEEE Custom Integrated Circuits Conference (CICC)*, Boston, MA, USA, 2025, pp. 1-3

[3] Rentao Wan, Mingoo Seok, “OASIS: Online Adaptive Video Compression via Closed-loop Feedback Control” NeurIPS, 2025, under review.
[4] L. Wang, M* Jana*, R. Wan, CT. Lin, M. Seok, “AnaRefresh: A 6-bit 64-by-64 Capacitor-based Analog Memory Macro with a Digital-Sample-Analog-Refresh Technique Achieving 2.82% Refresh Time Overhead and 24.6-pJ Refresh Energy Overhead”, ISSCC, 2025, *: Equal contributing authors, under review.