4.7 POWER-EFFICIENT ANALOG-TO-INFORMATION IMAGE SENSORS: THE FUEL OF AI MICROSYSTEMS

Ángel Rodríguez-Vázquez and Juan A. Leñero-Bardallo

ABSTRACT

This project explores innovative architectural concepts for image sensors with embedded intelligence. The main attributes of these architectures are the use of non-Von Neumann paradigms and the extensive usage of event-driven concepts. The goal is to achieve vision with unprecedented throughput and energy efficiency.

INTRODUCTION

Miniaturized, *intelligent* microsystems are at the central focus of semiconductor roadmaps and strategic nanoelectronic agendas (International Roadmap for Devices and Systems, 2020). They are enabled by the convergence of *heterogeneous* technologies and targeted to give electronic systems the ability to interact with the cyber-universe of analog information carrier signals. Besides being intelligent, in the sense of capable of *adapting* their response to arbitrary stimuli and unstructured environments, these systems must *communicate*, have a minimal *size* and *weight*, minimum *energy* consumption, and, if possible, energy *autonomy*. The target is to support human–machine interaction with electronic systems and appliances, have these systems connected, construct robots capable of detecting subtle surrounding changes and reacting rapidly, allow vehicles to identify risky situations, escape from them, etc. We also want electronic systems to resemble natural beings' outstanding capabilities regarding sensory data acquisition, analysis, interpretation, and reaction thereof. Challenges like these shape the domain of intelligent microsystems, a domain where the handling of information conveyed by light energy and the sense of *vision* is expected to play a significant role. Lately, the vision sense handles around 50% of the information that humans interchange with the environment.

STATE OF THE ART

Computer vision is one of the main areas within the ecosystem of intelligent sensory systems. Strategic agendas coincide when pointing out the widespread deployment of miniaturized computer vision systems as one of the main engines of R&D-and-Innovation in IT and nanoelectronics. The challenge is to conceive SWaP-optimized *artificial vision architectures* with embedded analysis capabilities and large efficiency in the number of operations completed per unit of time and energy. However, conventional processing architectures cannot meet these challenges. These traditional architectures adhere to the paradigm of encoding the analog information in the digital domain (using *digital imagers*) and relying on Von Neumann's digital processing architectures for analysis (A. Rodríguez-Vázquez et al., 2018). It implies encoding, storing, and transmitting the whole set of *raw* image data, following a *frame-by-frame* operation basis that uses enormous processing and energy resources while handling meaningless data. This approach significantly differs from that of *natural vision systems* as they operate by extracting *information*, instead of just raw data, right from the very front ends, at the edge of the information processing chain. Retinas, the front ends of natural vision systems, are *information-centric* front-end devices, while imagers used in conventional computer vision systems are *data-centric* devices.

The TIC179 research group and the company Teledyne-Anafocus (which was launched in 2001 by members of the TIC179 with the name Innovaciones Microelectrónicas S.L.) have conducted R&D on bio-inspired sensor front ends and smart imagers for decades. Contributions follow tracks that are concurrently explored by many other companies and academic institutions worldwide, namely:

- By incorporating *early-vision* processing tasks right in the focal plane, they are completed concurrently to light capture. The aim is to encode and communicate *features*, instead of frames. The approach uses a parallel processing paradigm alternative to Von Neumann's architectures but conferred with the *software-programming* capabilities of such a paradigm (Rodríguez-Vázquez et al., 2018).

- By using information encoding techniques similar to those employed by natural vision systems, including (i) the detection of spatial-temporal contrasts within visual scenes, (ii) the usage of pixels that trigger when they detect transient lighting changes, and (iii) use of peak luminance sensors that perform a light-to-frequency conversion (Leñero-Bardallo et al., 2018). Similar to feature-extraction front ends, all these techniques achieve large power efficiency and bandwidth.
- By exploring the ultimate physical limits of light carriers through detectors able to count individual photons and signaling the photon arrival events in time.

CONTRIBUTION

They are aligned to the challenge of improving the energy efficiency and performance of smart-imaging and vision sensor chips. These chips are lately the fuel of efficient sensor interfaces and hence crucial to incorporate Artificial Intelligence at systems targeted for the following AI technologies: (T1) IoT, (T2) robotics, and (T3) augmented/virtual reality. Targeted applications are, among others, in the following sectors: (S1) advanced industry linked to transportation, (S2) health and social welfare, and (S3) mobility and logistics.

Specific contributions of this project are summarized in the points below.

- The conception of the pixel and the chip architecture of a high dynamic range asynchronous image sensor with embedded solar energy harvesting capability. Its photodiodes operate in the reverse region of operation when measuring the lighting levels to which they are exposed. Also, they automatically switch to the photovoltaic regime to collect energy that feeds the sensor itself. Innovative circuit techniques allow these pixels to operate with down to 0.35 V voltage levels, compatible with the operating voltages of photodiodes in the photovoltaic region (Gómez-Merchán et al., 2020).
- The conception of two asynchronous solar sensors for vehicle navigation based on key-points tracking. They operate at high speed and employ photodiodes in the photovoltaic region of operation to reduce energy consumption. Only pixels that receive solar radiation generate output data. The first one measures the illumination

levels of the illuminated pixels to determine the sun's position by processing its outputs. The second provides directly as an output the coordinates of the centroid of the illuminated region, representing a great advance compared to the state of the art.
- The conception of new families of pixels and sensors capable of single-photon detection (through Singe Photon Avalanche Diodes) and embedded intelligence at pixel and sensor levels. These include distributed photon counters and distributed Time-to-Digital converters to capture 2D/3D scenes. They are targeted to support a new generation of solid-state LiDAR systems, on the one hand, and for the implementation of micro Positron Emission Tomography systems, on the other hand (Vornicu et al., 2020).

All architectural ideas explored in the project have been demonstrated through dedicated CMOS chips and camera modules built with them.

ACKNOWLEDGMENTS

This project has been carried out in collaboration with TELEDYNE-ANAFOCUS and with the support of the following researchers: Ana Gónzalez and José Ángel Segovia de la Torre.

WORKS CITED

Gómez-Merchán, R., et al. (2020). A comparative study of stacked-diode configurations operating in the photovoltaic region. *IEEE Sensors Journal*, *20*(16), 9105–9113. https://doi.org/10.1109/JSEN.2020.2987393

International Roadmap for Devices and Systems, IEEE, Edition 2020. https://irds.ieee.org/

Rodríguez-Vázquez, A., et al. (2018). CMOS vision sensors: Embedding computer vision at imaging front-ends. *IEEE Circuits and Systems Magazine*, *18*(2), 90–107. https://doi.org/10.1109/MCAS.2018.2821772

Vornicu, I., Carmona-Galán, R., & Rodríguez-Vázquez, A. (2020). CMOS SPAD sensors with embedded smartness. *International SPAD-Sensor Workshop.* Edinburgh, Scotland (UK), June 2020.