3-D Silicon Retina Modeling Spatiotemporal Dynamics of Ganglion Cells

Jongkil Park*, Sohmyung Ha†, and Gert Cauwenberghs†

*Electrical and Computer Engineering, †Bioengineering, University of California, San Diego, La Jolla, CA 92093
E-mail: jop021@ucsd.edu

Summary: This project combines advances in CMOS imaging and neuromorphic engineering towards the realization of a 3-D integrated event-based (asynchronous) silicon retina with real-time spatiotemporal processing and low power consumption, approaching the coding and energy efficiency of the mammalian retina for applications in retinal prostheses and neuromorphic visual information processing.

Despite tremendous advances in semiconductor technology and in our understanding of the mammalian retina, today’s imaging technology for artificial vision is far inferior to its biological counterpart. The mammalian retina is the gold standard in engineering design as the most efficient image processor with superior coding and energy efficiency. A low power silicon retina approaching some of the metrics of efficacy and efficiency of the mammalian retina is of critical importance for retinal prostheses and neuromorphic object recognition.

Since the pioneering silicon retina of Mahowald and Mead[1], large numbers of silicon retina chips have been developed each targeting various aspects of retinal modeling. To date, there have been only a few attempts at realizing more complete functionality of the retina in real-time on a single chip. Zaghloul and Boahen[2] developed a silicon retina that accounted for several of the spatiotemporal attributes of a range of ganglion cells, but at the expense of relatively large power consumption, almost a factor thousand times larger than that of the mammalian retina.

The 3D neuromorphic silicon retina emulates the detailed spatiotemporal dynamics of ganglion cell visual coding in the mammalian retina, while offering ultra-low power operation and high integration density as well as high fill factor. We combine state-of-the art approaches in CMOS imaging and neuromorphic computing on a single 3-D integrated substrate through wafer stacking of three traditionally disparate CMOS technologies that are each tailored for optimal performance: high light-sensitivity in the photoreceptor array on the top layer in an optoelectronic CMOS process, low-power and high-density asynchronous neural event coding and communication in the middle layer in a deep-submicron CMOS process, and dedicated memory for reconfigurable spatiotemporal dynamics in the base layer in DRAM technology.