Artificial retinas — fast, versatile image processors

Kazuo Kyuma, Eberhard Lange, Jun Ohta, Anno Hermanns, Bryan Banish and Masaya Oita

Artificial retinas combine video camera and image processing functions, allowing machines to function in their environment with unprecedented autonomy, or to augment quality control, surveillance and hazard monitoring. We review several retina devices and reveal how they execute basic manipulations of the image at processing speeds well beyond the capabilities of the human eye.

RECOGNITION and understanding of real-world images is one of the most important information processing technologies for the 21st century. Such images contain a very large amount of information which present image processing systems cannot analyse in real time. This is because present systems separate image sensing and image processing: images are sensed by a camera and processed by a computer. Consequently, performance is limited by slow camera frame rates and low transmission rates between camera and computer. By contrast, human beings can execute very difficult tasks in real time. For example, we can instantly recall the names of people by glancing at their faces, and can analyse motion information by selectively detecting a moving object. This is due to the inherent parallelism of our visual system, particularly of our retina, which combines image sensing and processing functions.

The artificial retinas described in this article, having evolved out of a continued research effort within our laboratory to explore and develop optical neural devices[1-3] capable of on-chip image processing, are devices that can simultaneously sense and process images. To date, several kinds of related devices have been developed, such as vision chips, focal plane processors, and silicon retinas[4-6]. However, the specialized functions of the devices reported so far, such as spatial and temporal filtering, have precluded them from use in large-scale industrial or commercial applications. Therefore, devices with more 'on-chip' functionality are required. The most notable distinction between our artificial retinas and other devices is their on-chip functional flexibility.

Device, function, applications

Figure 1 shows our artificial retina (RETINA2) which consists of a two-dimensional (2-D) array of variable sensitivity photodetectors (VSPDs)[7,8]. The VSPDs were formed by a side-by-side pair of diodes integrated onto and separated by a semi-insulating GaAs layer (pn-np structure). The size of each VSPD is 80 × 80 mm^2, and the total size of a 128 × 128 array is 10.24 × 10.24 mm^2.

The VSPD combines three devices: photodetector, spatial light modulator (SLM), and analog memory[3,9]. The photodetector current depends, both in sign and magnitude, on applied voltage, thus integrating detector and SLM into one variable sensitivity device. The analog memory effect encountered in these VSPDs stores conductivity information when a voltage is applied in the presence of an optical write pulse. This information can then be retrieved by injecting an optical readout pulse.

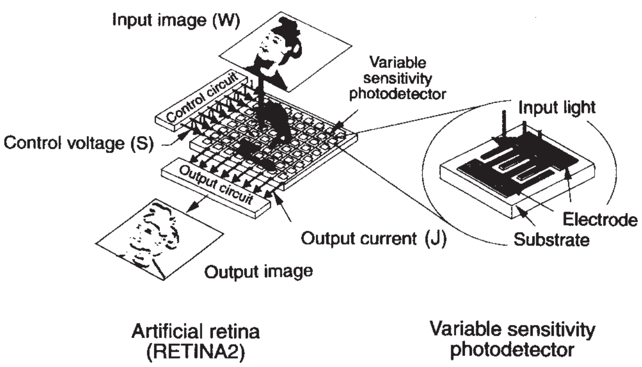

FIG. 1 Image processing in RETINA2 as a matrix-vector multiplication. The input image modulates the conductivities at the VSPD sites, thus generating a weight matrix W. This matrix is addressed with a control signal vector S, which also determines the operating mode of the retina. The resulting photocurrents are read out by the output circuit.

Image processing in RETINA2 is based on optical matrix-vector multiplication[8]. In Fig. 1, the input image is directly projected onto the chip as the weight matrix W. All VSPDs have one electrode connected along rows, yielding a sensitivity control vector S. Thus, the VSPD sensitivities can be set to arbitrary values at each row within a certain range. In addition, the remaining VSPD electrode is connected along columns, yielding an output current vector J defined by the matrix vector product J = WS.

The distinguishing feature of RETINA2 over other devices is the variety of processing that can be achieved through simply changing the control voltage pattern, S. Several examples of processing are summarized in Fig. 2 including TV camera-like image sensing, edge extraction, noise elimination (image smoothing), Fourier transform, pattern matching, and image compression/recognition. Two examples follow.

The first case is the edge-extraction operation shown in Fig. 1. The sensitivities of two adjacent detector rows are set to +1 and −1, respectively, whereas all other sensitivities are set to 0. In this case, the output current is proportional to the difference in light intensities of the two active rows. By shifting the control voltage pattern cyclically (0, +1, −1, 0, 0,), the horizontal edges of the input image are sensed. Thus, the system operates in a time-sequential and semi-parallel mode. The processing time (that is, frame rate) is 400 μs. The edge-extraction performance is depicted in Fig. 1.

The second example is an image recognition system using a neural network as a post-processor of the artificial retina output. As shown in Fig. 2*f*, all detectors are set to a uniform sensitivity of one. The photocurrents (proportional to light intensity) in each VSPD column are summed such that the output current vector, J, constitutes a vertical projection of the input image. Output signal J is a compressed representation of the input image W and is fed into the neural network classifier for recognition. This mode of operation has the advantage that the projection can be obtained in one cycle equal to the VSPD response time of 3 μs.

We have applied this system to optical character recognition, and have succeeded in recognizing 46 (Japanese) Hiragana characters with almost 100 per cent recognition rate. Furthermore, we have demonstrated a system that can recognize the full set of 1,945 Joyo Kanji characters by feeding the horizontal and vertical projections of the characters into a shift-tolerant neural network classifier.

The application fields of RETINA2 include automotive and avionics sensors, fac-

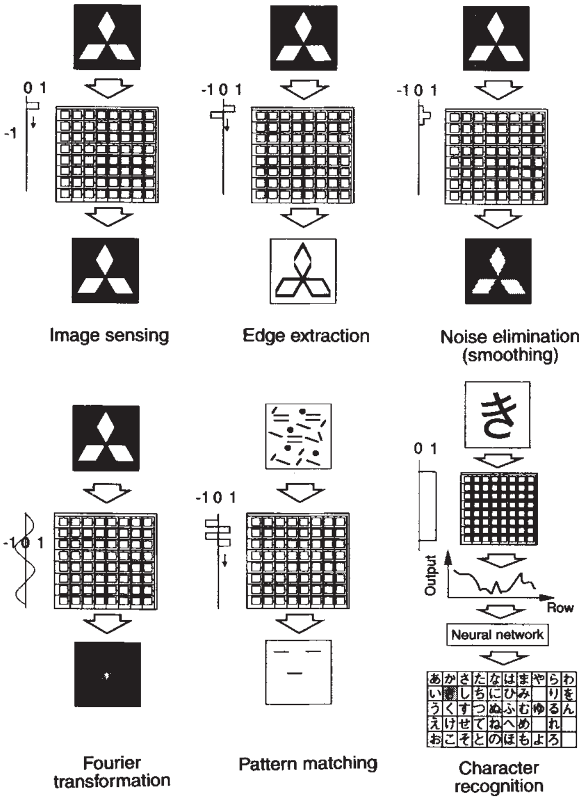

FIG. 2 Examples of image processing options in RETINA2. In the image sensing mode, *a*, all VSPD rows are successively scanned with a single polarity voltage pulse, one at a time. The photocurrents are read out on all column lines simultaneously. Reading out N^2 pixels thus takes merely N clock cycles. To switch to other operating modes one needs only to alter the scanned voltage pattern. For edge extraction, *b*, two adjacent rows are addressed with opposite polarity. The smoothing operation *c* uses a voltage sequence (0.5, 1, 0.5) to suppress noise and smooth out edges. Addressing the VSPD rows with a sinusoidal control signal of wave vector k computes the corresponding k-vector component of the spatial Fourier transform, with one component calculated in one clock cycle. In pattern matching, *e*, the target item is detected by using it as the scanned control signal. Finally, character recognition is facilitated by performing a projection, *f*. Here, all rows are simultaneously addressed with the same voltage, and the entire operation takes only one clock cycle. The retina output (projection) is then fed into a simple neural network for classification.

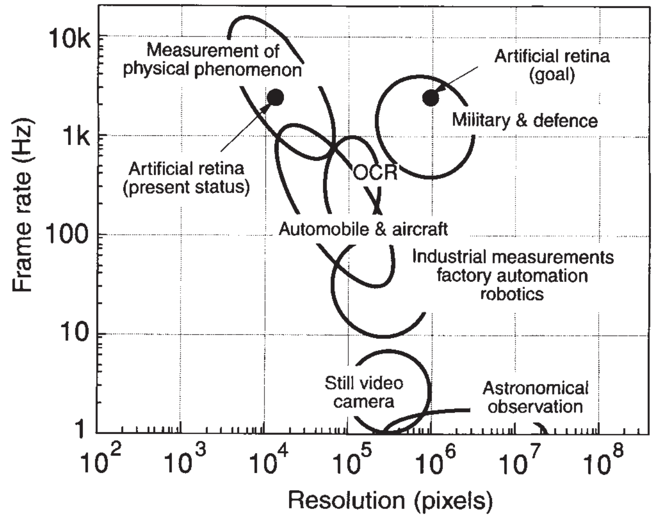

FIG. 3 Performance of the artificial retina and typical requirements in potential areas of application. It is hoped that improved resolution, to be achieved using silicon technology currently under development, will greatly enhance the scope of problems that can be addressed with the artificial retina.

tory automation and robotics, video still camera, and defence applications (see Fig. 3). One particularly interesting use is as an artificial eye in moving vehicles where it can assist the driver by recognizing road markers. With its particularly fast response, RETINA2 could prevent collisions by sensing obstacles and approaching vehicles.

Retinas in silicon technology

If artificial retinas are to find acceptance outside the laboratory, it is imperative to make use of the low-cost, high-resolution and yield of silicon technology, which furthermore provides ease of integration with a wide selection of processing electronics on the same chip. We are currently building two such retinas. The first is based on MOS (metal-oxide-semiconductor) imager cells augmented by on-chip electronics. The second employs high-gain, bipolar npn VSPDs (lateral three-layer structures with donor/acceptor/donor doping, respectively) along with compact CMOS (complementary metal-oxide-semiconductor) addressing electronics. Both display high photosensitivity, dense integration and frame rates which leave plenty of room for post-processing, should it be desired, while still completing the task in real time.

Next generation

More applications will open up as next generation chips combine not only image sensing and processing but also learning capabilities to extract features from the images in the real world. In order to realize such 'learning vision chips', we have recently fabricated RETINA3, consisting of a 2-D VSPD array vertically integrated on a one-dimensional LED array[3,10]. In preliminary experiments, image addition, subtraction and tracing of moving objects have already been achieved using the analog storage function of VSPDs. Furthermore, we have proposed an optical associative memory with learning capability consisting of SLMs integrated on the 2-D VSPD array[11,12]. The light transmittance of the SLM is adaptively varied in the learning process. Computer simulations indicate that the incomplete images W are successfully transformed into the corresponding point attractors by the nonlinear thresholding of the output J of the VSPD array, and by feedback to the input S.

In summary

Artificial retinas execute multiple, easily programmable forms of image processing with inherently high speed and parallelism. A new generation of high-resolution retinas that incorporate light-emitting devices will further boost processing capabilities. □

Kazuo Kyuma, Eberhard Lange, Jun Ohta, Anno Hermanns, Bryan Banish and Masaya Oita are in the Semiconductor Research Laboratory, Mitsubishi Electric Corporation, 8-1-1 Tsukaguchi Honmachi, Amagasaki, Hyogo 661, Japan. For more information on the artificial retina systems featured in the article, fill in reader service number 100.

1. Ohta J., Takahashi, M., Nitta, Y., Tai, S., Mitsunaga, K. & Kyuma, K. *Opt. Lett.* **14**, 844–846 (1989).
2. Nitta, Y., Ohta, J., Tai, S. & Kyuma, K. *Appl. Opt.* **32**, 1264–1274 (1993).
3. Kyuma, K., Ohta, J. & Nitta, Y. in *Associative Neural Memories* (ed. Hassoun, M. H.) (Oxford University Press, New York, 1993).
4. Wyatt Jr, J. L., Standley, D. L., & Yang, W. *Proc. IEEE int. Conf. Robotics and Automation*, Sacramento California 1330–1335 (1991).
5. Fossum, E. R. *Opt. Engng.* **28**, 865–871 (1989).
6. Mead, C. in *Analog VLSI and Neural Systems* (Addison-Wesley, Reading, Massachusetts, 1989).
7. Lange, E., Funatsu, E., Hara, K. & Kyuma, K. *Proc. IJCNN* **1**, 801–804 (1993).
8. Lange, E., Nitta, Y. & Kyuma, K. *IEEE Micro* (in the press).
9. Ohta, J., Nitta, Y., Tai, S., Takahashi, M. & Kyuma, K. *IEEE J. Lightwave Technol.* **9**, 1747–1754 (1991).
10. Nitta, Y., Ohta, J., Tai, S. & Kyuma, K. *Proc. IJCNN* **1**, 805–808 (1993).
11. Zhang, W., Ishii, T., Takahashi, M. & Kyuma, K. *Opt. Lett.* **17**, 673–675 (1992).
12. Oita, M., Nitta, Y., Tai S. & Kyuma, K. *IEICE Trans. Electronics* **E77-C**, 56–62 (1994).