

M I C R O P R O C E S S O R

www.MPRonline.com

THE INSIDER'S GUIDE TO MICROPROCESSOR HARDWARE

FASTER THAN A BLINK

Parallel Processors and Bonded Sensors Enable Ultrafast Imaging

By Tom R. Halfhill {2/12/07-01}

If a picture is worth a thousand words, what are 100,000 pictures per second worth? Plenty, to anyone who can design a digital-imaging system capable of achieving such spectacular frame rates. Applications include robotic vision, intelligent video surveillance, scientific

analysis of momentary events, monitoring industrial processes, interactive games, and guidance systems for unmanned vehicles and missiles.

To meet the challenge of recording, analyzing, and enhancing thousands of frames per second, scientists from Hungary, Spain, and the U.S. founded AnaLogic Computers Ltd. in Budapest in 2000. Three years later, they founded AnaLogic's successor company, Euteucus Inc., in Berkeley, California. With grants from the U.S. Missile Defense Agency and the Office of Naval Research, they developed Cellular Visual Technology (CVT).

CVT combines a massively parallel processor architecture with optimized image-processing software. Some implementations use an innovative semiconductor fabrication process to bond the image sensor directly onto the parallel-processor array, creating a stacked multilayer chip. Figure 1 is a photo of C-TON, the first such chip designed for the commercial market. The main goal of Euteucus, however, is to license the technology as intellectual property (IP) for integration into custom ASICs and FPGAs.

The company's cofounders and chief technology officers, Dr. Akos Zarandy and Dr. Csaba Rekeczky, divide their time between Budapest and Berkeley. Zarandy, a coinventor of the technology and vice president of technology for Euteucus, delivered a technical presentation on CVT at Fall Microprocessor Forum last October. This article also draws on an interview with Rekeczky, another coinventor and Euteucus vice president of research.

Not Your Typical Digital Camera

Euteucus is taking a radically different approach from conventional digital cameras and camcorders. Typically, a digital camera has a CCD or CMOS image sensor covered with a grid of

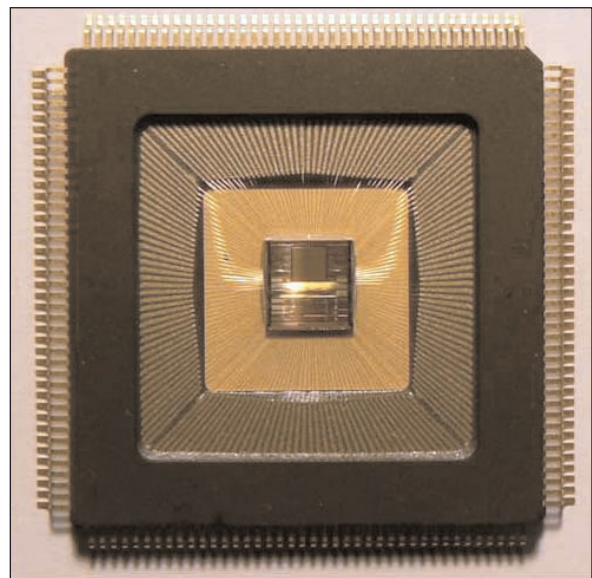


Figure 1. This photo shows the Euteucus C-TON chip before the image sensor is bump-bonded onto the parallel-processor die. This 0.18-micron implementation has 64 image processors. Euteucus has applied for an international patent on this technology. (Source: Euteucus)

tiny sensor sites. Each sensor site corresponds to roughly one pixel. During an exposure, light falling on the sensor sites is converted to electricity—in other words, the sensor sites convert photons into electrons. The more photons a sensor site captures in its light well, the more electrons it generates, so the brighter the pixel. After an analog-to-digital conversion, a conventional microprocessor—typically an embedded RISC core—reads the values from each sensor site and assembles the image, pixel by pixel. The processor usually performs additional operations on the data to enhance the image before saving the file in the camera's flash memory or other storage medium.

Conventional digital cameras have low frame rates, often measured in seconds per frame instead of frames per second. It takes time to sample the sensor sites and assemble the image. To get around this problem, almost all digital cameras have buffers that temporarily store one or more images for post-processing while the camera readies itself for the next exposure.

Digital camcorders have deeper buffers and, by sacrificing pixel resolution, they can manage the 25 to 30 frames per second required for PAL, SECAM, or NTSC video. High-definition (HD) digital camcorders can record somewhat higher frame rates and somewhat higher resolutions by applying more processing power to larger image sensors, which accounts for the higher prices of HD camcorders. But capturing and processing images at superfast frame rates are beyond the capabilities of conventional equipment.

Image processing is the main obstacle. All the photons arrive at the sensor sites at nearly the same instant—photons at different sensor sites are counted in the same “time window”—but the microprocessor responsible for assembling the image operates on the data serially. Image sensing is massively parallel, but image processing is not. Therefore, the logical solution is to apply massively parallel processing, so that each sensor site or cluster of sites has its own processor. In addition to taking that approach, Eutecus employs techniques that mimic some characteristics of human vision.

Massively Parallel Image Processing

First, we should note that Eutecus achieves its fantastic frame rates partly by massively parallel processing but also

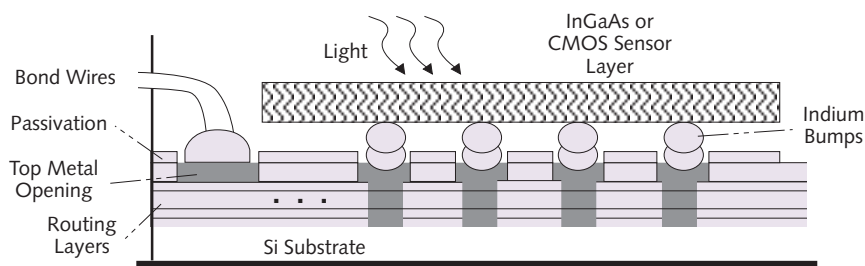


Figure 2. Using an innovative manufacturing technique called 3D bump-bonding, Eutecus grafts an image-sensor die onto another die containing the massively parallel processor array. The bumps, made of indium, tightly couple the two dies together, creating a multilayer single-chip solution. This implementation supports the highest frame rates, especially if the sensor is fabricated in an exotic InGaAs process.

by sacrificing pixel resolution and color. These compromises are inevitable with the current state of semiconductor technology. The company's latest CVT sensor-processor can resolve an image of only 64×64 pixels, in either 8- or 16-bit grayscale or stark 1-bit black and white. Eutecus believes it's possible to manufacture a chip in the same fabrication process (0.18 micron) that could capture as many as 256×256 pixels—about 0.06 megapixel.

That resolution may seem absurdly low compared with today's popular 10.2-megapixel digital SLRs, but it's sufficient for a surprising number of applications. In robotic vision, for example, even a low-resolution image is good enough to keep a robot from colliding with a wall or running over small objects. In other applications, detecting changes in a fast-moving process during an extremely brief time period is more important than recording fine detail. One customer is using a CVT imaging system to study the way cracks develop when a pane of glass shatters. In addition, CVT can make the most of its limited resolution by rapidly evaluating small points of interest within a larger area, much as the human eye perceives details within its wide field of vision.

Despite severe compromises in resolution and color, CVT is an impressive achievement that relies heavily on massively parallel image processing. Chip designers can implement the technology in two basic ways. The first method, which delivers the highest frame rates, consists of a single chip that stacks the image-sensor die on top of the processor die using a fabrication technique called 3D bump bonding.

As Figure 2 shows, the sensor die is bonded to the processor die on thousands of tiny indium bumps, creating a tightly coupled multilayer chip. The sensor can be a conventional CMOS or CCD device, but, for maximum performance, Eutecus uses a more exotic indium-gallium-arsenide (InGaAs) material. InGaAs sensors are about ten times more sensitive to light than silicon sensors are, so they can record usable images at the much shorter exposures that high frame rates require. The processor die to which the sensor is mated can be fabricated in a conventional CMOS process.

The second basic method of implementing CVT uses separate chips for the image sensor and processor array. The logical architecture is the same as the first method, but it's a different physical architecture. Although a fast parallel interface connects the chips together, it's not as fast as the bump-bonded interfaces, so the maximum frame rate is one or two orders of magnitude slower (about 1,000 to 10,000 frames per second), depending on the specific implementation. However, the two-chip configuration has a few advantages: it can use higher-resolution image sensors, it's less costly to manufacture, and chip designers can implement the parallel processor array

as an ASIC or synthesize it in an FPGA. The FPGA option allows fewer array processors but greatly reduces time to market.

A third possible implementation of CVT would integrate the sensor-site array and parallel processor array on the same die. This arrangement would couple the two arrays together even more tightly than the multilayer chip does, but it's not practical with existing semiconductor technology. Image sensors and ASICs require different design and fabrication parameters, even when both are manufactured in CMOS. In addition, integrating the sensor sites and processors on the same die would limit how closely the chip designers could pack the components together, reducing the chip's pixel resolution. And the sites would have smaller light wells for capturing photons, which would reduce their sensitivity. Nevertheless, single-die integration is an avenue for future research and development.

Overcoming the Drawbacks of Fabrics

The sensor array is fairly conventional, so we'll focus on the processor array. Euteucus weaves dozens of "processor cells" together in a tight fabric of on-chip interconnects. Figure 3 is a block diagram of the Euteucus C-TON chip. This particular design has 64 processor cells in an 8×8 array, but other designs can have smaller or larger processor arrays. (An early prototype has a 4×4 array.)

Each cell in the C-TON chip receives image data from 64 sensor sites, so the image resolution is 64×64 pixels. An integrated 8-bit microcontroller serves as the global control processor, handling off-chip I/O and broadcasting program instructions to the array. The C-TON design has 2KB of global program memory. C-TON is fabricated in conventional 0.18-micron CMOS and runs at 100MHz. Its 32-bit I/O bus also runs at 100MHz, providing 400MB/s of bandwidth.

Inside each cell is a 24-bit processor based on a proprietary single-instruction, multiple-data (SIMD) architecture optimized for image processing. A multiplexer collects the output from all the sensor sites connected to the cell and funnels the signals through a generic 8-bit analog-to-digital converter (ADC), which in turn connects to the processor. In addition, each cell shares a 1KB block of local memory with another cell and has a communications interface to all neighboring cells. Cells can communicate with each other through these shared memories, transferring 8 bits of data per clock cycle. There are no direct connections between distant cells.

It's important to grasp that when the image sensor is bonded to the processor die, as in the C-TON chip, raw image data from the sensor pours into the processor array through

the thousands of indium bump connections, not through the chip's conventional I/O bus. Each cell has a private I/O interface to the sensor sites directly above it. This clever arrangement solves a common problem with massively parallel processor arrays—propagating data through a large fabric.

Other massively parallel designs receive data at the edges of their fabrics and pass the data through the network. But in a bump-bonded multilayer chip, the image data enters the fabric everywhere at once, arriving at the destination (a processor cell) without unnecessary detours or delays. True, the image data must still traverse the fabric to leave the chip through its conventional I/O interface. However, by pre-processing the image data locally, the processor cells can reduce the amount of raw image data that must move off chip.

Layers of Arrays Distribute Resources

At the abstract design level, a sensor-processor chip like C-TON has multiple layers of hardware resources: image sensors, multiplexers, ADCs, processors, and memories. The image sensor is a physically separate bump-bonded die. The multiplexers and ADCs are mixed-signal components, so they require a distinct layer in the other die. The processors and SRAMs are part of that same die, built conventionally in digital CMOS. In the future, it might be possible to stack all these layers in one die. Figure 4 is an abstract view of the layers, with cross-sectional views of today's bump-bonded C-TON chip and a speculative view of a future 3D-layered chip.

The processor in each cell has a 24-bit ALU for performing 8- and 16-bit signed arithmetic operations (addition,

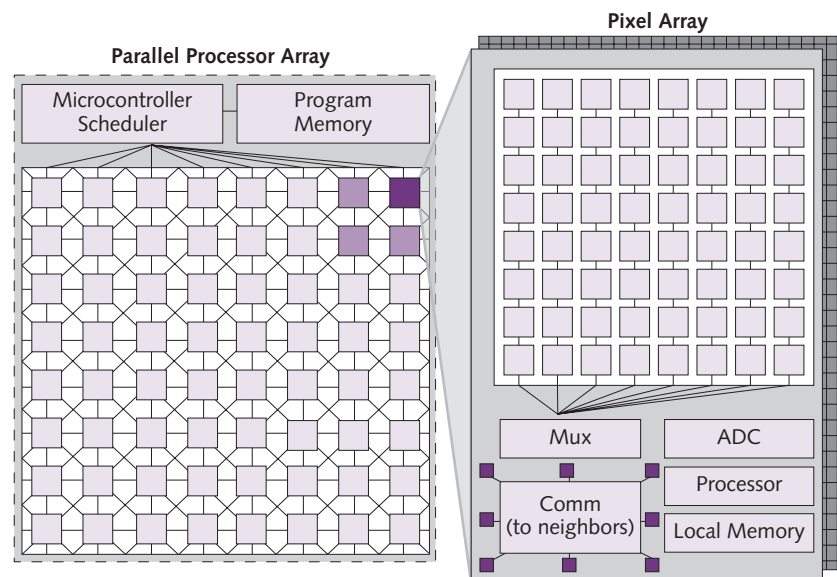


Figure 3. C-TON block diagram. For image processing, Euteucus employs a massively parallel processor array consisting of dozens of "processor cells." The array is scalable; this particular design weaves 64 cells together in an 8×8 array. As the inset shows, each cell has a multiplexer connected to the sensor sites, an analog-to-digital converter (ADC), an image processor, local memory, and an I/O interface to neighboring cells. C-TON has an image sensor bonded to its die, connecting 64 sensor sites to each processor cell.

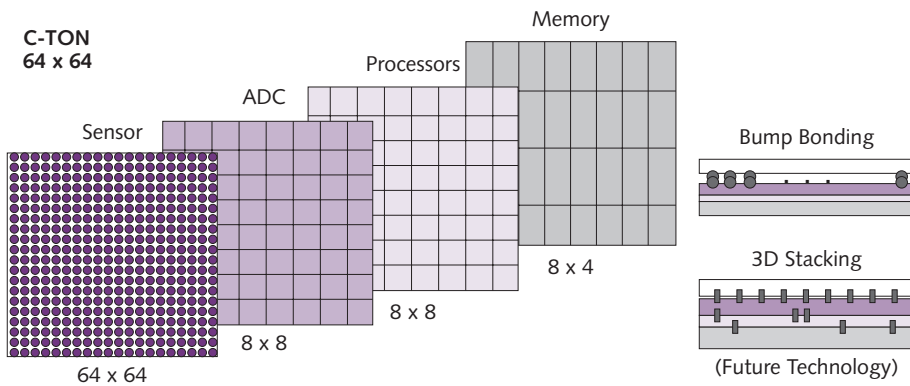


Figure 4. At the abstract level, a Eutecus sensor-processor chip resembles a multilayer cake, with arrays of different components in each layer. In the C-TON chip, a 64×64 array of sensor sites is the top layer and bonds to the lower layers on 4,096 indium bumps. Immediately below the sensor sites is an 8×8 array of multiplexers and analog-to-digital converters (ADC). The chip dedicates one ADC for each block of 64 sensor sites. Below that is an 8×8 array of 24-bit image processors—one for each ADC. Below that is an 8×4 array of 1KB SRAMs, so two processors share each 1KB block of local memory. Future designs might stack all these layers in one die.

subtraction, and multiplication). Another function unit, called the morphologic unit, performs simple 1-bit operations on black-and-white images and assists the ALU when processing grayscale images. A third function unit performs comparisons. Each processor can execute one instruction per clock cycle and read data from three sources: the cell's local memory, the memories shared with neighboring cells, and the image sensor (via the cell's ADC). Table 1 shows some performance metrics for C-TON when executing low-level image-processing tasks.

Although all the cells in the fabric must run the same global program, they can manipulate their pixels in slightly different ways by using data-driven operators and bit-masking techniques. This capability is another important advantage of connecting sensor sites directly to processors in the array. Figure 5 shows one possible benefit of providing individual pixels with special treatment. By individually adjusting each

Operation	Eutecus C-TON	Notes
Sobel Edge Detection	6.2 μ s	8-bit grayscale
Convolution (3x3 Pixels)	8.3 μ s	8-bit grayscale
Convolution (9x9 Pixels)	32 μ s	8-bit grayscale
Find Minimum (3x3 Pixels)	4.2 μ s	8-bit grayscale
Binary Dilation (3x3 Pixels)	3.1 μ s	1-bit b&w
Skeletonization	24.8 μ s	1-bit b&w

Table 1. The Eutecus C-TON chip can perform basic image-processing tasks in a few microseconds when operating on grayscale or 1-bit black-and-white pixels. To obtain the numbers in this table, Eutecus measured performance at the chip's nominal clock frequency of 100MHz. Eutecus provides function libraries of such primitives to help programmers write higher-level image-processing software.

pixel, the processors can effectively expand the dynamic range of the image, recapturing lost tonality and detail.

To save power, Eutecus uses extensive clock gating throughout the processor array. Individual processors can shut down when not needed. Eutecus says the 100MHz C-TON chip consumes a maximum 300mW when all processors are busy. At times, power consumption drops below 50mW, making C-TON suitable for battery-powered systems. (Figure 6 is a layout photo of C-TON.) Nevertheless, heat dissipation is a potential problem with this unusual design. Image sensors produce noisier pictures when warmed, because the excited sensor sites generate “hot”

(falsely illuminated) pixels, especially in darker areas of the image. To reduce this effect, developers can use the image processors to increase the signal-to-noise ratio.

Many Implementations Are Possible

C-TON is an early example of Eutecus technology, only one of many possible implementations. The size of the fabric and the number of sensor sites associated with each cell can vary, depending on the design. A cell could process data for an 8×8 block of pixels, as C-TON does, or for a single pixel. An earlier prototype chip, also fabricated in 0.18-micron CMOS, resolves only 32×32 pixels.

A next-generation design could implement the processor array as an ASIC in 90nm CMOS without bonding the image sensor to the die. The fabric would communicate with the sensor through the chip's conventional 32-bit I/O bus. Eutecus says this two-chip solution could have a 180×180 array of cells (32,400 processors) and would run at 120MHz, supporting a maximum resolution of 720×720 pixels (16 pixels per processor). This design would sacrifice frame rate for higher pixel resolution and lower manufacturing costs. Even so, it still should be capable of capturing more than 1,000 frames per second and performing a few morphologic operations on the images.

However, Eutecus doesn't want to be a fabless semiconductor company. The primary business model is licensing CVT technology as soft IP. To that end, Eutecus offers the processor array and associated elements as synthesizable VHDL models, ready for customers to integrate into their own ASICs and FPGAs. Eutecus provides development tools that allow customers to scale the array to any size they need. To aid software development, the company provides “Instant Vision” function libraries optimized for image processing. These low-level functions, written in assembly language, allow customers to write their application software in high-level C or C++.



Figure 5. High frame rates aren't the only advantage of applying massively parallel processing to digital imaging. The array processors can adjust the intensity of individual pixels, improving the photographic quality of the image. Many other effects are possible by applying intelligent image-processing algorithms. (Source: Euteucus)

Designs that synthesize the processors in an FPGA instead of carving them into an ASIC will always require an off-chip image sensor, because it's not practical to bond the sensor to the programmable-logic chip. Also, the limited capacity of an FPGA (relative to an ASIC) reduces the maximum size of the processor array. Nevertheless, programming an FPGA is much faster and less expensive than spinning an ASIC.

One such design synthesizes a 20 × 20 array of processors in a Xilinx Virtex-4 LX200. The global control processor is a RISC or DSP core synthesized in the same FPGA. At 150MHz, this implementation could execute 60 billion operations per second. Maximum image resolution would be 160 × 160 pixels (64 pixels per processor). Euteucus says the maximum frame rate would be about 3,000 frames per second.

Imitating Human Vision

If a customer can tolerate lower frame rates, the two-chip implementation (an unbonded image sensor with either an ASIC or FPGA) has another advantage, besides lower cost and faster time to market. A separate sensor can have much higher pixel resolution than a bonded multilayer chip would support. CVT allows developers to exploit the higher resolution by imitating some vital characteristics of human vision.

Conventional digital-image processors scan a whole image, paying little or no attention to the contents of the image. But humans (and other animals with single-lens eyes) don't see that way. Instead, our eyes rapidly flick from one part of a scene to another, evaluating a few key points to quickly assemble a composite image. Scientists call these rapid eye movements "saccadic jumps." They are necessary because, although our eyes can perceive a wide visual field of about 210 degrees, we can focus on only one small part of the field at a time. The focus point is called the fovea, and that's where our brains concentrate most of our analytical

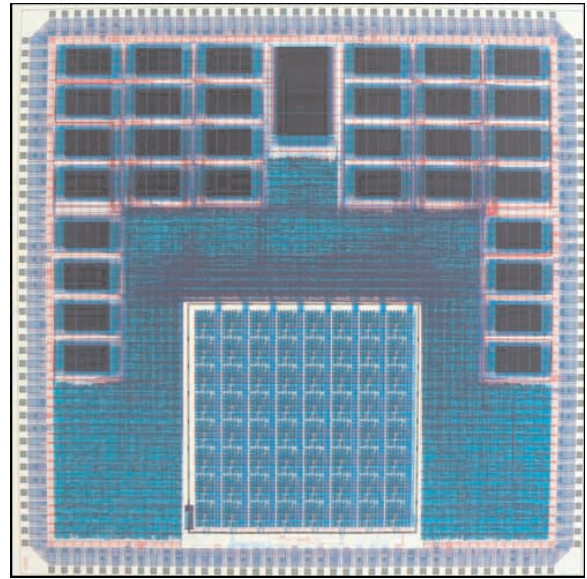


Figure 6. C-TON layout photo. At bottom is an 8 × 8 array of mixed-signal components, including multiplexers and ADCs. On the surface of the chip, the image sensor (not shown) is bonded to 4,096 indium bumps connected to this array. The large dark block at top center is 2KB of SRAM, the global program memory for all 64 image processors. The other 32 dark blocks are local memories for the processors. Each block is 1KB of SRAM shared by two processors. The processors aren't visible in this photo, because they are lost in the sea of wiring surrounding the other blocks. (Source: Euteucus)

power. Peripheral vision is mainly for detecting motion or sudden changes in illumination, which could signal a survival threat. Figure 7 illustrates the difference between saccadic vision and conventional digital vision.

When an image sensor has more pixels than a CVT array can process at once, developers can mimic the saccadic

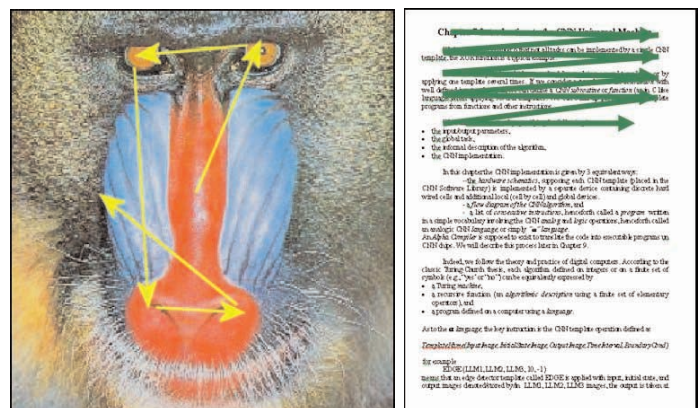


Figure 7. By evaluating a few key points within a larger visual field, organisms with single-lens eyes can rapidly assemble a composite image and respond appropriately. Conventional digital vision works differently—most cameras scan the whole image, with little regard to the contents. Insects with compound eyes have a visual field approaching 360 degrees, but they lack the ability to resolve fine detail. (Source: Euteucus)



Price & Availability

Eutecus is currently testing samples of C-TON, the first bonded sensor-processor chip intended for commercial applications. C-TON has an 8×8 -processor array and captures 64×64 -pixel monochrome images. Eutecus is licensing its IP cores for FPGAs now. IP cores for ASICs are also available, subject to customer specifications. Eutecus hasn't publicly disclosed licensing fees. Development tools will include four libraries of pixel-processing primitives and a development system that's also a reference design for an intelligent surveillance camera. For more information, contact CEO and President Stephen D. Hester (shester@eutecus.com), or visit:

- www.eutecus.com
- www.analogic-computers.com

jumps of human vision by processing groups of pixels in smaller areas. In effect, the processor array becomes a movable fovea point, jumping from one part of the sensor to another. Intelligent algorithms can evaluate the contents of each fovea point by processing the pixels in various ways. The first step might be to scan a lower-resolution overlay of the image to identify possible things of interest. Or the software could compare the frame to a previous frame, looking for evidence of movement, then zoom in on areas that have changed. Figure 8 is a diagram of this process.

Eutecus is pursuing a commercial application for this technique: intelligent video surveillance. A camera can be programmed to detect motion within its broad field of view,

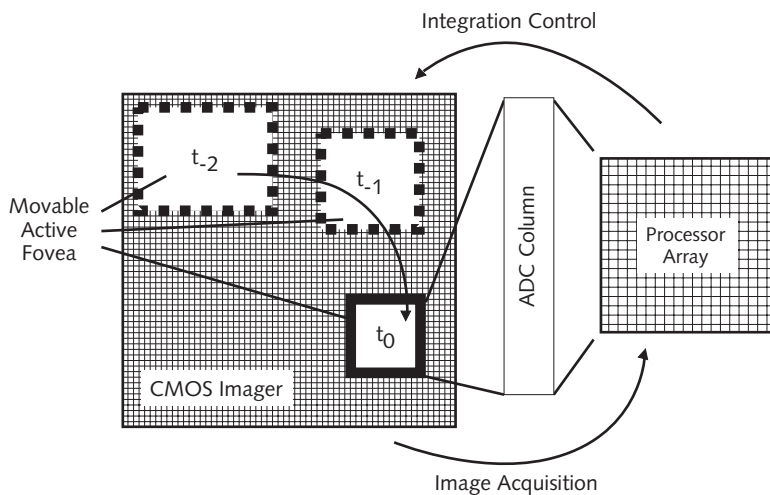


Figure 8. Eutecus CVT technology allows developers to mimic the saccadic jumps of human vision by applying the processor array to different parts of a high-resolution image. By concentrating on a few key points of interest, the software can quickly analyze the image. Applying the processor array to larger or smaller groups of pixels allows the software to zoom in or out on specific details. This technique exploits the ability to read data from sensor sites as if the sensor were a random-access memory.

then zoom in on anything that moves and analyze it in more detail. If the camera finds something of interest, it can alert security personnel. Note that with the appropriate software, the camera can make these decisions autonomously instead of relying on a centralized computer or merely transmitting the images to a wall of monitors watched by a bored security guard.

This isn't science fiction. Some conventional digital cameras already perform simpler forms of real-time image analysis. For instance, Canon and Fuji make consumer-level cameras that can automatically detect up to ten human faces in the viewfinder, lock focus on them, adjust exposure, and follow the faces as they move within the camera's field of view. Eutecus, with its superfast frame rates and individual pixel processing, vastly expands the possibilities of this technology.

However, programmers must make a trade-off: the more analysis their software performs, the more time it takes to process each image, so the lower the frame rate. If a program drives the 100MHz C-TON chip to its peak performance of about 100,000 frames per second, each processor in the array has only 10 microseconds to analyze and/or modify its pixels. As Table 1 shows, it's possible to perform some primitive operations in such a brief moment, but more-sophisticated tasks will require either lower frame rates or faster chips.

Programmers writing software for ultrafast frame rates will probably have to use assembly language instead of calling prewritten library functions from a C++ program. Currently, Eutecus is working closely with early customers to develop software. In the future, Eutecus acknowledges, lower-level development tools will likely be necessary.

Eutecus Is Entering the Market

Eutecus has completed an early prototype chip with a small 4×4 array of processors. This proof-of-concept chip has a bump-bonded image sensor capable of capturing images in 32×32 pixels. Engineering samples of the more advanced C-TON chip, with its 8×8 array and bonded 64×64 image sensor, are undergoing testing. These chips demonstrate that Eutecus technology works, and they will help customers design products and write software while spinning their own ASICs or programming FPGAs.


Licensing IP is the company's main goal. Eutecus is shipping the first versions of its IP blocks: one for ASIC implementations (with or without a bonded sensor) and one for FPGA integration. Improved versions of both IP libraries are due this quarter. Meanwhile, Eutecus is shipping its first Instant Vision software-development libraries, too. There are four libraries for signal and image processing, multitarget tracking, and feature classification. In addition, Eutecus offers a development system that can serve as a reference design for an intelligent surveillance camera. (Eutecus calls it the Bi-i intelligent camera.)

The technology for delivering superfast frame rates is highly specialized, but that doesn't mean the market is small. On the contrary, *MPR* believes this technology will tantalize developers and lead to new applications that weren't practical or even imagined before. Consider the historical precedents.

When Louis Daguerre unveiled the basic technology of photography in 1839, the world hailed his ability to capture lifelike images in daguerreotypes. But technological limitations prevented early photographers from recording subjects that couldn't remain still for long time exposures. By the late 1800s, wider-aperture lenses and more-sensitive emulsions permitted instantaneous exposures. Technologically, instantaneous photography was merely a refinement, but in practical terms, it was as revolutionary as the original invention. It conquered time, allowing photographers to capture life on the fly and observe things never seen before. The most famous early example was Eadweard Muybridge's startling stop-action photos proving that a galloping horse momentarily lifts all four feet off the ground—and in a way that centuries of painters had never envisioned.

In the 20th century, high-speed photography using analog technology expanded human vision much further. Harold Edgerton's electronic speedlights froze a bullet penetrating an apple. Special high-speed film cameras revealed the morphology of water droplets and the terrible beauty of nuclear explosions. But, for the most part, these feats required expensive equipment and were limited by mechanical film transports and the camera's inability to respond to what it was seeing.

Ultrafast adaptive imaging could reinvent photography for the 21st century. Soon, it will be possible to make affordable cameras that record *and intelligently respond* to events too elusive for living creatures to see. No longer must the camera be a passive observer. Bonding a sensor array directly to a processor array is a particularly impressive technical innovation, because it can make image processing as massively parallel as image sensing.

Like all new technologies, high-speed digital imaging will have applications both exciting and frightening. But it's another logical extension of human senses, and it gives engineers a new way to infuse artificial intelligence into machines. 

To subscribe to Microprocessor Report, phone 480.483.4441 or visit www.MPRonline.com