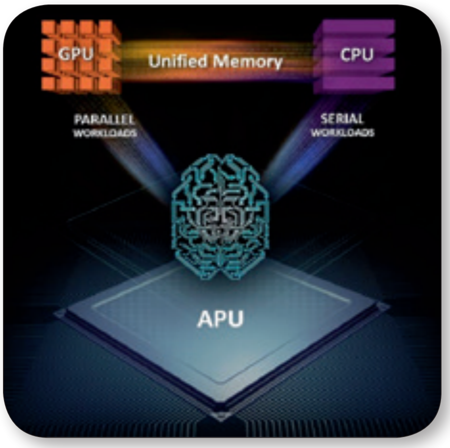
Intelligent vision systems enable totally new eras in robotics

By Fredrik Bruhn, Unibap

Intelligent vision systems are key in one aspect of the three technological trends fuelling the evolution of robotics; high resolution sensors, powerful heterogeneous system architectures (HSAs), and highly efficient brushless DC motors. HSAs like the AMD G-Series SoC offer highly integrated processor architectures in a unified platform.



Vision system innovations are employing the latest heterogeneous system architectures such as the AMD SoCs, which allow software engineers to make full use of the hardware features in a single, unified environment.

■ The ability to comprehend is no longer limited to the animal kingdom; machines are increasingly able to recognize, manipulate and influence the world around them. Technological trends have created the right environment to support advanced robotic systems that now play a crucial role in modern life. The range of applications for robotic systems is vast, as they move from automating the manufacturing process to being the hands and eyes of highly experienced surgeons. The potential for robots to aid an ageing population cannot be overlooked.

Fundamental to these robotic extensions of ourselves will be three key technologies, which are now advanced enough to support what some are already calling the robotic revolution. First comes vision; machine vision systems are now enabled by low cost, high performance sensors that provide much greater resolution than those of even just a few years ago. Next comes the ability to process the data generated by these advanced sensors, an area where massive advancements have been made in recent times, especially in the area of executing deep learning algorithms. Lastly comes movement; here, the great leaps that have been made in the development and efficiency of brushless DC motors provides the third key enabler for advanced robotics. Spatial awareness, enabled through stereoscopic vision, is so natural in the animal kingdom that it makes perfect sense to adapt the same principle for machines, which has given rise to the intelligent vision system. In this application, two high resolution camera sensors provide stereoscopic visual data which is then processed by high performance digital processors. Such systems are now being used with robot arms in assembly applications, while the same technology is fuelling the burgeoning autonomous vehicle industry. Acting as the eyes to robotic systems, intelligent vision systems must now perform a large amount of the data processing closer to the sensors, before passing the processed information on to the main system. This is made necessary because of the large amount of data now being generated by vision sensors, and made possible thanks to the advances made in processor technology. An intelligent vision system would once have been simple frame grabbing with perhaps some pixel binning, performed by a digital signal processor, at the time the most efficient engine for complex algorithms requiring parallel processing.

It is now something more akin to the human eye and visual cortex, enabled by heterogeneous system architectures (HSAs) that combine powerful general purpose microprocessor cores (MPUs) with Graphical Processing Units (GPUs) in a unified architecture. Creating this artificial visual cortex requires a combination of advanced digital processing platforms. Thanks to their parallel nature and hardware efficiency, FPGAs are used for processing individual pixels straight out of the sensors. Like the human eye, cameras have evolved to see in color: red, green and blue (RGB), encoded to display information in way suitable for the human eye. For intelligent vision systems, this representation is less useful than hue (the color circle), intensity (old grey scale) and saturation (how much color or grey) (HIS) and so the first task for a vision system is to convert RGB data into HIS data (using a computer, this conversion could require one core per sensor, but using an FPGA there is almost no area penalty and a delay of just six processing clock cycles).

Vision systems commonly employ lenses to improve their effectiveness, while the use of fish-eye lenses is becoming more common to further extend their field of view. Correcting for the effects of lenses is the next process and at this stage the two images will also be matched to create the stereoscopic image. This processing is typically carried out in the FPGA, while all subsequent processing would be handled by a heterogeneous SoC, such as the AMD G-Series SoC. It is through advances in HSA design, like those found in the AMD G-Series SoC, that they become suitable engines for

Figure 1. Example of what the IVS70 stereo camera sees with standard office corridor lighting and non-optimized decoding at 5 Mpix.

emerging intelligent vision systems. By building on established designs, decades of research and evolutionary artificial intelligence, HSAs provide the ideal platform for the next generation of machine vision systems.

The complexity involved with understanding the world through images should not be underestimated. The vision systems found in nature have evolved over millions of years, yet their digital counterparts have only been in development for mere decades. Nevertheless, software running on advanced processing platforms can now be seen as comparable in its ability to tackle this complex challenge. Neural networks have been in use for many years, but recent leaps in processing power mean their use is no longer compromised by the platform's ability to match their potential, enabling the adoption and development of even bigger and, more importantly, much deeper (more layered) neural networks. Indeed, while limitations in processing power meant neural networks may have needed to be simplified in the past, modern SoCs are more than able to support highly complex networks with many layers.

Other forms of artificial intelligence are also seeing the benefits of more powerful processing architectures. A leader in this field is x86 architecture, which has always been at the forefront of adopting new technologies. It successfully combines instructions optimized for streaming and vector operations, developed over many years, with new technologies such as Shared Virtual Memory. All of these innovations are employed in the latest heterogeneous system architectures, which allow software engineers to make full use

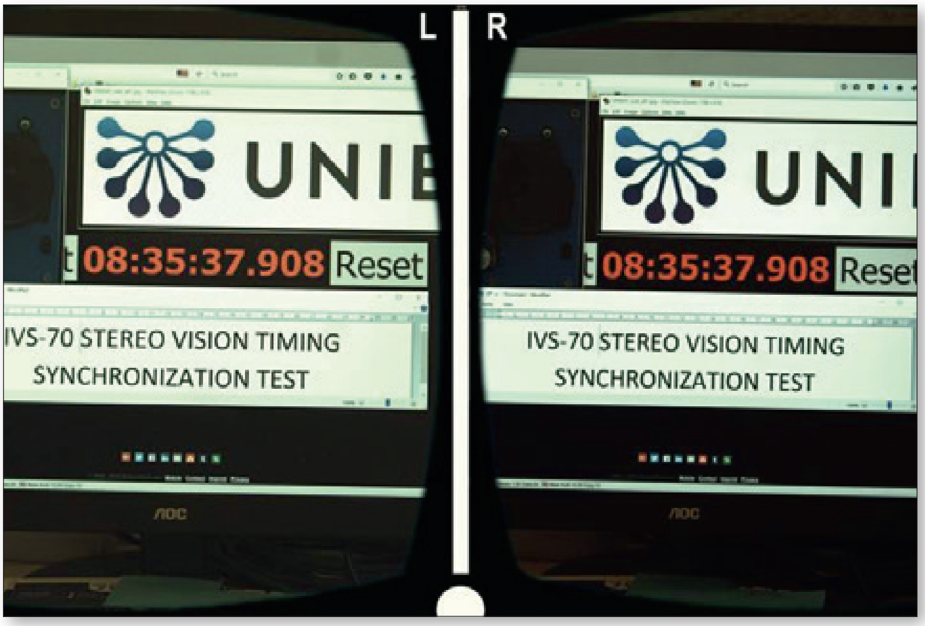


Figure 2. Time synchronization is perfect at the microsecond level, even if the computer screen can show only milliseconds. Picture taken with a 10 Mpix stereo-pair (2 x 5.2 Mpix) of 70 mm lenses.

of the hardware features in a single, unified environment. The latest generation of AMD SoCs represents how the theory of HSA is put into practice. Heterogeneous processing platforms essentially combine processors with different architectures in a single device, which AMD refers to as the Accelerated Processing Unit (APU). This brings together the powerful x86 CPU and its ability to efficiently execute sequential programming, with its highly optimized Graphical Processing Unit (GPU) designed to tackle parallel processing. By harnessing the potential of the GPU for intelligent vision systems, the HSA delivers greater power efficiency and, therefore, a total increase in processing capability.

The software ecosystem for advanced applications like intelligent vision systems also benefits from the HSA approach. Open standards such as OpenCV (Open Source Computer Vision) and OpenCL (Open Computing Language) make software development much simpler and so harnessing the power of an HSA less challenging. Developers can create advanced software applications without needing to focus on partitioning code between the various processing elements in an HSA; the platform and low level software has been developed to help with code partitioning and execution, in order to get the highest performance. This allows many powerful and complex algorithms to be efficiently ported to an HSA.

As intelligent vision systems develop and are deployed in significant numbers to enable the next generation of robotics, society will rapidly come to rely on them. Perhaps the most apparent of these applications will be autonomous vehicles, but their use will be widespread. Their safe operation will be imperative and the integrity of the data stored essential, because decision making (a fundamental feature of advanced robotics) will rely heavily on the integrity of the data available. There would be no human operator present to blame, for example, if an accident occurs due to the malfunction of a care-giving robot.

Highly dense integrated devices are increasingly prone to the phenomena know as single events. These occur when a single ionizing particle comes into contact with a transistor or other integrated element causing a change of logic state. The result, termed a single event upset, can change logic 0 to logic 1 in a memory device, for example, thereby changing the way a piece of code executes or, in other words, the decision an intelligent system makes. Protecting against the effects of single events is taken extremely seriously by the aerospace industry due to the potential consequences and, in part, to the closer proximity of the electronic systems to the cause: cosmic particles colliding with atoms in the atmosphere. While the probability of a single event occurring varies between systems, they could potentially occur once every 100 hours of operation. Protecting against the effects of single events hasn't been a high priority for most applications outside the aerospace industry, but for machine vision systems it could become so.

Independent tests carried out by the NASA Goddard Space Flight Center have demonstrated that the AMD G-Series SoC can withstand a total ionizing radiation dose of 17Mrad(Si). This compares favorably with the specification normally applied for standard space flights, of just 300krad. While a single event could hit anywhere, integrated memory such as SRAM is most susceptible due to its high density, while its importance to code execution is apparent. AMD employs advanced error correction (ECC RAM) in the G-Series SoC, which is able to compensate for the effects of single events. This makes the G-Series even more applicable to intelligent vision systems.

The development and deployment of robotic systems is set to extend beyond industrial automation to all vertical industries. Intelligent vision systems are a key component of that evolution, forming one aspect of the three technological trends fuelling the evolution of robotics; high resolution sensors, powerful heterogeneous system architectures (HSAs), and highly efficient brushless DC motors. HSAs like the AMD G-Series SoC offer a highly integrated and powerful combination of processor architectures in a unified platform. When coupled with open source software ecosystems, developers are empowered to create solutions that allow more advanced deep-learning algorithms to be used, moving machine vision systems and robotics from a sense-compare-decide model to a sense-plan-act behavioral scheme. Machine vision systems like the intelligent vision system IVS-70 from Unibap, enabled by the AMD G-Series SoC and Microsemi SmartFusion2 FPGA, demonstrate how advanced integrated devices and software are coming together with high performance optical systems to deliver machine vision solutions that will empower the robot revolution. ■