

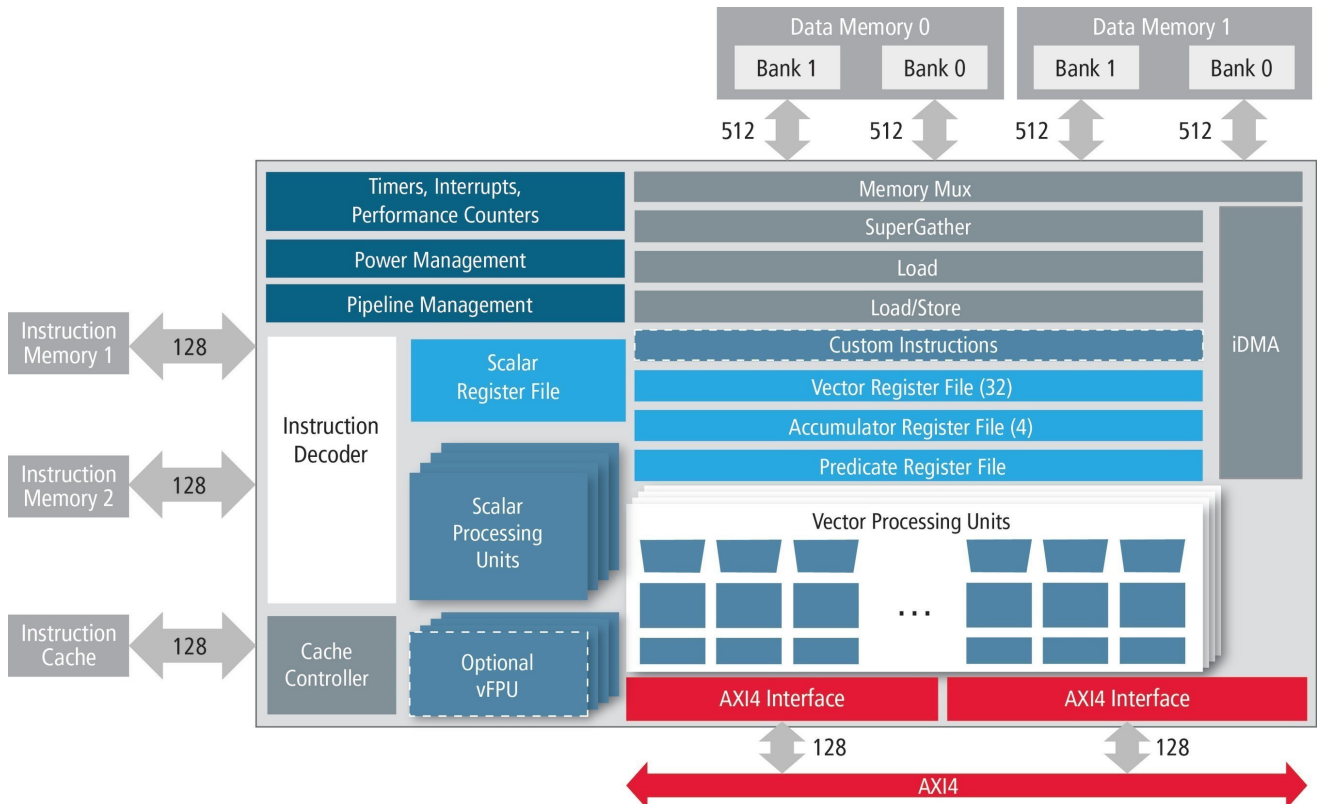
Dedicated Processors Add Power to Machine Vision

Vision is an ever-improving aspect of machines and autonomous systems. The growing role of dedicated vision processors is likely to lead to faster vision systems.

Systems that incorporate vision capabilities have grown more sophisticated over time, [such as Intel's Myriad X](#), but they're not always easier to implement. Graphical processing units (GPUs) and CPUs have been harnessed to process raw information and then squeeze out key attributes from which a system can

make decisions. Relatively new in the field are vision processors, devices optimized for exactly these kinds of machine-vision tasks.

With capable vision processors and increasingly sophisticated software, applications like facial recognition, object detection, and pose estimation, among other challenging



Vision processing units (VPUs) are often built around digital signal processors like Cadence's Tensilica P5. (Cadence Design Systems)

systems, can become reliable functions.

Understanding a Vision Processing Unit

Dedicated vision processing units (VPUs) help systems to quickly process, analyze, and make decisions using visual inputs (*see figure*). All depend on algorithms to properly interpret visual information quickly and accurately.

The emerging technology of the vision processor is optimized around the needs and abilities of AI, and it can be contrasted with well-known GPUs. While GPUs often focus on rasterization and texture mapping in support of 3D graphics, vision processors perform tasks like running machine-vision algorithms such as convolutional neural networks (CNNs) and scale-invariant feature transform (SIFT).

Vision processors can work with edge servers to support intelligent retail, safety and security, and many kinds of automation. And, according to an analyst firm, the market for VPUs is expected to expand at a CAGR of 12.96% through 2030, reaching US\$2872.44 million by 2028.

The key to that growth, of course, is in the applications and software technology that deliver specialized functionality.

Pose Estimation and Other Analytics

Increasingly, artificial intelligence provides the needed sophistication. In pose estimation, for example, an important subset of facial recognition, the face in various attitudes and angles is analyzed in three dimensions to confirm identity and perhaps signal other things about the person being observed.

In a broader extension, human pose estimation can be useful for security tasks or retail analysis. The process includes detecting a figure and identifying and analyzing features like the attitude of the shoulder or whether knees are bent. This is not only a matter of identifying the features, but using models to understand what sequences of movements and positions “mean” in terms of subsequent behavior or actions, or predicting what might happen next.

Similar kinds of analytics can be invoked to better understand and respond to the movement and action of motor vehicles. Velocity; direction; the appearance, color, and actions of indicator lights; and the position of steerable wheels can all help provide indicators of intent or of likely future actions.

With vision processors, system designers can incorporate more features and intelligence into systems, often without incurring high costs or using much power. Many vision processors are matched with model development tools, and APIs that make custom programming comparatively easy. Scaling, development, and customization is also easier if these tools are available.

Vision processors now often support multiple cameras, yielding interpretive results that can match humans. Be-

cause vision processing helps drive intelligence locally—at the edge—it can support systems that are generally more adaptable as well as very timely results.

Expanding Vision Processor Apps Beyond Facial Recognition

Visual processors can be used in many critical real-time applications, e.g., self-driving vehicles, where they can ensure that road hazards are detected and that the host vehicle safely navigates among all other vehicles on the road.

Vision processing units typically include compute and memory elements as well as data paths that can accelerate machine-learning algorithms and the raw activity of image processing. Many include parallel processing abilities to handle multiple tasks simultaneously and deliver near real-time results.

Some harness well-known AI frameworks such as TensorFlow and OpenCV, which in turn make it easy to employ standard tools and libraries.

One of the key features of vision processing is that most of the processing is accomplished locally, which is typically a goal of “edge computing.” It reduces bandwidth demands and latency, while improving privacy and enabling real-time responses.

Real-time capabilities make vision processors a good fit for several well-defined kinds of applications:

- *Unmanned autonomous vehicles (UAVs)*, commonly called drones, need to rapidly process visual information to avoid obstacles, identify subjects of interest, and find their way. Vision processors with the right application software can accomplish much of this with little or no human input, improving productivity and safety.
- *Advanced driver-assistance systems (ADAS)*, now often standard equipment for many vehicles, can achieve better results to help drivers navigate safely when a vision processor is employed to deliver rapid, sophisticated, and accurate assessment of available visual information.
- As discussed above, *facial recognition* is a prime focus for vision processors. New and better capabilities enabled by vision processors can quickly and more accurately identify individuals and clarify their activities—particularly helpful for security-focused activities.
- *Fully autonomous vehicles* need the power that vision processors can provide to replicate and even improve upon human capabilities. Initial capabilities include lane-departure alerts, but vision processors can also spot road hazards, pedestrians, and formal indicators (e.g., road signs), as well as situational indicators (e.g., a vehicle broken down or operating unusually).
- When it comes to the *Internet of Things*, vision processors can enable breakthrough capabilities to see and understand, and gather meaning from, a complex visual

situation.

- *Medical* and *health* applications can also be enhanced by vision processors in areas such as robotic surgery or even diagnoses.

In other words, vision processors are a highly adaptable technology that's just at the early stage of adoption. More remarkable applications to make retail checkout unnecessary, for example, may be on the near horizon.

The specific vision processor you choose will depend on goals, how many cameras must be controlled, ruggedness and power requirements, and the availability of relevant software.

As you consider vision processors, look for the factors that can contribute to project success such as scalability and energy efficiency. Also consider the whole ecosystem. Mature vision processors may have full-feature software and tools, AI models, and related applications that can reduce development time and improve time-to-market.

References

[TRFH: towards realtime face detection and head pose estimation](#)

[A Heterogeneous Architecture for the Vision Processing Unit with a Hybrid Deep Neural Network Accelerator](#)