cadence®

Choosing the Right DSP for High-Resolution Imaging in Mobile and Wearable Applications

By Pulin Desai, Cadence Design Systems

From smartphones to smart watches, from advanced driver assistance systems (ADAS) to virtual-reality gaming consoles, drone control, and a host of security devices, the application areas that rely on high-resolution imaging (1080p, 4K, and beyond) are growing. There's great opportunity to further enhance high-resolution imaging quality to increase the effectiveness of computer vision applications, especially those for mobile and wearable devices. What's needed is a next-generation DSP that can strike the right balance between power consumption and performance. In this paper, we will discuss criteria to consider when choosing the ideal DSP for high-resolution imaging in mobile and wearable applications.

Contents

Introduction

When Steve Sasson, an Eastman Kodak engineer, unveiled the first digital camera in 1975, he gave the world a whole new way to look at photography. Who knew back then how much image resolution would improve and how ubiquitous cameras would become? Today, high-resolution imaging is so sophisticated that we're relying on it for everything from face detection in smartphones to face recognition in security systems to traffic sign recognition in our vehicles, and for the autonomous vehicle of the future.

According to Yole Developpement, the CMOS image sensor industry is expected to grow at a CAGR of 10.6% to become a US$16.2B market by 2020. While smartphone applications comprise the bulk of the market, automotive, medical, and surveillance applications also present market opportunities. On the technology side, camera module height and pixel sizes—currently as small as 0.9µ—are reaching a limitation. With the increased pace of change in new sensor technologies, unique image processing methods are needed to extract the highest possible image fidelity.

Figure 1 outlines the latest approach for the camera pipeline. Image and vision processing algorithms can be implemented in an electronic system using a combination of general-purpose CPUs and/or GPUs, or the hardware pipeline or RTL. Because sensor technologies are changing at a rapid rate and vary widely with different suppliers, the RTL pipeline struggles to keep pace in terms of size and functionality.

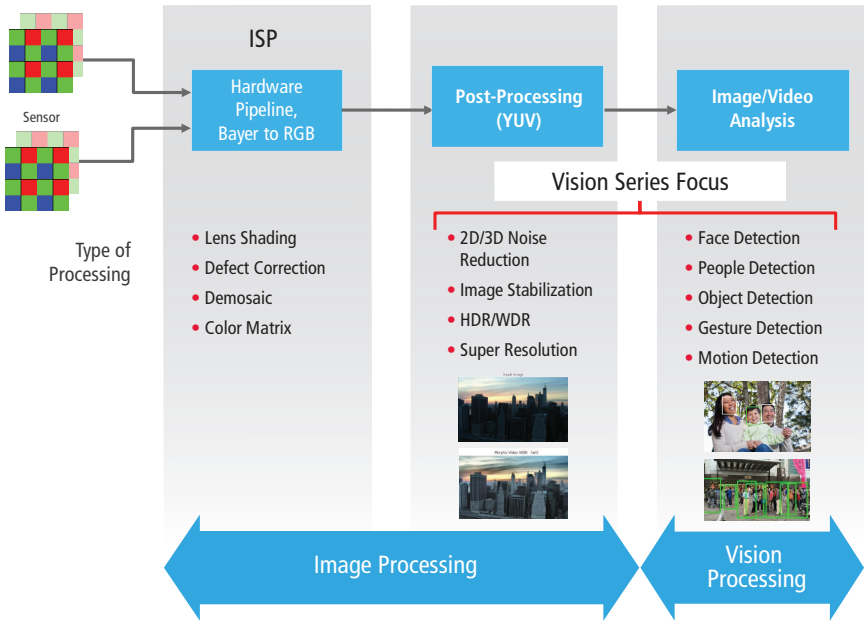

Figure 1: The latest approach for the camera pipeline.

In this paper, we'll discuss a more viable method for 32MP and above image and vision processing: offloading these functions to specially designed DSPs. By choosing the right DSP, you can maintain the functionality needed for your computer vision design without sacrificing performance or power.

Solving the Power/Performance Riddle

While the CPU/GPU route is flexible, one of the biggest drawbacks is the performance/power tradeoff. As Figure 2 shows, running computationally demanding image and vision processing algorithms taxes the CPU/GPU, weighing down the performance for other tasks and consuming a substantial amount of power. Moreover, if your end device is a wearable, the CPU/GPU option will not support the long battery life needed, will likely cause the device to run hot (and slow down the CPU/GPU frequency), and will consume a substantial amount of power, particularly for applications such as video chat.

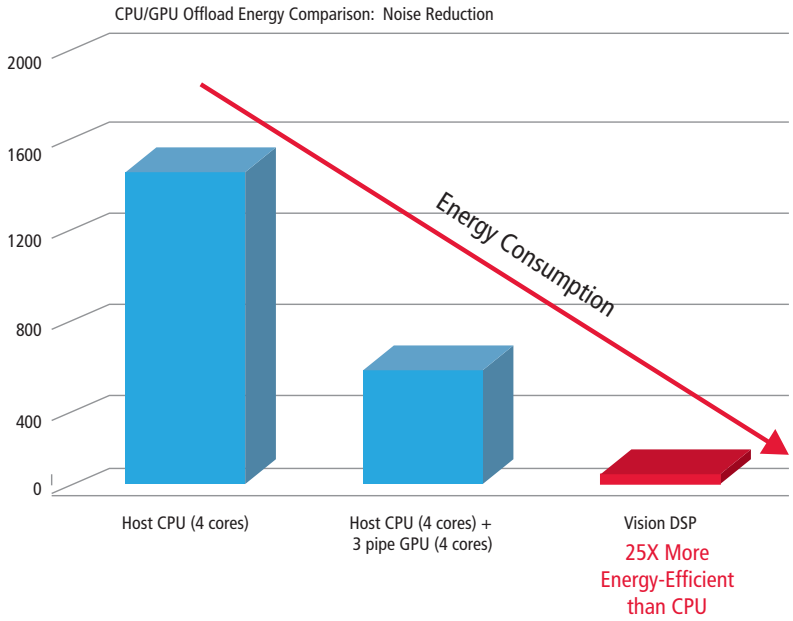

Figure 2: Offloading image and vision processing functions to a DSP results in substantial energy consumption savings.

The other option, running the algorithms on an RTL-based hardware design, provides a performance/power advantage, but limits your flexibility. Hardware design requires a lengthy design and verification process, and each function would require its own gates—a costly proposition. And should anything change in your algorithm—to address a new requirement or sensor, or fix a bug, for example—you'd have to perform an expensive hardware re-spin.

Offloading to a specially designed DSP presents a solution that balances performance and power requirements, provided you have the right DSP. A camera pipeline with a DSP consumes a smaller amount of total logic, as the DSP can be reused for multiple algorithms.

What Kind of DSP Can Support Computer Vision Applications?

Computer vision—which involves duplicating the abilities of human vision by electronically perceiving and understanding an image—has been around since the 1950s. First used in academia, computer vision is becoming prevalent in mobile wearable devices, as well as automotive and drone applications. In today's consumer devices, computer vision is driven by low energy consumption (for a longer battery life), high performance, and killer use cases. Wouldn't you enjoy video chatting with a far-away family member on a tablet featuring razor-sharp 4K resolution? Or, consider the possibilities if the same technology could support functional measurements so that users could determine, for instance, whether a piece of furniture seen online could fit in their bedroom.

Camera processing systems for these applications consist of two key components—image processing and vision processing—each with its own requirements that vary depending on the end device.

Image Processing

- Sophisticated noise reduction to accommodate small pixel size
- Video image stabilization capabilities
- Fast auto-focus capabilities
- High dynamic range (HDR)
- Support for dual cameras in the same device
- Support for wide dynamic range (WDR), digital video stabilization, and multiple advanced noise reduction methods

Vision Processing

- Face detection, with resolution ranging from the standard video graphics array (VGA) level to 720p
- People and object detection
- Gesture tracking
- Eye tracking capabilities
- Motion detection
- Convolutional neural network (CNN)

Continually improving image resolution is also driving requirements for high memory bandwidth. While today's smartphones can deliver as many as 16MP, tomorrow's phones are expected to have resolution as high as 24MP to 32MP. Similarly, smartphones are already delivering 1080p video and will eventually move to 4K, with videos running at 120fps.

With these types of imaging and vision processing functions and the memory bandwidth demands, flexibility is key for a computer vision DSP. After all, optics, sensors, and algorithms are constantly changing. Supporting these high-resolution image and vision processing functions and memory bandwidth demands requires a highly flexible computer vision DSP with a variety of features:

- Support for multiple algorithms, such as face detection, gesture detection, motion detection, image stabilization, HDR, WDR, and 2D/3D noise reduction

- Programmability, so that a single DSP can support multiple end applications—design requirements will vary
- High performance and low power consumption
- Ability to be placed inside the ISP processing path for processing in the Bayer domain to support functions such as noise reduction and HDR or to support new types of sensors (IR sensors, or those that detect a mix of R, G, B, and IR pixels, for example)
- Small form factor

Developing such DSPs calls for highly sophisticated development tools and a highly optimized imaging and vision library. Some applications require high-precision math, so the solution would need to be able to perform vector floating-point operations.

New Image and Vision Processing DSP

Cadence's Tensilica® Vision P5, the newest DSP in the Tensilica family, is an example of a specially designed processor for image and vision processing applications. Figure 3 shows the architecture and key features of this fourth-generation DSP.

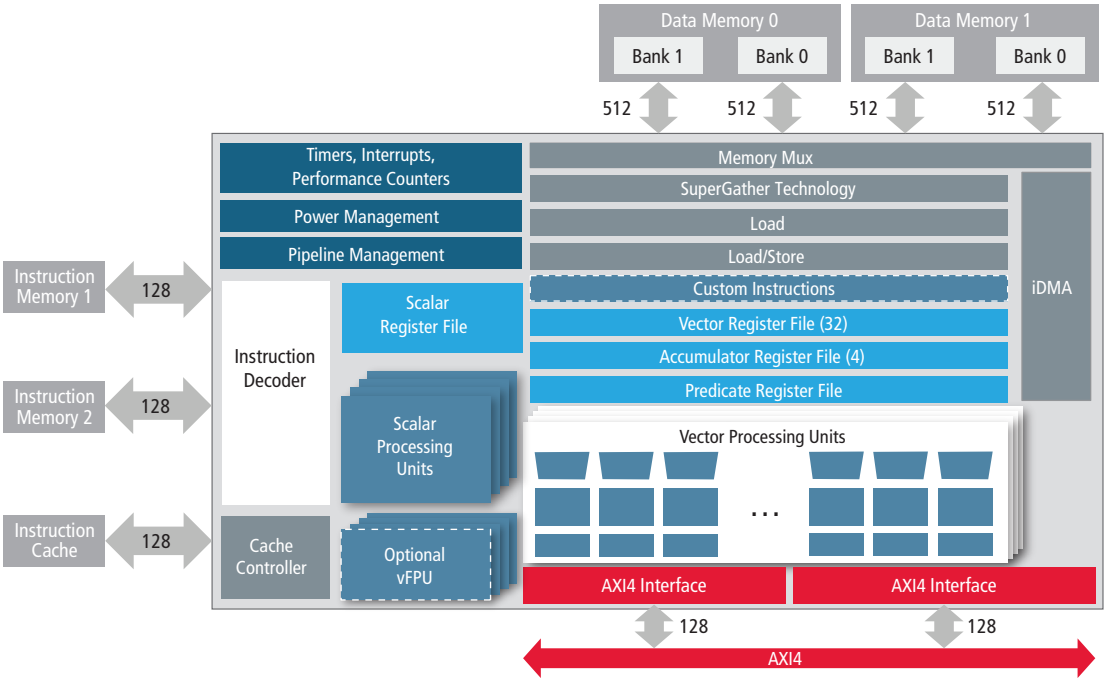

Figure 3: Tensilica Vision P5 architecture

The Tensilica Vision P5 offers features ideal for image processing and analytics as well as video pre- and post-processing:

- Step function in vision performance over the previous-generation Tensilica IVP-EP DSP
- 64-way SIMD engine
- 1024-bit memory interface with SuperGather™ technology
- Enriched support for 32-bit data types
- Customizable for even higher performance
- Support for a large number of OpenCV and OpenVX library functions, as well as a number of imaging/vision kernels for filtering, object detection, and optical flow
- Multi-processor ready

Developed on the configurable Cadence Tensilica Xtensa® processor, the Tensilica Vision P5 has an integrated 2D iDMA that provides high-bandwidth data transfer between local and system memory. The processor features an optional vector floating-point unit, error correction code (ECC) support for automotive qualification, and an ARM® AMBA® AXI4 interface to the system bus. The Tensilica Vision P5 offers a significant performance improvement over its predecessor (see Figure 4).

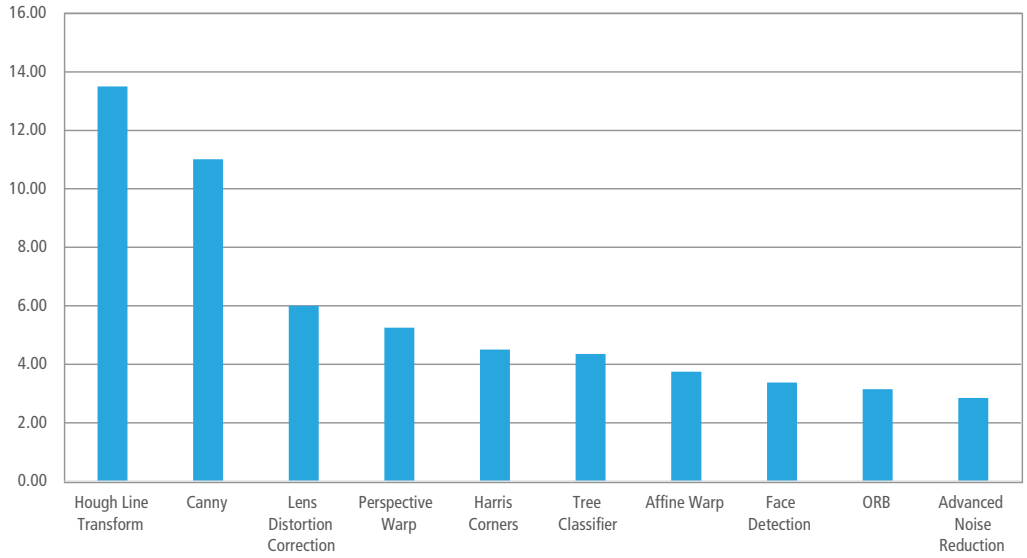

Figure 4: Performance enhancement of Tensilica Vision P5 vs. previous-generation DSP.

Use Cases

Morpho, Inc., an image processing technology company based in Tokyo, Japan, used an earlier generation Tensilica image/vision processing DSP for its video WDR, a single image-based luminance enhancement technology. By porting the video WDR to the DSP, Morpho achieved an 8X performance gain compared to its previous CPU-based solution. On the CPU, 1080p resolution at 30fps required 2.2GHz, compared to only 282MHz with the DSP, as shown in Figure 5. Since the Tensilica Vision P5 provides 2X to 4X better application-level performance, the performance number could be as low as 70MHz.

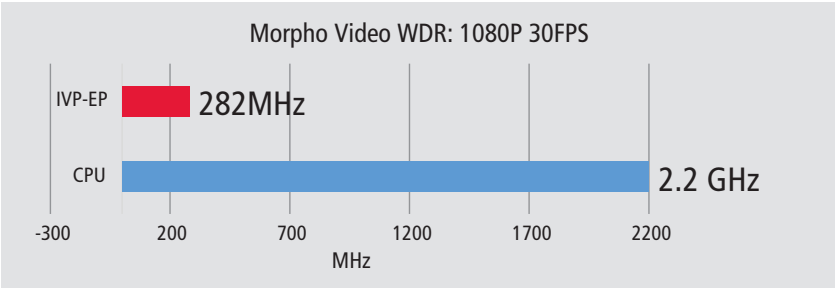

Figure 5: Morpho achieved a substantial performance gain in its video WDR application after migrating to Tensilica DSP technology

Uurmi Systems also took advantage of the Tensilica image/vision processing DSP family's energy efficiency, achieving a 15X fps/watts improvement over a mobile CPU+GPU solution in its fog removal application, as shown in Figure 6. Originally implemented in the previous-generation Tensilica IVP-EP DSP, this application is expected to achieve even greater benefits with the Tensilica Vision P5 because the new DSP is 5X more energy-efficient vs. its predecessor.

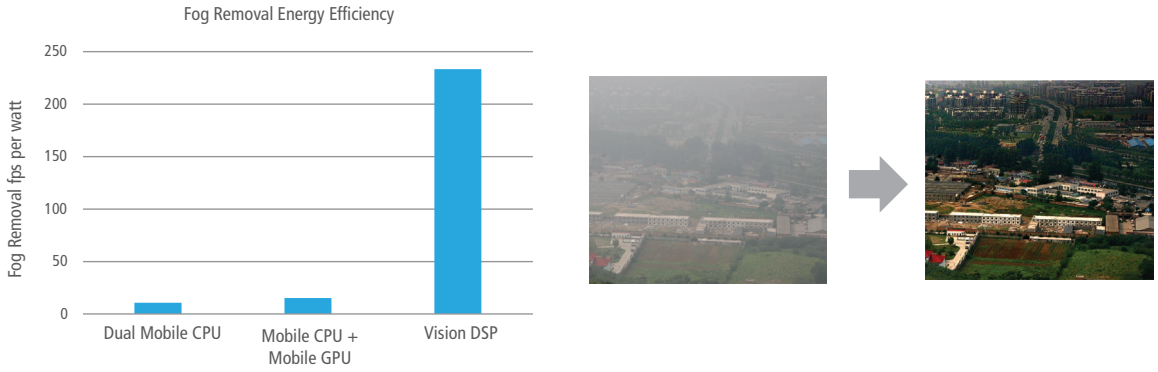

Figure 6: Uurmi Systems lowers energy consumption in fog removal application

Summary

Running computer vision algorithms on CPUs or a combination of CPUs/GPUs is a common approach, but has limitations, particularly with power consumption in mobile and wearable applications. In contrast, offloading to a DSP such as the Tensilica Vision P5 can yield substantial energy consumption savings while supporting the high workloads of image and vision processing applications.

For Further Information

Learn more about the Tensilica Vision P5 DSP at http://ip.cadence.com/ipportfolio/tensilica-ip/image-vision-processing

Footnote

1 Yole Developpement. (2015). The CMOS Image Sensors industry is about to change, with major investment in manufacturing & design [Press release]. Retrieved from http://www.yole.fr/CIS_2015.aspx#.VeiDtM6aSm4

cadence®

Cadence Design Systems enables global electronic design innovation and plays an essential role in the creation of today's electronics. Customers use Cadence software, hardware, IP, and expertise to design and verify today's mobile, cloud and connectivity applications. www.cadence.com