Enhancing Performance in Automotive Vision Systems with Advanced Image Signal Processing

Manjunath Somayaji, Ph. D., Senior Director of Imaging R&D, indie Semiconductor

Previously, Donald McClymont, indie Semiconductor's co-founder and CEO authored a blog, In Pursuit of the Uncrashable Car, in which he described how important it is to integrate multiple sensing and processing modalities into vehicle architectures. The spirit of such a lofty goal is evident in the automotive industry's constant pursuit of safer vehicles towards initiatives like Vision Zero, wherein no more human lives are lost to automotive accidents. Automotive original equipment manufacturers (OEMs) thus continue to push the performance requirements of vehicular sensing and processing modalities. One such modality is 'vision' – that is, camera-based sensing in and outside the vehicle.

An automotive camera's performance depends on many factors. A quality camera needs a good lens, a high-fidelity digital image sensor, and an advanced camera video processing engine to retain, extract and highlight critically relevant information of a vehicle's environment. Central to the video processing engine is the Image Signal Processor – or ISP – that can make or break a car's ability to detect and respond to safety-critical situations on the road.

Several attributes contribute to a reliable ISP. For one, it should capitalize on an image sensor's ability to see farther while ensuring near-real-time responses to ever-changing road and environmental conditions. It should not be overwhelmed by the brightest of lights while simultaneously preserving features in the darkest regions of a scene (such as shadows or poorly lit areas). It should support a variety of image sensor technologies and spectral combinations while maintaining a high degree of color fidelity. Moreover, it should be capable of supporting both human vision applications such as surround view, backup assist and e-mirrors, and also machine vision (sensing) applications which, is being deployed in advanced driver assistance systems (ADAS). Each of these applications can have very different requirements. Oftentimes it may require support for multiple video streams concurrently at very high pixel throughput rates. The performance bar can indeed be high when human safety is at stake.

Seeing Farther, Faster

It is well-known that the faster a vehicle is moving, the longer the braking distance is required to stop it. Twice the speed equals four times the energy needed to bring the vehicle to a complete halt. In such scenarios, the farther ahead one can identify hazards, the greater the ability for a timely response. ADAS systems rely on high-resolution cameras for this very reason. Moreover, these cameras must operate at sufficiently high frame rates to enable real-time detection and reaction to constantly evolving situations. Figures 1a and 1b illustrates this point. As shown in Figure 1b, the higher-resolution image, an ADAS system has a greater chance of success reading road signs or detecting the pedestrian. For this reason, it is not uncommon to see image sensors being deployed in vehicles with cameras achieving eight or more megapixel resolutions at 30 frames per second (fps). That is two hundred and forty million pixels every second! An ISP would need to process every single one of these pixels and do so in real-time with very low latency (delay); at low power to not overheat the camera; in a small-sized form factor; and at a price-point that will align with mass-market deployment for automotive OEMs.

Figure 1a: In this low-resolution image, machine vision could struggle to read signage or detect pedstrians.

Figure 1b: in this higher-resolution image, a machine vision application could perceivably read the signs or detect the pedestrian in the blue shirt.

Very few commercial ISPs today can meet this need. indie's offering of advanced camera video processor (CVP) system-on-chips (SoCs) feature proprietary technology optimized to address these challenging requirements while remaining compact and cost-effective. The GW54x0 family of processors – can easily handle 8MP@30fps. However, indie's newest generation CVPs, the iND880xx family, can support multiple cameras at such high resolutions and frame rates. The iND880xx features multiple ISP pipelines within a single low-power chip, which allows mixing and matching of many sensors at different resolutions and frame rates to handle multiple applications simultaneously – for both human vision and machine vision (sensing).

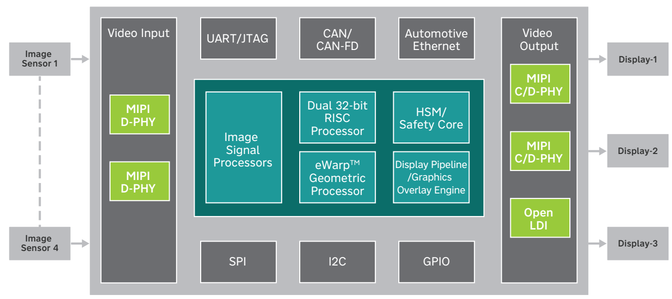

Figure 2: Functional diagram of indie's newly released iND880xx Camera Video Processor (CVP)

High Dynamic Range (HDR) Imaging

Humans have been gifted with amazing, high-dynamic-range color vision. Our eyes have a logarithmic response to light – that is, we can effortlessly see extremely bright regions and dark areas at the same time within a scene. Silicon-based image sensors on the other hand, have a linear response to light. This means a native image sensor pixel struggles to see what the human eye is capable of. Changing the camera's exposure to light – measured in f-stops or decibels – does little to change its dynamic range, which is a property of the pixel's photon-collecting capacity. Human eyes can accommodate up to 20 f-stops – or 120dB – of dynamic range, whereas most native image sensor pixels – automotive or otherwise – struggle to achieve anything meaningful above 12 f-stops, or 72dB.

The impact of such a deficiency in an automotive camera can be detrimental to vision-based viewing and sensing. Take the examples in Figure 3a and 3b where a car is exiting a parking garage. In Figure 3a, the image was taken with a state-of-the-art consumer camera, and in Figure 3b, the image was taken with an automotive camera powered by our GW5xx vision processor. It is clear the consumer camera is not designed for automotive safety needs, as any machine-vision detection algorithm reliant on it would likely fail to detect the vehicle exiting the garage with potentially severe consequences.

Figure 3a: This image was taken with a state-of-the-art consumer camera, a machine vision algorithm would most certainly fail to detect the vehicle exiting the garage, leading to potentially dangerous consequences.

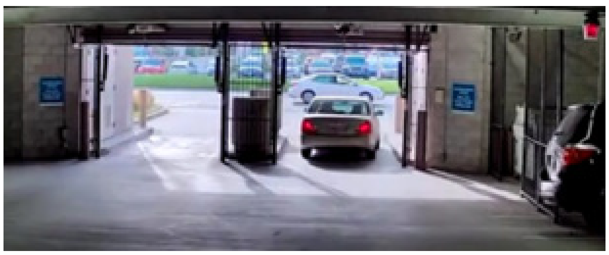

Figure 3b: This image, taken with an automotive camera powered with indie's GW5xx vision processor, is reproduced clearly, capturing details both inside and outside the garage.

Image sensor providers have long sought to address this shortcoming through various HDR image capture technologies. Some capture multiple exposures, others use split pixels, yet others use multiple charge storage techniques in the pixel circuitry. The result is one or more image frames, low dynamic range in themselves, but each bracketing a different region of the scene's overall dynamic range. These must be combined either in the sensor or the ISP to create the high dynamic range imagery that is critical to automotive safety. Today's automotive image sensors can generate up to 24 f-stops or 144dB of dynamic range, perhaps even more. A higher dynamic range means more bits per pixel and a wider bus. An ISP must be able to handle such a wide pixel bus while maintaining fast data throughput and process the images into meaningful low-bit-depth outputs suitable for downstream consumption without losing details in the process. The iND880xx family of solutions can be paired with the latest HDR image sensors from all sensor vendors given its 144dB ISP pipelines.

Alternative Color Filter Array Support

Low light performance is another important attribute in automotive cameras. Detecting a pedestrian that may be wearing dark clothing at nighttime can be quite a challenging task. Yet, government regulations such as those from the National Highway Traffic Safety Administration (NHTSA) in the United States and the European Union's General Safety Requirements II (GSR II) are making automatic emergency braking (AEB) a legal requirement, with the latest NHTSA mandate also extending to nighttime pedestrian AEB. While the traditional Bayer pattern color filter array (CFA) (for more information, please read our blog Improving Driver Safety with Advanced Signal Processing and Color Filter Arrays) has stood the test of time for its color reproducibility and resolution, it comes at a cost of throwing away two-thirds of incident light in each of the red, green and blue pixels (see Figure 4). To address this, vendors have produced image sensors with unconventional CFAs in their quest to collect more photons in low light conditions. These unconventional sensors yield much better low light performance than Bayer sensors, as illustrated in Figure 5a and 5b.

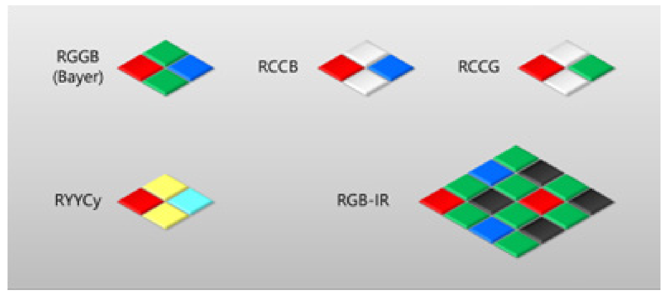

Figure 4: Color filter array (CFA) patterns supported by indie's ISPs.

Figure 5a: Low-light performance of an image sensor with a conventional Bayer CFA.

Figure 5b: Low-light performance of an image sensor with an unconventional (RCCB) CFA.

While the Bayer pattern is a variant of red, green, and blue color filters in a mosaicked pattern, the newer CFAs incorporate clear pixels (as in RCCB or RCCG) or yellow and cyan pixels (as in RYYCy). There is also the RGB-IR image sensor wherein color and infrared (IR) images are simultaneously captured in a single sensor. The ability to successfully deploy non-Bayer CFAs has been a challenge for many years. Color must be reproduced when photons representing only a single fundamental additive color (either red, green, or blue) are inherently absent from a color channel's signal. Color noise from the weaker signals can contaminate the stronger channels' signals during restoration, thereby negating any low-light advantage. It is also easy for colors to be misconstrued by the color correction algorithms. In RGB-IR sensors, all color channels are contaminated by the IR signal and must be cleaned through proper IR estimation to faithfully recover the RGB signals. indie's vision processors are at the leading edge of this technology, overcoming the above challenges in their ISPs. indie's previous generation GW54xx family and the iND880xx family are capable of reproducing high-fidelity colors from all the CFAs shown in Figure 4.

A good performance ISP must not only reproduce color faithfully, but it must also minimize image artifacts due to challenges inherent in the scene. Take for example the night scene in Figures 6a and 6b. The glow of lights can give rise to extensive halos, masking nearby features, while bright white lights can suffer from purple fringing. Red lights can turn yellow when using RCCB image sensors, creating conflicting data for the perception processing. An ISP must be able to overcome such challenges routinely. The iND880xx family, with its state-of-the-art color processing technology, has overcome these challenges.

Figure 6a: A night scene with halos around lighting and purple artifacts on the street lights.

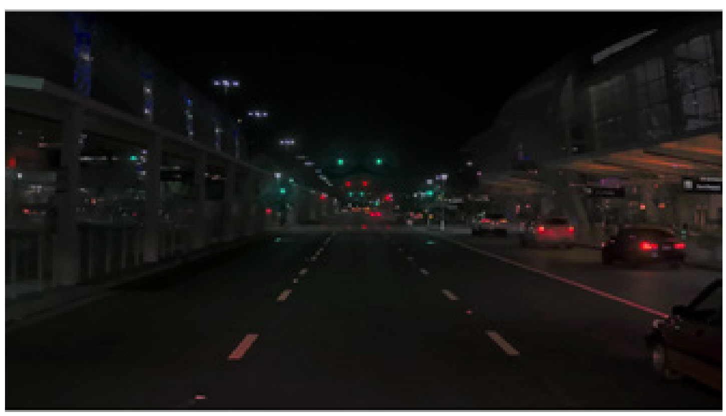

Figure 6b: Here is that same night scene with advanced color artifact processing technology. The signal lights are free of halos; red lights are faithfully reproduced as red; and the streetlights are free of purple fringe artifacts.

Today, indie is at the forefront of camera video processing technology. The ISP plays a central role in indie's aspiration towards safer cars and a future where automotive fatalities are a thing of the past. indie's leadership in this regard is reflected in the over 40 million of its vision processors deployed in cars globally. The next-generation iND880xx and iND881xx vision processors have further pushed the boundaries by addressing hitherto-unsolved problems in camera image quality, thereby inching closer to our vision of the Uncrashable Car.