WHITE PAPER

Getting the Most Out of Your Image-Processing Pipeline

By
Jianping Zhou,
Member of Technial Staff,
Texas Instruments

Overview

Often, the difference between two digital cameras can be measured in more than just the number of pixels each captures. Today, multi-megapixel cameras are available for a wide range of applications including digital still cameras, camcorders, camera phones and video surveillance equipment. Now, the image-processing pipeline plays a key role in the overall image quality since the increasing complexity with image processing is required by a corresponding increase in image resolution.

Developers are able to optimize each of the image-processing pipeline's multiple stages to control image and video quality, while continuing to minimize complexity in order to accommodate diverse user preferences. The integrated, hardware-based image-processing pipeline provides the performance as well as the flexibility required to ensure that developers can produce the highest quality images available for their applications.

The image-processing pipeline performs the baseline and enhanced image processing, which takes the raw data produced by a camera sensor and generates the digital image that will then be viewed by the user or undergo further processing before being saved to nonvolatile memory. This pipeline is a series of specialized algorithms that adjusts image data in real-time and is often implemented as an integrated component of a system-on-chip (SoC) image processor. With an image pipeline implemented in hardware, front-end image processing can be completed without placing any processing burden on the main application processor. This allows the cycles to encode and perform advanced processing functionalities such as video analytics including object recognition and object tracking.

On the other hand, quality has no standard metric and different applications approach quality in different ways, since the same camera can even be used in a variety of ways. A consumer digital still camera needs to be able to take quality pictures in various lighting scenarios, such as indoors, in bright sunlight and in relative darkness. In order to maximize quality for these scenarios, the image-processing pipeline needs to be readily flexible and configurable to provide the highest quality for each individual picture.

Front-end digital image processing not only involves reading pixels off the sensor, but the image quality needs to be dynamically configurable to match variations in user preferences.

To maximize the quality of the products, TI digital media processors based on DaVinci™ technology, such as the TMS320DM6446 processor, integrate an image-processing pipeline that has been implemented in a configurable manner to combine the performance advantages of hardware-based processing with flexible control to enable fine tuning of algorithms for quality. In many cases, fine tuning the image-processing pipeline allows developers to compensate for the use of lower-cost and consequently, lower-quality components.

Understanding the Image-Processing Pipeline

The image-processing pipeline is designed to exploit the parallel nature of image-processing algorithms and enable a camera to process multiple pictures simultaneously while maximizing final image quality. Additionally, each stage in the pipeline begins processing as soon as image data is available so the entire image does not have to be received from the previous sensor or stage before processing is started. This results in an extremely efficient pipeline with deterministic performance that increases the speed with which images can be processed, and therefore the rate at which pictures can be taken by users.

Performance, however, is only one factor that influences overall camera quality. The flexibility of DaVinci technology processors enables developers to specifically tune individual stages to match the particular sensor and lens combination of a camera. Additionally, developers can tune sensor/lens combinations across a wide range of operating conditions under which the camera might be used.

Taking a hardware-based, configurable approach guarantees performance and flexibility. Certainly, ASIC implementations provide excellent performance and low cost, but their fixed nature fails to achieve the best quality possible under a wide range of operating conditions and also limits their ability to adapt to multiple applications. On the other hand, while software-based approaches provide the required flexibility, they require too many MIPS on the main application processor and consume too much power.

As front-end image processing is fairly well-defined, the algorithms involved are well-suited to a configurable approach. This even includes those stages which tend to be proprietary between vendors, such as CFA interpolation. Through a configurable approach, developers maximize overall image quality by adjusting specific parameters at each pipeline stage. Additionally, TI DSPs also allow the developers to disable certain stages when necessary.

Getting the most out of the image-processing pipeline is a process of fine tuning each pipeline stage for every sensor and lens combination. Note that if the camera is intended to be used for a fixed application, one set of parameters may be enough. For more versatile cameras, however, developers will want to take multiple measurements and calculate the appropriate parameters for each set of applicable operating conditions. For example, most digital still cameras offer user-selectable modes to adjust the camera for cloudy, sunny or night time lighting. Tuning the image-processing pipeline for each of these modes enables users to assist the camera in achieving the best image quality. Alternatively, a camera can automatically evaluate current lighting conditions and make an intelligent selection between the available modes to maximize the quality of the particular image being captured.

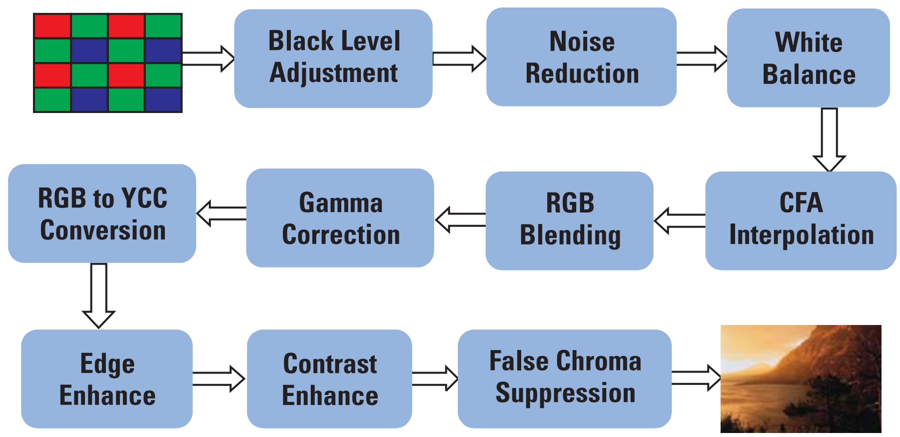


Figure 1. DaVinci imaging-processing pipeline.

Fine Tuning the Image Pipeline

The image-processing pipeline integrated in TI's DaVinci™ digital media processors has 10 distinct stages. By understanding the function of each of these stages and how each stage can affect the final image, developers can maximize quality.

Black-Level Adjustment

This stage in the pipeline adjusts for dark current from the sensor and for lens flare, which can lead to the whitening of an image's darker regions. In other words, sensor black is not the same as image black. The most common method for calculating this adjustment is to take a picture of a completely black field (typically accomplished by leaving the lens cap on), resulting in three base offsets to be subtracted from the raw sensor data. Failure to adjust the black level will result in an undesirable loss of contrast.

$$R'_{i,j} = R_{i,j} - O^{R}_{i,j}$$
$$G'_{i,j} = G_{i,j} - O^{G}_{i,j}$$
$$B'_{i,j} = B_{i,j} - O^{B}_{i,j}$$

Noise Reduction

There are numerous sources of noise that can distort image data – optical, electrical, digital and power – which must be removed before they are amplified in later pipeline stages. The actual noise level present in an image, however, plays a critical role in determining how strong the noise filter must be since the use of a strong filter on a clean image will actually distort and blur the image rather than clear it up.

Noise reduction is achieved by averaging similar neighboring pixels. Through the use of an Optical Electrical Conversion Function (OECF) chart (Figure 2 on the following page) and a uniform lighting source, the noise level can be characterized for different intensities.

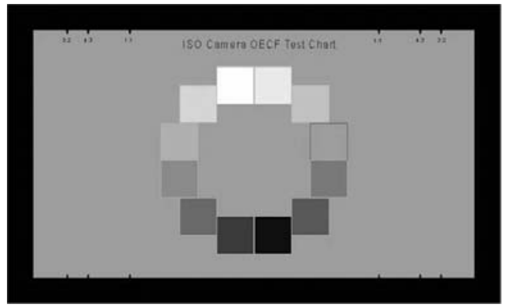
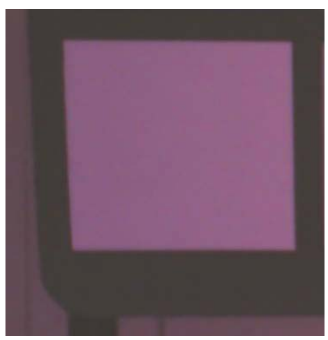


Figure 2. Standard Optical Electrical Conversion Function (OECF) chart.

If the noise level is high for a particular intensity, then more weight is given to the average pixel value of similar neighbors. On the other hand, if the noise level is low, more weight is given to the original pixel value. The OECF chart is comprised of 12 uniform gray patches and produces 12 corresponding power levels based on the noise standard deviation at the mean value for each intensity/luminance level. These 12 power levels are then used to reduce noise across an image using either a linear or square-root model, depending on the sensor and gain (or ISO) level.



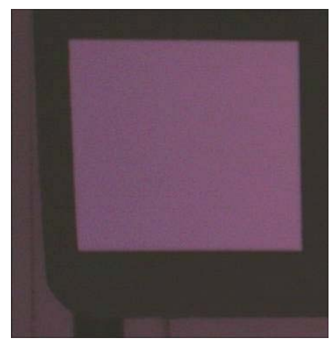


Figure 3. With noise filtering on *With noise filtering off*

White Balance

Different types of lighting – such as incandescent, fluorescent, natural light sources, XE strobe and W-LED flash – have a pronounced effect on color. The most difficult to tune is in mixed-light conditions. White balance automatically compensates for color differences based on lighting so white actually appears white. See Figure 4 on the following page

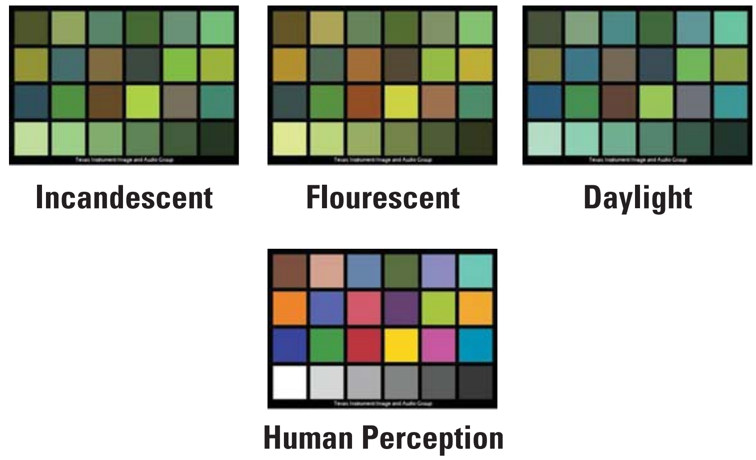


Figure 4. Standard Optical Electrical Conversion Function (OECF) chart.

Fine tuning white balance begins by measuring the average RGB values across the six gray patches on a ColorChecker chart (Figure 5 bottom row).

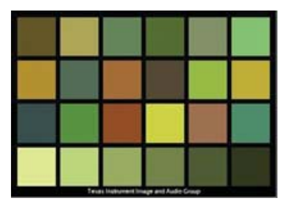


Figure 5. Colorchecker chart.

Using mean square error minimization, the appropriate gains for each color can be calculated. Setting the green gain to a default of one eliminates one set of calculations later in the image-processing pipeline.

$$\min f_R(W_R) = \sum_{n=1}^{6} (W_R \times R_n - G_n)^2$$

$$\min f_B(W_B) = \sum_{n=1}^{6} (W_B \times B_n - G_n)^2$$

Figure 6. Mean square error minimization.

The resulting gains are applied to each image pixel:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} R \times W_R \\ G \times W_G \\ B \times W_B \end{bmatrix}$$

CFA Interpolation

Typically, digital cameras employ only a single sensor to capture an image, so the camera can only obtain a single color component for each pixel even though three components are necessary to represent RGB color.

CFA interpolation is the process of interpolating two missing color components for each pixel based on the available component and neighboring pixels. CFA interpolation is primarily a transform function that does not vary based on sensor or lighting conditions, and therefore no tuning of this image-processing pipeline stage is required. However, it is still one of the most complex algorithms in the image-processing pipeline, and the quality of its output is highly dependent upon the expertise of the silicon vendor. TI digital media processors provide superior CFA interpolation technology to ensure the images are as close to what the user sees as possible.

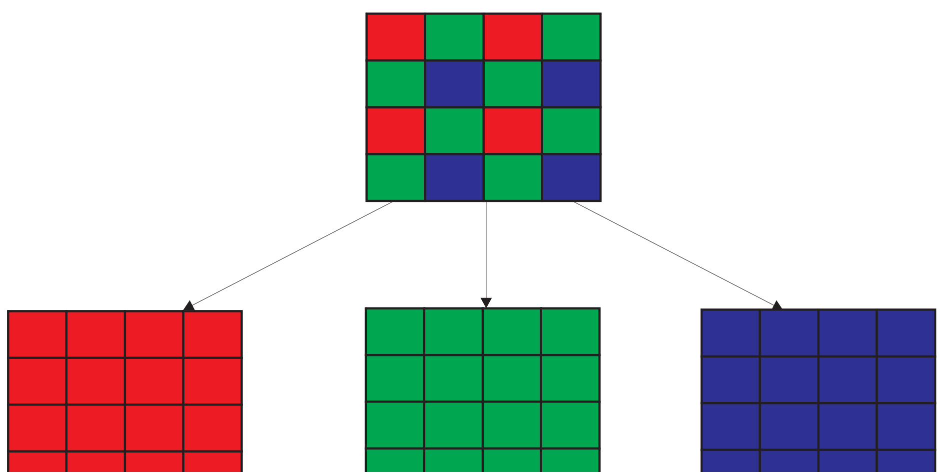


Figure 7. CFA interpolation.

RGB Blending

Different sensors produce different RGB values for the same color. Tuning this pipeline stage involves creating a blending matrix to convert the sensor RGB color space to a standard RGB color space such as the Rec709 RGB color space.

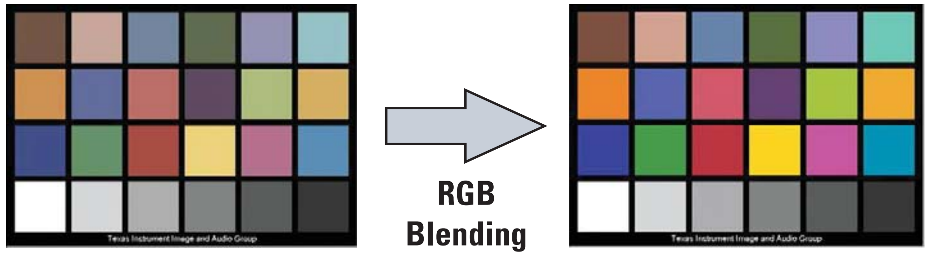


Figure 8. Sensor RGB color space *Rec709 RGB color space.*

The blending matrix is calculated by starting with a ColorChecker chart and obtaining average RGB values for 18 different color patches (the top three rows of the chart) that have already been white balanced. Next, inverse Gamma correction is applied to the reference RGB values. The blending matrix is then constructed using constrained minimization.

$$\min f(\mathbf{M}) = \sum_{n=1}^{18} \sum_{i=1}^{3} \left(\sum_{j=1}^{3} M_{i,j} \times RGB_{j,n} - RGB_{i,n}^{ref} \right)^2$$

$$\textit{subject to } \sum_{j=1}^{3} M_{i,j} = 1$$

Figure 9. Constrained minimization.

The blending matrix is applied as follows:

$$\begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} & M_{13} \\ M_{21} & M_{22} & M_{23} \\ M_{31} & M_{32} & M_{33} \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

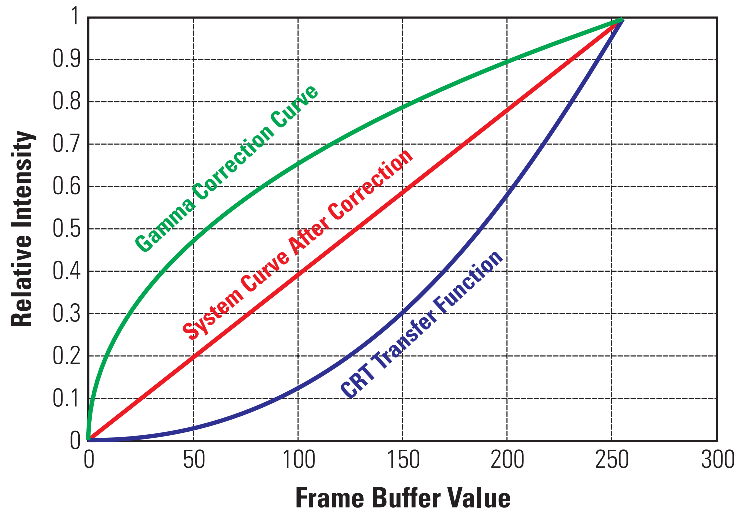
Figure 10. Blending matrix.

The final result is consistent color between cameras using different sensors.

Figure 11. Before RGB tuning *After RGB tuning.*

Gamma Correction

Gamma correction compensates for the nonlinearity of relative intensity as the frame buffer value changes in output displays.



$$R' = R^{1/\gamma}$$
$$G' = G^{1/\gamma}$$
$$B' = B^{1/\gamma}$$

Typically, displays are calibrated using a standard gamma correction such as Rec709 or SMPTE240M. Calibrating the image-processing pipeline to the same standards ensures optimal image quality across the majority of displays.

Under most circumstances, this image-processing pipeline stage does not require tuning. Tuning only comes into play when specialized displays utilize a different gamma correction, such as those used in airport computers or military applications. Developers will need to configure the pipeline with the appropriate gamma correction lookup table for the display.

RGB-to-YCC Conversion

Images also need to be adjusted for the human eye, which is more sensitive to luminance (Y) than color (Cb, Cr) information. This pipeline stage separates luminance from color for different processing using different precisions.

$$\begin{pmatrix} Y \\ C_b \\ C_r \end{pmatrix} = \begin{bmatrix} 0.2989 & 0.5866 & 0.1145 \\ -0.1687 & -0.3312 & 0.5000 \\ 0.5000 & -0.4183 & -0.0816 \end{bmatrix} \times \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

Figure 13. Standard formula for RGB-to-YCbCr conversion.

This transformation is based on a standard formula and is a straight conversion that requires no tuning.

Edge Enhancement

Edge enhancement affects the sharpness of an image. The default parameter for this pipeline stage performs well across a wide range of applications, so tuning is not usually required. For those applications which require (or prefer) a sharper or smoother image, edge enhancement is a single parameter whose strength can be adjusted appropriately with larger values providing stronger edge enhancement.

Modifying the default parameter will require corresponding adjustments to the noise filter since edges in low-quality images can be difficult to distinguish from noise. Alternatively, the block can be disabled depending upon the application. For example, in low-light conditions where images have a lower signal-to-noise ratio (SNR), edge enhancement will boost noise making it more visible to users.

Contrast Enhancement

Contrast enhancement is comprised of two parameters: contrast and brightness.

Since the optimal contrast and brightness vary based on the particular lighting conditions, as well as upon user preference, these parameters are often implemented so that they can be dynamically selected by the user. For normal daylight lighting conditions, high contrast without clipping highlights often provides the highest image quality.

False Chroma Suppression

The final stage in the image-processing pipeline corrects various color artifacts. For example, when a picture of foliage is taken against a bright sky, purple fringes can appear on high-contrast edges. Typically, tuning is not required for the majority of applications under normal lighting condition. If low-lighting conditions exist and the camera will be capturing very low-quality images, the resulting color artifacts on textures or object edges can be removed using a strong false chroma suppression parameter.

Quality as Defined by Preference

While there are aspects of image quality that are clearly objective – a washed-out picture with all the wrong colors is obviously not a true representation of what a user sees – fine tuning image quality is still a highly subjective process and perhaps more art than science since so much is predicated upon user preference. A bright image does not always mean it is of higher quality than a dark image, nor is a sharp image valued more than a smooth one. For example, portraits tend to be more appealing when they are smoother while landscape images usually benefit from being sharper.

Ultimately, the final determinate of image quality is how satisfied the end user is with the results. Quality depends on the scene being shot, the particular lighting conditions and the particular preferences of the user. How the image-processing pipeline is implemented has a significant impact on perceived final image quality, and consequently, a digital camera must have the flexibility to control and adjust to individual requirements.

Such flexibility, however, cannot come at the expense of performance or highly complex algorithms that are too difficult to work with. By offering developers a configurable image-processing pipeline, DaVinci™ technology digital media processors balance the ability to adjust quality without limiting the usefulness of algorithms. Additionally, implementing the pipeline in hardware avoids burdening the main application processor, freeing these cycles for more value-added functionality such as red-eye removal, compensation for slight movements or shaking, or advanced video analytics, to name a few.

TI's digital media processors were designed to be flexible enough to accommodate changing user preferences in a dynamic fashion. Developers can provide multiple operating modes to enable users to select different pipeline configurations based on their own individual preferences. With TI processors, developers have the ability to disable various pipeline stages in order to address the specific needs of specialized applications. For example, consider an industrial application, which must be able to address specific noise issues. Rather than recreating the entire image-processing pipeline, developers are able to simply disable the existing noise filter and add a specialized or proprietary noise filter of their own.

The flexibility of the image-processing pipeline also results in significant time-to-market and cost-reduction savings since the same pipeline can be leveraged across multiple applications by accommodating differences in individual sensor and lens combinations, output displays and varying lighting conditions. In this way, developers can also reduce the number of components that need to be managed across product lines and lower overall bill of materials (BOM) through economies of scale.

TI's DaVinci™ technology digital media processors maximize control of quality and minimize complexity while accommodating diverse user preferences. The integrated, hardware-based image-processing pipeline provides the performance as well as the flexibility required to ensure that developers can produce the highest quality images available for their applications.