1.1 (Invited)

Image Processing Chain in Digital Still Cameras

Albert J. P. THEUWISSEN

DALSA Corp.
Prof. Holstlaan, 4,
5656AA Eindhoven, The Netherlands
albert.theuwissen@dalsa.com

Abstract

Nobody is perfect, neither are the components used in digital still cameras. Nevertheless the photographer wants to see the best quality image taken with his/her Digital Still Camera (DSC). Today the number crunching power of signal processors is high enough that many of the imperfections created by the lens and the image sensor can be easily corrected. For instance lens vignetting, defect pixels and color imperfections are no longer limitations to obtain high-quality results. The Digital Signal Processor (DSP) corrects everything, even the good pixels ....

Keywords : DSC, Image Sensors, Adaptive Image Processing, CCD, CMOS Image Sensor.

Introduction

In the past the quality of a picture taken with a DSC was determined to a large extent by the quality of the lens and the image sensor. But progress in digital signal processing power and in the understanding of the physics behind the various defects and limitations of the components that make up the DSC, allow many defects to be corrected. Ultimately it remains true : "To make a good image, one has to take a good image". But along the road from photons to digital numbers the signal can pass through several calculation cycles to improve the quality of the end result.

This paper describes the main artifacts that can be introduced by the lens and the image sensor, and shows how they can be corrected.

Artifacts degrading the image quality ......

There are several kinds of artifacts that degrade the quality of the image captured by the camera. The first are real imperfections added to the image, e.g. those caused by the lens, by the color filter and by the sensor. Other artifacts are introduced by the physical nature or working principle of the image sensor, e.g. by the temporal and spatial integrations of the image signal, and by the spectral distributions of the color filters.

The most important artifacts are first described, and then methods for correction are highlighted.

Artifacts of the lens

The main problem with the cheaper lenses used in consumer DSC and mobile phone cameras is the vignetting. The transmission of the lens, as well as its resolving power is highest in the center part of lens. Towards the edges these parameters become worse. The transmission fall-off follows a $\cos^4$-law : where the argument of the cosine is the angle between the optical axis of the lens and the point in the image plane. The lower the quality of the lens is (read the cheaper the lens), the stronger the vignetting effect is. A 25% reduction in transmission at the edge of the image plane relative to the center is not exceptional !

Artifacts of the illumination

The human eye does not perform an absolute measurement of the incoming light information, while an image sensor does. Moreover, the human eye interprets several combinations of Red (R), Green (G) and Blue (B) light as white. (This effect is also known as metamerism.) This is not the case for imagers !

The effect of balance in R, G and B is also amplified by the fact that the sensor does not have equal quantum efficiencies for the various spectral bands.

Artifacts with the color sampling

The absorption coefficient of silicon is very strongly dependant on the wavelength of the incident light. In principle one can use this characteristic to separate the incoming photons into the three basic color planes (R, G and B). But in practice the color separation of silicon itself is not selective enough to allow color imaging even in low-light level conditions. For this reason color separation is typically done by means of color filters on top of every pixel of the imager. A typical configuration is the Bayer pattern, in which every 2x2 matrix of pixels has 1 R, 2 G and 1 B "colored" pixels. These filters introduce several artifacts [1].

- Transmission curves of the color filters : The color filters themselves are made out of dyed photoresist or evaporated color pigments. The pass band of the color spectra overlap each other and the filters do not represent pure colors. The blue pixels will respond slightly to green light as do the red pixels, etc.

- Sampling grid : Every pixel of the image sensor samples only 1 single color, while on the other hand every pixel in a hard copy is composed of the combination of the 3 colors In other words during the sampling of the information by the image sensor, the information about 2 of the color components at every location is missing !

Artifacts of the Image Sensor

Most of the sensor artifacts arise from noise sources and pixel non-uniformities.

- Temporal and Fixed-Pattern Noise (FPN) : DSC applications are very much more demanding than the classical video applications as far as noise generated in the imaging chain is concerned. With a standard video application 30, 50 or 60 frames/s are generated and displayed. Most displays as well as the human eye are not capable of processing images at 30 frames/s or greater. That results in the fact that the displays and the human eye filter out most of the temporal noise in video images. That is not the case when only a single image is presented to a display or to the human brain. In still images the temporal noise of the imaging chain is frozen and all temporal noise becomes fixed-pattern noise !

But besides the temporal noise becoming fixed-pattern noise in still applications, the sensors themselves also generate extra fixed-pattern noise. For CCDs most of the FPN is random and originates from dark-current non-uniformities. For CMOS imagers the dark current is higher than for CCDs and so the contribution of dark current to FPN is also higher. But the architecture used to readout a CMOS image sensor generates extra components of FPN, some of it random and some of it column-wise.

- Sick and Dead pixels : In a real image sensor it can happen that a certain amount of the pixels turn out to be sick : they show a light sensitivity that it less then the average value of all pixels. The device can even show pixels that are not sensitive at all, so-called dead pixels. And in the worst case a complete column of pixels can even be dead !

.... but that can be repaired as well.

Most the artifacts described above can be repaired or counteracted in the processing of the data during and after the readout of the sensor. A basic fundamental understanding of the various artifacts, (their nature and their dependencies) helps in creating appropriate algorithms to "clean up" the images. Unfortunately every time the sensor data is manipulated, the risk exists of decreasing the signal-to-noise ratio of the signal. Especially when correction requires that signals be differenced, a loss in signal-to-noise ratio needs to be accepted.

In most cases the correction of artifacts is done in the reverse order to which they were introduced in the sensor signal. To cope with the aforementioned artifacts one can use the following strategies :

Correction of column FPN

Especially in the case of CMOS image sensors, column FPN can be worse. If appropriate electronic processing in the analog domain on-chip is insufficient to suppress this FPN, then the column FPN can be "measured" by reading one or more dummy lines from the sensor and the values are stored in digital memory. Column FPN finds its origin in threshold variations of MOS transistors resulting in gain and off-set variations between the various column amplifier, and since this FPN is relatively constant over time, the stored digital values can be used to correct real image data.

Pixel level FPN

FPN can have two origins : spatial variations in electrical parameters (threshold variations of the in-pixel amplifiers), and spatial variations in dark current [2]. Both can be corrected by subtracting a reference frame stored in digital memory, however the dark current contributions are very sensitive to temperature variations and hence the reference frame will be temperature dependant. The advantage of subtracting a reference frame is the automatic cancellation of the column FPN as well.

Sick and defect pixels

The only way out to correct for defect pixels is an appropriate algorithm in which the defect pixels (with a known XY location) are interpolated with the knowledge of the pixel values of its neighbors in the same color plane [3].

If sick pixels are a problem, correction can be done by means of look-up tables. The XY addresses of the sick pixels are stored in a digital memory, as well as their gain factor needed to compensate for their "sickness".

Dark current correction

This is probably the simplest artifact to correct, although the dark current is very strongly temperature dependent. Correction is done by using the dark current content of dark reference pixels to define the black level of the video signal. Every sensor has black reference columns and rows of pixels. These are exactly the same as the active pixels and integrate during the same period as the active pixels. The only difference is that a metal light shield shields the reference pixels from light. In this way the reference pixels contain only dark current.

The dark reference pixels are read out together with the actual video in each frame. In this way the temperature dependency of the image pixels can be compensated : the dark reference pixels always have the same temperature as the active pixels [4].

Although correction for dark current itself is relatively simple, correction for dark current shot noise is not possible !

Color interpolation

Every pixel samples just one color – the two remaining ones needed to build up a color image for that pixel are not sampled. The two missing color components need to be reconstructed by interpolation.

These interpolation algorithms can be of a very sophisticated complexity [5]. In many cases adaptive interpolation schemes are used. That means that the interpolation algorithm for a particular pixel can change from image to image, and depends on the actual content of the image. For every unique image and every pixel in that image, the image information in the vicinity of the pixel under consideration is analyzed. Depending on the outcome, a particular algorithm is chosen. In many cases the interpolation is done in a direction (vertical, horizontal, left diagonal, right diagonal) that contains the least change in image content. For example, if the image contains (locally) vertical lines, the interpolation is done preferably along the vertical direction.

Color matrixing

To cope with imperfections in the transmission characteristics of color filters, with the optical cross-talk between pixels, and with the spectral sensitivity differences between pixels, the color obtained after the interpolation needs to be "recalibrated". This can be done by means of multiplying the original color vector (R_o, G_o and B_o) for every pixel with a 3 x 3 color matrix to obtain the new values of the three base colors (R_n, G_n and B_n) [6].

This manipulation of the data can introduce significant loss in signal-to-noise. Especially where there is large overlap in the spectral transmission curves of the filters, or where the pixels suffer from cross-talk, the off-diagonal matrix coefficients become negative and large. And differencing large values generates results with lower S/N ratio.

Contouring

The sparse sampling grids for the 3 colors in combination with the interpolation have a kind of low-pass filter effect on the image data. The resulting image is spatially smoothed. High-frequency components are filtered out, and the picture lacks some sharpness. But a correction for this affect can also be applied. Specifically, the data after color matrixing (R_n, G_n and B_n) is passed through a high-pass filter. The result after high-pass filtering (R_{HF}, G_{HF} and B_{HF}) is added to the original data (R_n, G_n and B_n). In this way the contours of the various objects in the image are amplified and part of the lost sharpness is regenerated.

White Balancing

Every color camera that makes use of solid-state image sensors needs to be "told" how to combine R, G and B to generate white. Most of the time only the R and the B channels are adapted – the G channel is used as a reference. For example, for illumination by a clear blue sky, more B is used to build up white, while in the case of illumination by candle light, R contributes more strongly to white.

To create the R and B coefficients needed for correct white balancing, several methods are used :

- by means of a number of pre-settings on the camera,
- by pointing the camera to a white object and forcing the camera to accept this as the correct combination to generate the white color,
- by an automatic measurement system in which the camera tries to find the optimum combination of white by searching for "gray" objects in the scene.

Vignetting

The intensity fall-off of the lens can be easily compensated for. In a fixed-focus, fixed-iris lens the vignetting is fixed and can be pre-programmed in the camera. All R-, G- and B-values are multiplied by a constant value (which becomes gradually larger than 1 towards the outer edges of the image) to regenerate the original intensities.

If the lens also suffers sharpness fall-off towards the edges, a position-dependant sharpening algorithm can also be used to create extra contouring towards the outer parts of the image. Correction of vignetting effects becomes more difficult if the camera uses a lens that is not fixed focus, fixed-iris. Correction for zoom-lenses is particularly complicated. Nevertheless, there are cameras on the market that use the setting of the iris and the setting of the focal length of the lens to adapt the correction algorithms. Once the effort is made to include dedicated correction algorithms in a DSP, it is a small step to make these corrections adaptive to the camera settings.

Conclusion

Several imperfections in the imaging chain as well as in the basic integrating nature of solid-state image sensor generate raw image data that needs many corrections and improvements before it can be displayed in hard copy. Many of the correction algorithms rely strongly on the basic generation mechanisms of the artifacts that need to be corrected. In recent years many of these tools have been applied in an adaptive way : the use of one or another algorithm depends not only on the exact camera/lens setting, but also on the actual content of the image. As a consequence, the algorithms can change from pixel to pixel within one image, and almost certainly from image to image.

Acknowledgement

The author thanks his colleagues from DALSA Tokyo KK, DALSA Professional Imaging Eindhoven and DALSA Waterloo for their discussions concerning this subject. The contributions of Auke van der Heide (DPI) in particular are gratefully acknowledged.

References

[1] A. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*, Kluwer Academic, 1995.

[2] H-S. Wong, "Technology and Device Scaling Considerations for CMOS Imagers", *IEEE Transactions on Electron Devices,* Vol. 43, Dec. 1996, pp. 2131-2142.

[3] A. Tanbakuchi etal., "Adaptive Pixel Defect Correction", *Proceedings of SPIE,* Vol. 5017, Santa Clara, 2003, pp. 360-370.

[4] A. Theuwissen etal., "Ultra-high Resolution Image Capturing and Processing for Digital Cinematography", *Digest Technical Papers ISSCC,* Digest of Technical Papers, ISSCC, 2003, pp. 162-163.

[5] T. Kuno etal., "Aliasing Reduction Method for Color Digital Still Cameras with a Single-Chip Charge-Coupled Device", *Journal of Electronic Imaging,* Vol. 8, Oct. 1999, pp. 457-466.

[6] A.J. Blanksby and M.J. Loinaz, "Performance Analysis of a Color CMOS Photogate Image Sensor", *IEEE Transactions on Electron Devices,* Vol. 47, Jan. 2000, pp. 55-64.