Digital camera pipeline resolution analysis

Toadere Florin
INCDTIM Cluj Napoca
Str. Donath nr. 65-103, ClujNapoca
Romania
toadereflorin@yahoo.com

Abstract: - our goal of this paper is to make a resolution analysis of an image that passes trough a digital camera pipeline. The analysis covers the lens resolution, pixel dimension, light exposure and colors processing. We pass an image trough a photographic objective; we focus the light on a Bayer color filter array then we interpolate, we sharp and we set the integration time. The color processing analysis covers: the colors balancing, the colors correction, the gamma correction and the conversion to XYZ.

Key-Words: - photographic objective, pixel size, Bayer CFA, dynamic range, colors processing

1 Introduction

Digital image is an electronic snapshot taken of a scene. The digital image is sampled and mapped as a grid of dots or pixels. A pixel is the smallest piece of information in an image. The term pixel is also used for the image sensor elements on a digital camera. Also cameras are rated in terms of "megapixels". In terms of digital images, spatial resolution refers to the number of pixels utilized in construction of the image. The spatial resolution of a digital image is related to the spatial density of the image and optical resolution of the photographic objective used to capture the image. The number of pixels contained in a digital image and the distance between each pixel is known as the sampling interval, which is a function of the accuracy of the digitizing device. An image may also be resampled or compressed to change the number of pixels and therefore the size or resolution of the image [19].

In Fig.1 we have an illustration of how the same image might appear at different pixel resolutions, if the pixels were poorly rendered as sharp squares [19].
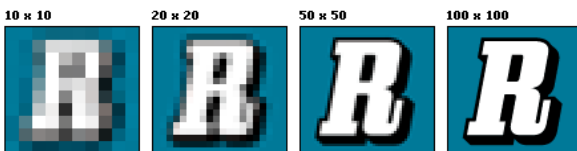

Fig.1 image at different pixel resolutions

The optical resolution is a measure of the photographic objective's ability to resolve the details present in the scene. The measure of how closely lines can be resolved in image is called spatial resolution, and it depends on properties of the system creating the image. For practical purposes the clarity of the image is decided by its spatial resolution, not the number of pixels in an image. The spatial resolution of computer monitors is generally 72 to 100 lines per inch, corresponding to pixel resolutions of 72 to 100 ppi. Optical resolution is sometimes used to distinguish spatial resolution from the number of pixels per inch. In optics spatial resolution is express as contrast or MTF (modulation transfer function). Smaller pixels result in wider MTF curves and thus better detection of higher frequency energy [1-4, 19].

An optical system is typically judged by its resolving power, or ability to reproduce fine detail in an image. One criterion, attributed to Lord Rayleigh, is that two image points are resolved if the central diffraction maximum of one is no closer than the first diffraction zero of the other. Rayleigh's criterion applied to point images demands that the images be separated by 0.61distance between centers of the dots [1, 3].
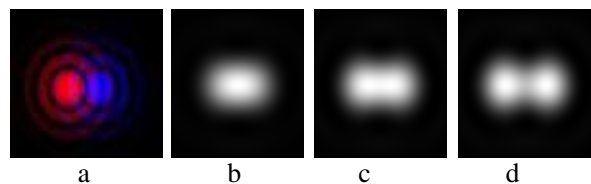

Fig.2 a) diffraction figure, b) 0.4 separation, c) 0.5 separation, d) 0.6 separation

In Figure 2, we see the diffraction figure of the two points in conformity with Rayleigh criterion and three possible situations, corresponding to separations of 0.4, 0.5, and 0.6. The first case is not resolved because there is no evidence of two peaks.

The second case is barely resolved, and the third case is adequately resolved.

The image sensor is a spatial as well as temporal sampling device. The sampling theorem sets the limits for the reproducibility in space and time of the input image. Signals with spatial frequencies higher than Nyquist frequencies cannot be faithfully reproduced, and cause aliasing. The photocurrent is integrated over the photodetector area and in time before sampling. Nowadays technologies allow us to set the integration time electrically and not manually like in classic film camera. In photography, exposure is the total amount of light allowed to fall on the photographic image sensor during the process of taking a photograph. Factors that affect the total exposure of a photograph include the scene luminance, the aperture size, and the exposure time [17, 19].

2 The image capture system

The MTF and PSF (point spread function) are the most important integrative criterion of imaging evaluation for optical system. The definition of the PSF and the MTF can be found in [1-4]. We don't enter in details of how we compute the all optical system PSF's, we just present the definition of the circular aperture, light fall of and CCD MTF. Further information can be found in [1-4]. Knowing those information we go further having for the optical part:

$$PSF_{opt} = PSF_{lens} * PSF_{filter} * PSF_{\cos 4f} * PSF_{CCD} \quad (1)$$

2.1 The photographic objective design

In this paper we use an Inverse telephoto lens Fig.3 whose functionality is simulated using Zemax [11].
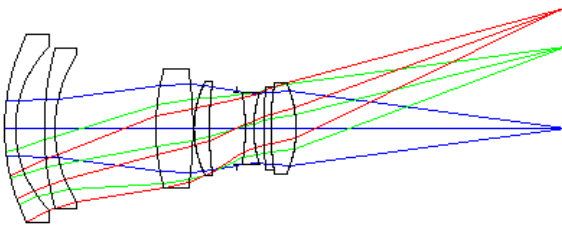

Fig.3 The inverse telephoto lens

The visual band optical system has $60°$ FOV, f number 2.5. 1/2 inch CCD with C optical interface is selected, i.e. its back working distance is 17.52 ± 0.18 mm. Example tracings for rays are plotted at 0, 22 and 30 degrees and the wavelengths for all calculations are 0.450, 0.550 and 0.650 microns.

This is the lens construction which gives a short focal length with a long back focus or lens CCD distance. It enables wide-angle lenses to be produced for small format cameras, where space is required for mirrors, shutters, etc. To understand how this photographical objective functions, we compute the PSF with the Zemax optical design program [2].

2.2 The photographic objective aperture

The aperture is the hole through which light enters the camera body, and the word most often refers to the photographic lenses is f/number, which is an indication of how much light will reach the sensor. The f/number is equivalent to the size of the hole divided by the focal length of the lens and is present to any photographical objective.

The circular aperture has the formula:

$$c(r) = circ\left(\frac{r}{r_0}\right) \quad (2)$$

r is the circle radius;

r_0 is the cut off radius;

and the PSF is calculated as:

$$c(x, y) = \lambda \frac{r_0}{r} J_1\left(\lambda \frac{r_0}{r}\right) \quad (3)$$

λ is the wavelength;

J_1 is the Bessel function of order one.

A perfect optical system is diffracted limited by relation:

$$d = 2.44 \lambda N \quad (4)$$

were:

$$N = \frac{F}{\#} = \frac{f}{d} = \frac{1}{2NA}$$

is the focalization ratio and it is present on any photographic objective.

F is the focus length;

d is the aperture diameter;

NA is the numerical aperture;

N is $\sqrt{2}$ multiple: 1.4, 2, 2.8, 4, 5.6……

The constant 2.44 is used because it corresponds to the first zeroes of the Bessel function J1(r) for a circular aperture [1-4, 9].

2.3 The light fall off

Cos4f law states that light fall-off in peripheral areas of the image increases as the angle of view increases, even if the lens is completely free of vignette. The peripheral image is formed by groups of light rays entering the lens at a certain angle with

respect to the optical axis, and the amount of light fall-off is proportional to the cosine of that angle raised to the fourth power. As this is a law of physics, it cannot be avoided [2, 9, 17].

$$E_i = \pi \cdot L \frac{1}{1 + 4(f/\#(1-m))^2} (\cos\phi)^4 \qquad (5)$$

L is source light radiation, m is the magnification.

2.4 The CCD MTF

Our interest is to see what happens to an image that passes through the optical part of a CCD image sensor. We start by computing MTF and PSF. We consider a 1-D doubly infinite image sensor (Fig. 4. a)) where: L quasi neutral region, Ld depletion depth, w aperture length, p pixel size [5-8, 16, 17].
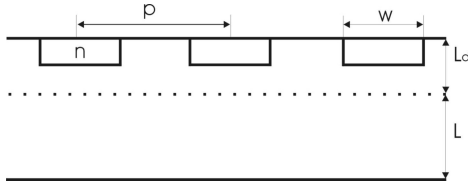

Fig.4 a) the CCD sensor model

To model the sensor response as a linear space invariant system, we assume $n+/p$-sub photodiode with very shallow junction depth, and therefore we can neglect generation in the isolated $n+$ regions and only consider generation in the depletion and p-type quasi-neutral regions. We assume a uniform depletion region. The monochromatic input photon flux $F(x)$ to the pixel current $iph(x)$ can be represented by the linear space invariant system. $iph(x)$ is sampled at regular intervals p to get the pixel photocurrents. After certain manipulation [5, 6] we have:

$$MTF(f) = \frac{|H(f)|}{H(0)} = \frac{D(f)}{D(0)} w^2 \sin c(wf) \qquad (6)$$

$\frac{D(f)}{D(0)}$ is called the diffusion MTF and sinc(wf) is called the geometric MTF. We also have:

$$MTF_{CCD} = MTF_{diffusion} \cdot MTF_{geometric} . \qquad (7)$$

Note that $D(0) = n(\lambda)$ with $n(\lambda)$ the spectral response of the CCD. By definition: spectral response is a fraction of photon flux that contributes to photocurrents as a function of wave length. Thus $D(f)$ can be viewed as a generalized spectral response (function of spatial frequency as well as wavelength). In our analyses we use 2D signals (images) and we shall generalize 1D case to 2D case. We know that we have square aperture at each photodiode with length w.

$$MTF = \frac{|H(f_x, f_y)|}{H(0)}$$
$$= \frac{D(f_x, f_y)}{D(0)} w^2 \sin c(wf_x) \sin c(wf_y) \qquad (8)$$

2.5 The pixel dimensions

To find the maximum size of a pixel in the CCD image sensor we use Equation (4). The sensor is located in focal plane of the lenses, the wavelength $\lambda = 555nm$ and the magnifying coefficient $M = 1$. Applying these values to Equation (4) we obtain:

$$d = 2.44 \cdot 555 \cdot 11 = 10.833 \mu m ;$$
$$d_1 = d \cdot M = 10.833 \mu m .$$

To deliver sufficient sampling the pixel size should be smaller then:

$$p = \frac{d_1}{2} = \frac{10.833}{2} = 5.4 \mu m .$$

In our analyses we use the next parameter values:

$p = 5.4 \mu m$, $Ld = 1.8 \mu m$, $L = 10 \mu m$, $w = 4 \mu m$
$\lambda = 550nm$ and $y = 2.3mm$.

1/2 inch CCD with C optical interface is selected, i.e. its back working distance is 17.52 ± 0.18 mm. The visual band optical system has $60°$ FOV, f/number 2.8 [5, 6, 9].

According to the relation between the FOV of object space and image height shown in Equation (9), if FOV and the size of CCD are selected, the effective focal length is determined.

$$2y = 2f \tan|\omega| \qquad (9)$$

$2y$ is the diagonal size of CCD, f is the effective focal length and 2α is the full field of view in object space.

Taking the effective focal length f in mm and the CCD pixel size p in microns, we can calculate the CCD plate scale as

$$P = \frac{206265 \times p}{1000 \times f} \qquad (10)$$

where 206265 is the number of arcseconds in 1 radian and 1000 is the conversion factor between millimeters and microns [7].

2.6 The Bayer CFA

Color imaging with a single detector requires the use of a Color Filter Array (CFA) [12-17] which covers the detector array. In this arrangement each pixel in the detector samples the intensity of just one of the

many-color separations. In a single detector camera, varying intensities of light are measured at a rectangular grid of image sensors. To construct a color image, a CFA must be placed between the lens and the sensors. A CFA typically has one color filter element for each sensor configuration. Many different CFA configurations have been proposed. One of the most popular is the Bayer pattern, which uses the three additive primary colors, red, green and blue (RGB), for the filter elements. Green pixels covers 50% of the sensor surface and the others colors red and blues covers 25% each.
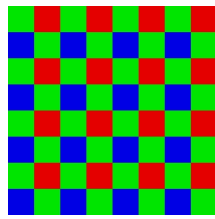

Fig.5 CFA Bayer RGB

2.7 The color difference space interpolation

The color difference space method proposed by Yuk, Au, Li, and Lam [18] interpolates pixels in green-red and green-blue color difference spaces as opposed to interpolating on the original red, green, and blue channels. The underlying assumption is that due to the correlation between color channels, taking the difference between two channels yields a color difference channel with less contrast and edges that are less sharp. Demosaicking an image with less contrast yields fewer glaring errors, as sharp edges, what cause most of the interpolation errors in reconstructing an image. The color difference space method creates KR (green minus red) and KB (green minus blue) difference channels and interpolates them; the method then reconstructs the red, green, and blue channels for a fully demosaicked image. Further information about the method can be found in [15, 20].

2.8 The sharpening

Sharpening is often performed immediately after color processing or it can be performed at an earlier stage of the image processing chain; for example, as part of the CFA de-mosaicing processing [4]. In this paper we sharp right before interpolation in order to eliminate the blur caused by the optical system components and to have a better view of the image transformation process.

Sharpness describes the clarity of detail in a photo. In order to correct the blur we have to sharp the image using a Laplacian filter [12-15]:

$$L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}. \qquad (11)$$

2.9 The CCD dynamic range

The dynamic range of a real-world scene can be 100000:1. Digital cameras are incapable of capturing the entire dynamic ranges of scenes, and monitors are unable to accurately display what the human eye can see. Sensor DR (dynamic range) quantifies its ability to image scenes with wide spatial variations in illumination. It is defined as the ratio of a pixel's largest nonsaturating photocurrent i_{max} to its smallest detectable photocurrent i_{min} [4-8, 13]. The largest saturating photocurrent is determined by the well capacity and integration time

$$i_{max} = \frac{qQ_{max}}{t_{int}} - i_{dc} \qquad (12)$$

The smallest detectable signal is set by the root mean square of the noise under dark conditions. DR can be expressed as:

$$DR = 20\log_{10}\frac{i_{max}}{i_{min}}$$

$$= 20\log_{10}\frac{qQ_{max} - i_{dc}t_{int}}{\sqrt{qt_{int}i_{dc} + q(\sigma_{read}^2 + \sigma_{DNSU}^2)}} \qquad (13)$$

$q = 1.6 \times 10^{-19}C$ is the electron charge, Q_{max} is the effective well capacity; σ_{read}^2 is the readout circuit noise and σ_{DNSU}^2 is the offset FPN due to dark current variation, commonly referred to as $DSNU$ (dark signal nonuniformity).

3 The color processing

3.1 The color balance

In photography color balance refers to the adjustment of the relative amounts of red, green, and blue primary colors in an image such that neutral colors are reproduced correctly. Color balance changes the overall mixture of colors in an image and is used for generalized color correction. The Von Kries method is to apply a gain to each of the human cone cell spectral sensitivity responses so as to keep the adapted appearance of the reference white constant [12-15]. The Von Kries method for white balancing can be express as a diagonal matrix. The elements of the diagonal matrix D are the ratios

of the cone responses (Long, Medium, Short) for the illuminant's white point. In our simulation we consider the monitor with point.

$$D = \begin{bmatrix} 0.5844 & 0 & 0 \\ 0 & 0.5753 & 0 \\ 0 & 0 & 0.514 \end{bmatrix} \qquad (14)$$

3.2 The color correction

We need to specify two aspects of the display characteristics therefore we can specify how the displayed image affects the cone photoreceptors [12, 15]. To make this estimate we need to know something about: (1) the effect each display primary has on your cones and (2) the relationship between the frame-buffer values and the intensity of the display primaries (gamma correction). To compute the effect of the display primaries on the cones, we need to know the spectral power distribution (SPD) of the display; an Apple monitor (Fig. 6 b)), and the relative absorptions of the human cones (Fig. 6 a)).
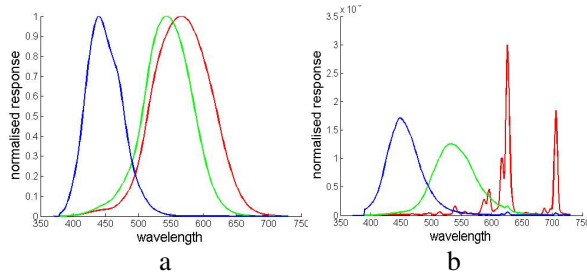
Fig.6 a) Cone sensitivity, b) Samples monitor SPD

Having this data, we can compute the 3 x 3 transformation that maps the linear intensity of the display R, G, B signals into the cone absorptions L, M, S.

3.3 The gamma correction

Phosphors of monitors do not react linearly with the intensity of the electron beam. Instead, the input value is effectively raised to an exponent called gamma. Gamma is the exponent on the input to the monitor that distorts it to make it darker. Since the input is normalized to be between 0 and 1, a positive exponent will make the output lower. The NTSC standard specifies a gamma of 2.2. By definition [12-15] gamma is a nonlinear operation used to code and decode luminance or tristimulus values in video or image systems. Gamma correction is, in the simplest cases, defined by the following power law expression:

$$V_{out} = V_{in}^{\ \gamma} . \qquad (15)$$

3.4 The color conversion

We convert the device-dependent RGB data into LMS (XYZ) format [12-15] using the color calibration information specified in color correction paragraph. RGB represent a color space. Red, green and blue can be considered as the X, Y and Z axes using equation (12) we can convert one into other.

$$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \begin{bmatrix} 0.431 & 0.342 & 0.178 \\ 0.222 & 0.707 & 0.071 \\ 0.229 & 0.130 & 0.939 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix} \qquad (16)$$

4 The simulation results

In this simulation we shall try to demonstrate in images the functionality of an image capture system from the resolution and luminosity point of view and then to process the colors [9, 10, 16, 17]. From image (Fig.7 a)) to image (Fig.8 b)) we see the role played by the optical part of the sensors. Even if we do not have deformation of the images, we have diffraction and changes in contrast, becoming worse as the image passes through sensors. Digital camera objective suffers from geometrical distortions and also in the CCD exists electrical and analog to digital noises which are not taken in to account here. In Fig. 8 c) we have a Bayer CFA sub sampled image. By using a good interpolation technique we can minimize the pixel artifacts (Fig.9 a)). We sharpen (Fig.9 b)) and we set the dynamic range (Fig.9 c)). By setting the integration time we determine the amount of light that enter in the digital camera. Then we need to recover the original color characteristics of the scene by color balancing. In (Fig.10 a)) we use the Von Kries matrix, a simple and accurate color balancing method. Another very important role is played by the color correction performing compatibility between human eyes cone sensitivity and the SPD sample monitor as in Fig. 10 b). Comparing (Fig.6 a)) and (Fig.6 b)) we see that in red colors spectrum we have big differences. Thus we expect to have some deficiencies to recover this color and, in some way, any other colors. Because the intensity of the light generated by a display device is not linear, we must correct it by gamma correction. In this analysis gamma is 0.45 and finally we have conversion to CIE XYZ as in Fig. 10 c).

All the images in this paper have the dimension of 256x256 pixels. The images are generated one from another following the order presented in this paper. The time necessarily to generate in Matlab all images is about 5 seconds.
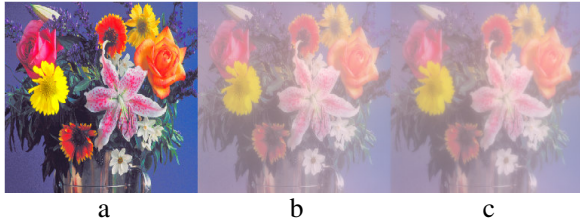
Fig.7 a) input image, b) image at the output of the lens, c) image at the aperture output
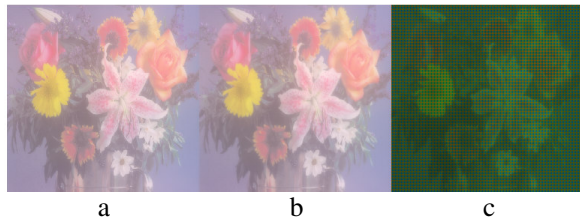

Fig.8 a) light fall off, b) CCD MTF, c) Bayer CFA

Fig.9 a) interpolation, b) sharp, c) dynamic range

Fig.10 a) color balance, b) color correction, c) conversion to XYZ and gamma correction

5 Conclusion

The analysis and simulation presented in this paper covers an important part of a digital camera pipeline related to the image acquisition system and color processing. This analysis can be useful in understanding and learning the functionality of the digital camera pipeline and to help people who design and implement such cameras. Further work is needed to simulate missing parts like: electrical noises and inverse problem reconstruction.

Acknowledgment

References:
[1] A. Vandrlugt, *Optical signal processing*, Wiley, 1991
[2] Joseph M. Geary, *Introduction to lens design with practical Zemax example*, William Bell, 2002
[3] J.W. Goodman, *Introduction to Fourier optics*, McGraw Hill, New York 1996
[4] T. C. Poon and P. Banerjee, *Contemporary Optical Image Processing with Matlab*, Elsevier, 2001
[5] www.isl.stanford.edu/~abbas/
[6] www.optics.arizona.edu/detlab/
[7] S. B. Howell, *Handbook of CCD astronomy*, Cambridge University Press, 2006
[8] G. Lutz, *Semiconductor radiation detectors*, Springer Verlag, 2007
[9] Toadere Florin, Simulation the optical part of an image capture system, *ATOM-N 2008 The 4th of the international conference advanced topics in optoelectronics, microelectronics and nanotechnologies*,29-31 August, Constanta, Romania.
[10] P. Maeda, P. Catrysse, B. Wandell: "Integrating lens with digital camera pipeline", *In Proceedings of the SPIE Electronic Imaging 2005 Conference*, Santa Clara, CA, January 2005
[11] www.zemax.com
[12] M. Ebner, *Color constancy*, Wiley & Sons, 2007
[13] Gaurav Sharma, *Digital Color Imaging Handbook*, CRC Press, 2003
[14] K. Castleman, *Digital image processing*, Prentice Hall, 1996
[15] www.white.stanford.edu/wandell.html
[16] Toadere Florin, Functional parameters enhancement in a digital camera pipeline image simulation, *NSIP 2007, International workshop on nonlinear signals and image processing,* September 10-12, Bucharest, Romania.
[17] C. Ting: Digital camera system simulator and applications, PhD thesis, Stanford University, CA, 2003.
[18] Yuk, C.K.M.; Au, O.C.; Li, R.Y.M., Color Demosaicking Using Direction Similarity in Color Difference Spaces, *Sui-Yuk Lam Circuits and Systems, 2007. ISCAS 2007.*
[19] http://en.wikipedia.org/wiki/
[20] http://scien.stanford.edu/class/psych221/projects/08/Demosaicing/methods.pdf