Resolution analysis of an image acquisition system

TOADERE FLORIN, RADU ARSINTE, NIKOS E. MASTORAKIS

UT Cluj Napoca
Str. Memorandumului nr.28
INCDTIM Cluj Napoca
Str. Donath, nr. 65-103,
Cluj Napoca, ROMANIA
Florin.Toadere@bel.utcluj.ro

UT Cluj Napoca
Str. Memorandumului nr.28
Cluj Napoca, ROMANIA
radu.arsinte@com.utcluj.ro

WSEAS Research Department
Agiou Ioannou Theologou 17-23
Zografou, 15773, Athens, GREECE
mastor@wseas.org

Abstract: - the aim of this paper is to make the resolution analysis of an image that propagates trough the image acquisition system of a digital camera pipeline. The analysis covers the lens resolution, the pixel dimension, the light exposure, the colors processing blocks and the JPG compression. The input image passes 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, the conversion to XYZ and the JPG compression.

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

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 [23].

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 [23].

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 [1-3, 8, 15].

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 0.61 distance between centers of the dots.

![Image](image.jpg)

Fig. 2: a) The 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. We have 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.

In figure 3 we have the configuration used in our analysis: the object, the optical sensor, the image sensor and color processing blocks.

![Image](image.jpg)

Fig. 3: The image capture system block scheme

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 [2, 7, 13, 14, 16, 18].

2 The image capture system

The image acquisition sensors are complex systems using optical, mechanical and electrical components. In the process of image acquisition we have: an input image, a photographic objective and a CCD image sensor. The MTF and PSF (point spread function) are the most important integrative criterion of imaging evaluation for optical system. In our analyses we use a LSI (linear shift invariant system) which is characterized by the mathematical relation

\[ g(x, y) = H[f(x, y)] \] (1)

where \( H \) is an operator representing a linear, position (or space) invariant system which is characterized by relation

\[ g(x, y) = \int \int \phi(a, b)h(x-a, y-b) \] (2)

\( \alpha \) and \( \beta \) are spatial frequencies (line/mm) which are defined as the rate of repetition of a particular pattern in unit distance.

\[ h(x-a, y-b) = H[\delta(x-a, y-b)] \] (3)

is the impulse response of \( H \); in optics, it is called the point spread function (PSF) [1-4, 6, 8, 15]. The net sensor PSF is a convolution of individual response from optical lenses: the lenses, the aperture, the light falloff, the CCD’s optical part

\[ PSF = PSF_{\text{lens}} \cdot PSF_{\text{filter}} \cdot PSF_{\text{cos}} \cdot PSF_{\text{CCD}} \] (4)

The PSF characterize the image analyses in space but also we can characterize the image in frequency by optical transfer function (OTF). OTF is the normalized autocorrelation of the transfer function and has the formula

\[ H(\alpha, \beta) = \frac{\int \int P(x+\alpha/2, y+\beta/2)P(x-\alpha/2, y-\beta/2)\, dx\, dy}{\int \int P(x, y)^2\, dx\, dy} \]

The numerator represents the area of overlap of two pupil functions, one of which is displaced by \( \alpha/2, \beta/2 \) and the other in opposite direction by \( -\alpha/2, -\beta/2 \) on x and y; and the other in opposite direction by \( -x \) and \( -y \). OTF is defined as the rapport between the area of the overlap of displace pupil function and complete area of the pupil function.
The changes in contrast that happen when an image passes through an optical system is expected to have a lot to do with the optical transfer function. The definition of the modulation transfer function (MTF) is

\[ MTF = \frac{\text{contrast of output image}}{\text{contrast of input image}} \]

which represent the ratio of the contrast of the output image to that of the input image. The relation between OTF and MTF is

\[ MTF = |OTF| \tag{5} \]

The modulation transfer function is identical to the absolute value of the optical transfer function. The net sensor MTF is a multiplication of individual absolute transfer function

\[ MTF = MTF_{\text{lens}} \cdot MTF_{\text{filter}} \cdot MTF_{\text{cos}} \cdot MTF_{\text{ccd}}. \tag{6} \]

In general, the contrast of any image which has propagated through an image acquisition system is worse the contrast of the input image.

2.1 The photographic objective design

The Cooke Tyler triplet photographic objective present interest because it represents the simplest lenses form capable of correcting the five fourth-order wavefront aberrations: spherical aberration, coma, astigmatism, field curvature, and distortion; and both lateral and transverse chromatic aberration \([5, 17, 22]\).

In the limiting case of three lenses in contact (zero element separation), the distortion and transverse chromatic aberrations are zero. The three powers can be used to control the total power, longitudinal chromatic aberration, and Petzval curvature. This yields a set of three equations in three unknowns

\[ \begin{align*}
\phi_1 + \phi_2 + \phi_3 &= \phi \\
\frac{\phi_1}{n_1} + \frac{\phi_2}{n_2} + \frac{\phi_3}{n_3} &= 0 \\
\frac{\phi_1}{V_1} + \frac{\phi_2}{V_2} + \frac{\phi_3}{V_3} &= 0
\end{align*} \tag{7} \]

where \( \phi_1, \phi_2, \phi_3 \) are the element powers, \( n_1, n_2, n_3 \) are the element indices of refraction and \( V_1, V_2, V_3 \) are the corresponding Abbe numbers. The first equation determines the total power, the second the Petzval curvature, and the third the longitudinal chromatic aberration. Solutions to this set of linear equations exist providing that three different glass types are chosen. More information about the design of the MTF and the PSF (Fig. 4) of a Cook Tyler triplet lenses can be obtained using professional software like: ZEMAX, OSLO and CODEV.

2.2 The photographic objective aperture

The circular aperture has the formula

\[ c(r) = \text{circ} \left( \frac{r}{r_0} \right) \tag{8} \]

\( r \) is the circle radius, \( r_0 \) is the cut off radius, and the PSF is calculated as

\[ c(x, y) = \frac{\lambda}{r_0} \frac{r_0}{r} J_1 \left( \frac{\lambda r_0}{r} \right) \tag{9} \]

\( \lambda \) is the wavelength, \( J_1 \) is the Bessel function of order one.

A perfect optical system is diffracted limited by the relation

\[ d = 2.44 \lambda N \tag{10} \]

where

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

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 \( J_1(r) \) for a circular aperture \([1-4, 6, 8, 15, 19-21]\).

![Fig. 4 a) the MTF of the Cook Tyler triplet lenses b) the PSF of the Cook Tyler triplet lenses](image)

![Fig. 5 a) the circular aperture, b) the PSF of the circular aperture](image)
2.3 The light fall off

The 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 [19, 21]

\[ E_i = \pi \cdot L \frac{1}{1 + 4f/\#(1-m)^2} (\cos \phi)^4. \]  

(11)

Where \( L \) is source light radiation, \( m \) is the magnification.

![Fig. 6 The light falloff a) the cos4f law, b) the PSF of the cos4f](image)

2.4 The MTF of the CCD

An image sensor is a spatial (as well as temporal) sampling device. Consequently, its spatial resolution is governed by the Nyquist sampling theorem. The spatial frequencies that are above the Nyquist rate, cause aliasing and cannot be recovered. Bellow the Nyquist rate is the useful radiation. The low pass filtering caused by integration and diffusion degrades the reproduction of frequencies below \( f_{Nyquist} \), degradation measured by the modulation transfer function. Consequently the cutoff frequencies \( f_{cx} \) and \( f_{cy} \) must be

\[ f_{cx} \geq 2f_s \]
\[ f_{cy} \geq 2f_y. \]

The lens (Fig. 8) acts as a lowpass filter with a cutoff frequency in the frequency domain given by

\[ f_s = f_y = \frac{2NA}{\lambda} \]

(12)

where the \( NA \) is the numerical aperture of the lens and \( \lambda \) is the shortest wavelength of light used with the lens.

We are interested to see what happens to an image that passes through the optical part of the CCD image sensor. We start by computing MTF and PSF. We consider a 1-D doubly infinite image sensor (Fig. 8) where \( L \) quasi neutral region, \( L_d \) depletion depth, \( w \) the pixel size [7, 14, 18-24].

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 \( i_{ph}(x) \) can be represented by the linear space invariant system. \( i_{ph}(x) \) is sampled at regular intervals \( p \) to get the pixel photocurrents. After certain manipulation we have

\[ MTF(f) = H(f) \frac{D(f)}{D(0)} w^2 \sin c(wf) \]

(13)

\[ D(0) \]

is called the diffusion MTF and \( \sin c(wf) \) is called the geometric MTF. We also have

\[ MTF_{CCD} = MTF_{diffusion} \cdot MTF_{geometric}. \]

(14)

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 \( w \)

\[ MTF = \frac{H(f_s, f_y)}{H(0)} \]

\[ = \frac{D(f_s, f_y)}{D(0)} w^2 \sin c(wf_s) \sin c(wf_y) \]

(15)

![Fig. 7 The modulation transfer function a) the MTF of the CCD, b) the PSF of the CCD](image)
imaging system is worse than the contrast of the input image.

2.5 The dimension of the pixel
A solid-state image sensor, like CCD or CMOS imagers is a semiconductor device that converts an optical image that is formed by an imaging lens into electronic signals. To reproduce an image with acceptable resolution, a sufficient number of picture elements or “pixels” must be arranged in rows and columns [14].

![Fig. 8 A simplified pixel cross section view](image)

*z* represents the distance between pixels, *w* represents the pixels width, *L* represents the quasi neutral region, *Ld* represents the depletion length.

In figure 8 we have the cross section through the pixel and we can see that it is part of a periodic structure of pixels [14]. The picture presents a complex device structure compound of the optical part lenses and colors filters and the analog part responsible with conversion from photons to charges and then in to voltage. Supplementary (not represented in the figure) we have conversion from analog to digital signal and numeric colors processing on the same chip [7, 10, 14].

To find the maximum size of a pixel in the CCD image sensor we use equation (10). The sensor is located in focal plane of the lenses, the wavelength \( \lambda = 550\,\text{nm} \) and the magnifying coefficient \( M = 1 \). Applying these values to equation (10) we obtain

\[
d = 2.44 \cdot 550 \cdot 11\,\mu\text{m} = 10.833\,\mu\text{m} \\
\frac{d}{L} = \frac{d}{M} = 10.833\,\mu\text{m}
\]

To deliver sufficient sampling the pixel size should be smaller then

\[
p = \frac{d}{2} = \frac{10.833}{2} = 5.4\,\mu\text{m}
\]

According to the relation between the FOV of object space and image height shown in equation (16), if FOV and the size of CCD are selected, the effective focal length is determined

\[
2y = 2f \tan\left(\frac{\omega}{2}\right)
\]

where \( 2y \) is the diagonal size of CCD, \( f \) is the effective focal length and \( 2\omega \) 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

\[
P = \frac{206265 \times p}{1000 \times f}
\]

where 206265 is the number of arcseconds in 1 radian and 1000 is the conversion factor between millimeters and microns [12, 23].

2.6 The fill factor
The microlens from figure 8 collimates incident light to the photodiode. The microlens condenses light onto the photodiode and effectively increases the fill factor. The microlens plays a very important role in improving light sensitivity on CCD image sensors [14].

The fill factor (FF) is defined as the ratio of the photosensitive area inside a pixel, to the pixel area,

\[
FF = \frac{W_x \cdot W_y}{L_x \cdot L_y} \cdot 100\%.
\]

The larger the fill factor the more light will be captured by the chip up to the maximum of 100%. This helps improve the dynamic range. As a tradeoff, the larger values of the fill factor mean more spatial smoothing due to the aperture effect.

2.7 The quantum efficiency
Photodetector convert incident radiant power in photocurrent. Incident photons generate e-h pairs in silicon material. Some of the generated carries are converted in to photocurrent [7, 14].

Quantum efficiency (QE) is a quantity defined for a photosensitive device such as photodiode or a
charge-coupled device (CCD) as the percentage of photons hitting the photoreactive surface that will produce an electron-hole pair

\[ QE(\lambda) = \frac{N_{\text{sig}}}{N_{\text{ph}}} \]  

(19)

where \( N_{\text{sig}} \) and \( N_{\text{ph}} \) are the generated signal charge per pixel and the number of incident photons per pixel, respectively.

Part of the incident photons are absorbed or reflected by upper structures above the photodiode. The microlens and photodiode structure determine the effective FF and the charge collection efficiency, respectively

\[ QE(\lambda) = T(\lambda) \cdot FF \cdot n(\lambda) \]  

(20)

where \( T(\lambda), FF, \) and \( n(\lambda) \) are the transmittance of light above a detector, the effective FF and the charge collection efficiency of the photodiode.

\[ N_{\text{sig}} = \frac{i_{\text{ph}} \cdot A_{\text{pix}} \cdot t_{\text{int}}}{q} \]  

(21)

\[ N_{\text{ph}} = \frac{P \cdot A_{\text{pix}} \cdot t_{\text{int}}}{h \nu} \]  

(22)

where

\( i_{\text{ph}} \) is the photocurrent in [\( A/cm^2 \)],

\( A_{\text{pix}} \) is the pixel size in [\( cm^2 \)],

\( P \) is the optical input power in [\( W/cm^2 \)],

\( t_{\text{int}} \) is the integration time,

\( \nu \) is the frequency,

\( h \) is the Plank’s constant.

If we increase the QE then the useful signal increase and consequently the dynamic range ratio increase.

### 2.8 The Bayer CFA

An image sensor is basically a monochrome sensor which responds to light energies that are within its sensitive wavelength range. Thus, a method for separating colors must be implemented in an image sensor to reproduce an image of a color scene. Digital camera with a single detector requires the use of a color filter array (CFA) [7, 14] 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. The green proportion was selected because the human visual system derives image details primarily from the green portion of the spectrum. The luminance differences are associated with green whereas color perception is associated with red and blue. RGB CFAs have superior color reproduction and higher dynamic range due to their superior wavelength selectivity properties [21-23].

![Fig. 9 The Bayer CFA](image)

#### 2.9 The color difference space interpolation

Color interpolation is the process of estimating a value at a location that does not have a measured value by using other actual measured values. It is clear that better estimation can be obtained if more measured pixels are used for the interpolation calculation [16].

The color difference space method proposed by Yuk, Au, Li, and Lam [25] 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.

#### 2.10 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 demosaicking 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 [1, 4, 20, 21].

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

\[
L = \begin{bmatrix}
-1 & -1 & -1 \\
-1 & 8 & -1 \\
-1 & -1 & -1
\end{bmatrix}.
\]  

(23)

2.11 The CCD dynamic range

Dynamic range is the ratio of the maximum to minimum values of a physical quantity. For a scene, the ratio is between the brightest and darkest part of the scene. 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 dynamic range (DR) 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 \( i_{\text{max}} \) to its smallest detectable photocurrent \( i_{\text{min}} \) or the ratio between full-well capacity and the noise floor [7, 14, 16]. The maximum amount of charge that can be accumulated on a photodiode capacitance is called full-well capacity. The initial and maximum voltages are \( V_{\text{reset}} \) and \( V_{\text{max}} \), they depend on the photodiode structures and operating conditions. The largest saturating photocurrent is determined by the well capacity and the integration time

\[
i_{\text{max}} = \frac{qQ_{\text{max}}}{I_{\text{int}}} - i_{\text{dc}}.
\]  

(24)

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

\[
DR = 20 \log_{10} \frac{i_{\text{max}}}{i_{\text{min}}}
\]

\[
= 20 \log_{10} \frac{qQ_{\text{max}} - i_{\text{dc}}I_{\text{int}}}{\sqrt{qI_{\text{int}}i_{\text{dc}} + q(\sigma_{\text{read}}^2 + \sigma_{\text{DNSU}}^2)}}
\]  

(25)

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

3 The color processing

Color is one of the most important aspects to the design of digital still cameras because the first impression of image quality is mostly affected by the quality of color [16, 23].

3.1 The colors saturation

In color theory the saturation or purity refers to the intensity of a specific hue. The saturation of a color is determined by a combination of light intensity and how much of it is distributed across the spectrum of different wavelengths [4, 9, 16]. Image color saturation is obtained by multiplying the images with matrix \( A \)

\[
A = \begin{bmatrix}
1.4333 & -0.2667 & -0.2667 \\
-0.2667 & 1.4333 & -0.2667 \\
-0.2667 & -0.2667 & 1.4333
\end{bmatrix}.
\]  

(26)

3.2 The white balance

One of the most challenging processes in a digital camera is to find an appropriate white point and to adjust color. The equivalent process that happens in the retina by adjusting the cones’ sensitivity is called chromatic adaptation. 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 apply a gain to the response of each of the human cone cell spectral sensitivity in order to keep the adapted appearance of the reference white constant [9, 16]. 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

\[
D = \begin{bmatrix}
0.5844 & 0 & 0 \\
0 & 0.5753 & 0 \\
0 & 0 & 0.514
\end{bmatrix}.
\]  

(27)

3.3 The color correction

We need to specify two aspects of the display characteristics in order to specify how the displayed image affects the cone photoreceptors [9, 22, 23]. 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. 11)), and the relative absorptions of the human cones (Fig. 10)).

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

3.4 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 [1, 4, 16, 21-23] 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

$$V_{out} = V_{in}^\gamma.$$  

(28)

3.5 The color conversion
We convert the device-dependent RGB data into XYZ format [1, 4, 16, 21-23] using the color calibration information specified in color correction paragraph. RGB represent the color space. Red, green and blue can be converted in X, Y and Z using the conversion matrix

$$
\begin{bmatrix}
X \\
Y \\
Z
\end{bmatrix} =
\begin{bmatrix}
0.431 & 0.342 & 0.178 \\
0.222 & 0.707 & 0.071 \\
0.229 & 0.130 & 0.909
\end{bmatrix} 
\begin{bmatrix}
R \\
G \\
B
\end{bmatrix}.
$$

(29)

3.6 The JPG compression
Compression can affect the digital picture resolution causing picture artifacts’ comparable with those presented in figure 1. In our simulation we use the JPG compression which is a commonly used method of lossy compression for digital image which means that some original image information is lost and cannot be restored. In order to avoid this situation we chose the maximum compression ratio whose result is an image that cannot be distinguished by the eye from the original [1, 4]. This algorithm, based on luminance-chrominance color spaces and discrete cosine transforms, performs a spatial frequency transform of the image, and then quantizes the frequency components in order to eliminate visually redundant information, which, as a result, produces a smaller image representation.

4 The simulation results
In our simulations we demonstrate using images the functionality of an image capture system from the resolution and luminosity point of view and then we process the colors [2, 18-24]. From image (Fig.12 a)) to image (Fig.13 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, which become 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 figure 13 c) we have a Bayer CFA sampled image. By using a good interpolation technique we can minimize the pixel artifacts (Fig.14 a)). We
sharpen (Fig. 14 b)) and we set the dynamic range (Fig. 14 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 figure 15 a) we use the Von Kries matrix, a classic 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 figure 15 b). Comparing (Fig. 10) and (Fig. 11) we see that the spectrums of the red colors 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 figure 15 c). In figure 16 we have the comparison between original and compressed image.

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.

5 Conclusion
The analysis and simulations presented in this paper covers an important part of a digital camera pipeline related to the image acquisition system and color processing and JPG compression. An important characteristic of these sensors is the tradeoff between the resolution, light sensibility and exposure with the dynamic range. Enhancing the input signal we enhance the dynamic range and consequently the resolution of the image. This analysis can be useful in understanding and learning the functionality of the digital camera pipeline and to help people who design and implement digital cameras.

Acknowledgment
This paper was supported by the project "Development and support of multidisciplinary postdoctoral programmes in major technical areas of national strategy of Research - Development - Innovation" 4D-POSTDOC, contract no. POSDRU/89/1.5/S/52603, project co-funded by the European Social Fund through Sectoral Operational Programme Human Resources Development 2007-2013."
References: