Simulation the functionality of a web cam image capture system

TOADERE FLORIN, NIKOS E. MASTORAKIS INCDTIM Cluj Napoca Str. Donath, nr. 65-103, ClujNapoca, 400293, ROMANIA Florin.Toadere@bel.utcluj.ro

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

Abstract: - The goal of this paper is to imagine the functionality of a web camera image capture system. In order to do that we illuminate a spectral image, we capture the reflected light using an achromatic doublet and finally we focus the light on a CCD. We are interested to see what happens as the image propagates trough the image capture pipeline. We present an illumination algorithm, we design the achromatic doublet, we use an aperture and we compute the CCD transfer function. Using the CCD integration times we take several images in rapid succession, at different exposure levels; then we blend the images in order to obtain a HDR (high dynamic range) image. We consider the image capture system to be axial and the light is orthogonal to the system.

Key-Words: - spectral image illumination, achromatic doublet design, CCD transfer function, HDR image, photon shot noise, colors enhancement

1 Introduction

Imaging systems are complex and require to transforms signals through a number of different devices. Consequently, understanding components in isolation, without reference to the characteristics of the other system components, provides only a limited view. In these types of complex systems a controlled simulation environment can provide the engineer with useful guidance that improves the understanding of design considerations for individual parts and algorithms [3].

In recent years digital cameras gained a lot of popularity, and tend to replace the classical photographic cameras. The optical part, still remain the same no mater if the camera is digital or classical. The most important characteristic of a photographic objective is the quality of a photography, which is characterized by resolution. Optical resolution describes the ability of an imaging system to resolve detail in the object that is being imaged. Each component of the lenses contributes to the optical resolution of the system, as will the environment in which the imaging is done.

In this paper we are interested to take the picture of the color checker image. In order to do that we simulate the functionality of the image capture system (the illumination and the optical part) as the illuminate object image passes trough it. The quality of photography depends on the input illumination condition and how the web camera can preserve the light. The exposure of the image sensor depends on the lens f/number, sensor exposure time, scene illumination level, and scene reflectance, as well as many other secondary factors.

Our system consists of an EG&G PerkinElmer flash lamp [13], a spectral image [14] used as input object, an achromatic doublet which captures the reflected light and focuses the output light on a CCD sensor. Several images are taken in rapid succession, at different exposure levels, when the user depresses the camera shutter button halfway down [2, 5-7, 10]. We blend together the images in order to obtain a HDR image. In the low illuminate images we have visible photon shot noise. The photon shot noise represents the photon to charge conversion noise and is directly relate to the illumination degree. We made a HDR image by blending together the noisy low illuminated images with no noise light saturated images. We sharp to enhance the clarity of details in the image; we make the histogram adjustment and colors saturation in order to improve the image colors and to reduce the photon shot noise [1, 3, 5].

2 The image capture system

2.1 The illumination algorithm

Human vision is sensitive to visible light, that of the electromagnetic spectrum with part wavelengths from about 400 to 700 nm. The illumination determines the amount of light that covers a surface. Color helps in the perception of the beauty of the digital image. The perceived color of the surface is determined not only by the color of the surface but also by the color of the light. Therefore, the type of lamp in the light fixture can significantly impact the perception of the color of the object. This effect should be taken into account when we see images on the computer screen. The perception of the object's shape differs with the light distribution on its surface and with the configuration of the resulting shadows. The direction of the light beam can easily affect the perception of the object's shape [3, 5, 9, 15].

Color vision is the capacity of an organism or machine to distinguish objects based on the wavelengths of the light they reflect or emit. Color derives from the spectrum of light interacting in the eye with the spectral sensitivities of the light receptors. The nervous system derives color by comparing the responses to light from the three types of cone photoreceptors in the eye L, M, S (long, medium and short) equivalent to R, G, B (red, green and blue) colors [3, 5, 9, 15].

The EG&G PerkinElmer flash is represented on 31 wavelengths because the spectral image is defined as a 256X256X31. In this paper we use a spectral image under the Spectral Binary File Format (.spb). This format has the following characteristics: file identifier is a 3 letter string SPB (Spectral Binary file) located at the beginning of the file. Image dimensions and wavelength values are stored in file header. Dimensions (x, y and n) are stored in uint32 format and wavelength values in foat32 format. Spectral image values are reflectance values stored as float32. Spectral image values are scaled between 0 and 1, where 1 describes maximum reflectance. Image data is written to the file in column order and values are stored in little endian form [14].

If we perceive light that is reflected from a surface, instead of light that is directly emitted from a light source, our eyes receive result of the scalar product of reflectance and radiance spectrum [5, 9, 15]. In continuous case the human eye response is:

$$c_i = \int_{\lambda \min}^{\lambda \max} (\lambda) r(\lambda) l(\lambda) d\lambda, \qquad i = S, L, M \qquad (1)$$

where:

 $S_i(\lambda)$ is the function of sensitivity of the *i*-th type of cones,

 $r(\lambda)$ is the fraction of the reflected illuminant energy,

 $l(\lambda)$ is the spectral distribution of light.

L, M, and S are the responses of the long, medium, and short cones of the eye [9, 15].



PerkinElmer

The image obtained using equation 1 is not enough from the monitor's colors possibilities of representation. In order to remediate this deficiency we have to make compatibility between monitor possibility of colors generation and how the human eyes cones perceive the colors' radiance. We need to specify how the displayed image affects the cone photoreceptors [3, 5, 15]. To make this estimate we need to know: the effect that each display primary has on your cones and 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 of the display; a CRT (cathode ray tube) monitor (Fig. 2), and the relative absorptions of the human cones (Fig. 1) [15]. 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.

$$\begin{bmatrix} 14.0253 - 13.5154 & 0.7385 \\ -4.1468 & 10.1490 & -1.3618 \\ -0.1753 & -0.5663 & 7.3776 \end{bmatrix}$$
(2)

In addition, the characteristics of the display device (screen display) where the digital image is viewed also affect the intensity distribution and interrelationship of contrast between light and dark regions in the specimen. Phosphors of monitors do not react linearly with the intensity of the electron beam. For CRT display monitors and televisions, the luminance produced at the face of the display is a power function, which is proportional to the voltage applied to the faceplate grid raised to an exponential power. The numerical value of the exponent of this power function is known as gamma [1, 5].



Fig. 2 the CRT monitor spectral reflectance

In conformity with the equation (1:3) we have the next algorithm steps:

1. Load the data into Matlab (spectral image, lamps spectrums and cones response) [9, 13, 14].

2. For each lamp spectrum illuminations, we compute the human eyes color response using equation 1.

3. We make compatibility between monitor possibility of colors generation and how the human eyes cones perceive the colors radiance using equation 2.

4. We make gamma correction, to correct the monitor luminance using equation 3.

2.2 The achromatic doublet design

The optical sensor is analyzed in space by the PSF (point spread function) and in the spatial frequency by the MTF (modulation transfer function). MTF and PSF are the most important integrative criterions of imaging evaluation for the optical system and are determined according to application.

In order to determine the doublet impulse system response and the transfer function, we use a LSI (linear shift invariant system) which is characterized by [1, 12]:

$$g(x, y) = H[f(x, y)]$$
(4)

H is an operator representing a linear, position (or space) invariant system.

$$g(x, y) = \int_{-\infty}^{\infty} f(\alpha, \beta) H[\delta(x - \alpha, y - \beta)] d\alpha d\beta$$

$$g(x, y) = \int_{-\infty}^{\infty} f(\alpha, \beta) h(x - \alpha, y - \beta) d\alpha d\beta$$

$$h(x - \alpha, y - \beta) = H[S(x - \alpha, y - \beta)] \text{ is the impulse}$$
(5)

 $h(x-\alpha, y-\beta) = H[\delta(x-\alpha, y-\beta)]$ is the impulse response of H; in optics, it is called the point spread function (PSF) [1, 2, 4, 8].

The sensor's PSF is a multiple convolution of individual response from: the doublet, the aperture and the transfer function of a CCD [17]

$$PSF = PSF_{doublet} * PSF_{aperture} * PSF_{CCD}.$$
 (6)

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

$$H(\alpha,\beta) = \frac{\iint P(u+\frac{\alpha}{2},v+\frac{\beta}{2})P(u-\frac{\alpha}{2},v-\frac{\beta}{2})dxdy}{\iint P(u,v)^2 dxdy}$$

The numerator represents the area of overlap of two pupil functions *P* (square or circle), one of which is displaced by $\frac{\alpha}{2}$ in direction *u* and by $\frac{\beta}{2}$ in direction *v*; and the other in opposite direction by $-\frac{\alpha}{2}$ and -

 $\frac{\beta}{2}$. The denominator represents the complete area

of the pupil function.

The change in contrast when an image passes trough an optical system is expected to have a lot to do with the optical transfer function that specifies the quality of the system.

The MTF (modulation transfer function) is defined as: the ratio of the contrast of the output image to that of the input image

$$MTF = \frac{contrast \ of \ output \ image}{contrast \ of \ input \ image}.$$

The relation between *OTF* and *MTF* is:

$$MTF = |OTF|. \tag{7}$$

In conclusion, the modulation transfer function is identical to the absolute value of the optical transfer function. The net sensor's MTF is a multiplication of individual transfer functions in a way similar to equation 6.

$$MTF = MTF_{doublet} \cdot MTF_{aperture} \cdot MTF_{CCD}$$
(8)

In general, the contrast of any image which has gone through an image capture system is worse that the contrast of the input image. We compute the doublet merit function which can be seen as the MTF.

2.2.1 The achromatic doublet glasses

An achromatic doublet [2, 4, 8, 19], represent a simple web cam photographic objective which is composed of two lenses made from different glass types, one a low-dispersion crown glass (V > 50) and the other a high-dispersion flint glass (V < 50). The combination is required to have total power f and zero axial chromatic aberration. Crown in front is preferred because crown glasses are harder and less susceptible to abrasion or weathering than flint glasses. We propose to design an achromatic doublet with a front positive lens of crown glass with BK10 glass (from Schott 497670) and a focal length of 12.2mm. We should find what flint glass will use and to determine the power for that second lens to achieve an overall focal length of 23mm.



Fig. 3 the doublet function geometrical representation

We have the next mathematical relations for the total power of the systems:

$$\phi = \phi_1 + \phi_2 \,. \tag{9}$$

Individual power of each lens function of total power is:

$$\phi_1 = \frac{V_1}{V_1 - V_2} \phi \tag{10}$$

$$\phi_2 = -\frac{V_2}{V_1 - V_2}\phi \,. \tag{11}$$

We change the values: $V_1 = \frac{1}{f_1} = \frac{1}{12.2}$; $V_2 = \frac{1}{f_2}$;

 $V_1 - V_2 = \frac{1}{23}$; and we obtain: $V_2 = 39.75$ and $f_2 = -22.9mm$.

Using the formulation for the power of the first lens relative to the overall power gives an Abbe number for the second lens of 39.8. From the Schott catalog, the closest Abbe number to this is 39.4 for N-BASF6. Using the fairly common FK5 glass (less expensive and more available) would lead to a focal length of 23mm. It point to the fact that we may have to compromise somewhat on design based on the glasses available and how customized we want to figure the lens.

2.2.2 The aberrations analysis

In this paper we simulate the doublet functionality [2, 4, 8, 19], and in function of the lenses glasses we design the parameters and we compute the merit function. In order to analyze aberrations we consider the exit pupil to be the mechanical finite dimension of the lenses. We use the Seidel aberrations formula:

$$W(x, y) = W_{020}r^{2} + W_{111}hy + W_{200}h^{2} + W_{040}r^{4} + W_{131}hyr^{2} + W_{220}h^{2}r^{2} + W_{222}h^{2}y^{2} + W_{311}h^{3}y +$$
(12)
$$W_{400}h^{4}$$

Using this formula we can express the thin lenses aberrations function of power, in our analyses we use: axial chromatic, coma and spherical aberrations

$$W_{020} = \frac{1}{2\lambda} y_a^2 \frac{\phi}{V},$$
 (13)

$$W_{131} = \frac{1}{4\lambda} y_a^2 \phi^2 L(a_5 B - a_6 M), \qquad (14)$$

$$W_{040} = \frac{1}{16} y_a^2 \phi^3 (a_1 + a_2 (B - a_3 M)^2 - a_4 M^2).$$
(15)

 λ is the wave length, y_a is the aperture, ϕ is the power, V is the Abbe number, $L = -nu_a y_c$ is the Lagrange invariant, $B = \frac{c_1 + c_2}{c_1 - c_2}$ is the bending, n_g is

the glass refraction indices, $M = \frac{1+m}{1-m}$ is the magnification,

$$a_{1} = \left(\frac{n_{g}}{n_{g}-1}\right)^{2}, a_{2} = \frac{n_{g}+2}{n_{g}(n_{g}-1)^{2}}, a_{3} = \frac{2(n_{g}^{2}-1)}{n_{g}+2}$$
$$a_{4} = \frac{n_{g}}{n_{g}+2}, a_{5} = \frac{n_{g}+1}{n_{g}(n_{g}-1)}, a_{6} = \frac{2n_{g}+1}{n_{g}}.$$

2.2.3 The aberrations correction

We have the mathematical relation that describes the optical design which implies Seidel aberrations [2, 4, 8, 11, 19]. The defect vector f is a set of m functions f_i that depend on a set on n variables. The function is of the type:

$$\sigma^2 = f^t \cdot f \tag{16}$$

A is a $(n \times m)$ matrix of first derivatives:

$$A_{ij} = \frac{\partial f_i}{\partial x_j}$$

and s are changes in the variables from the current design. The gradient g is a $(n \times 1)$ vector given by:

$$g = \frac{1}{2} \nabla \sigma^2 \tag{17}$$

its components are:

$$g_{i} = \frac{\partial \sigma^{2}}{\partial x_{i}} = 2 \left(f_{1} \frac{\partial f_{1}}{\partial x_{i}} + f_{2} \frac{\partial f_{2}}{\partial x_{i}} + \dots f_{m} \frac{\partial f_{m}}{\partial x_{i}} \right)$$
$$g = A^{t} f .$$
(18)

Method of Least Squares

$$g = A^{t} (f_{0} + As)$$
$$g = g_{0} + A^{t} As$$
$$C = A^{t} A$$
$$g_{0} + Cs = 0$$

is a set of simultaneous linear equations known as the normal equations of least-squares. Providing that the matrix C is not singular, these equations can always be solved, and the formal solution s may be written:

$$s = -C^{-1}g_0. (19)$$

The basic idea of the damped least-squares is to start with the basic equation for the least squares condition. g_0 is the gradient at the starting point and augment the diagonal of the matrix *C* by the addition or factoring of a damping coefficient. Modifications of the form $c_{ii} + p$ for example, are called additive damping. In the case of additive damping, the equation for the damped least-squares solution reduces to:

$$g_0 + ps + Cs = 0. (20)$$

As the damping factor p increases, the third term in the equation above becomes small and the solution vector becomes parallel to the gradient vector.

$$s = -\frac{1}{p}g_0.$$

2.2.4 The merit function

We make the lens focal length 23mm with an f/2 aperture ($y_a = 5mm$). Let the half field angle u_c be 0.1 (5.73°) and the wavelength be 555nm. Let the glass index of refraction $n_g = 1.5$. Assume the object is at infinity (M = 1).

The defect functions are longitudinal chromatic, coma and spherical aberrations. The wave front errors are equations 13-15. To solve this problem we must solve the next equation system [4, 8]:

$$\begin{cases} f_1 = \phi - (\phi_1 + \phi_2) \\ f_2 = \frac{\phi_1}{V_1} + \frac{\phi_2}{V_2} \end{cases}$$
(22)

The merit function is the solution of equation 22 (Fig. 4) and can be called the achromatic doublet's MTF. From Fig. 4 a) we see that MTF presents a strong discontinuity due to the aberrations. The PSF

is represented in (Fig. 4 b)) and we see that the PSF's spatial configuration is spread all along the surface like strips. Knowing this we expect to have a strong blurred image at the output of the doublet.



Fig. 4 a) the doublet MTF, b) the doublet PSF

2.2.5 The aperture

In order to take a photo we need an 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 [2, 4, 8, 19].

The circular aperture has the formula:

$$c(r) = circ\left(\frac{r}{r_0}\right) \tag{23}$$

r is the circle radius;

 r_0 is the cut off radius.

The PSF is calculated as:

$$c(x, y) = \lambda \frac{r_0}{r} J_1\left(\lambda \frac{r_0}{r}\right)$$
(24)

 λ is the wavelength;

 J_1 is the Bessel function of order one.

A perfect optical system is diffracted limited by the relation:

$$d = 2.44\lambda N \tag{25}$$

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

f 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 as shown in Fig. 5 b). For the photographic objective f/number is present inside the lenses (Fig. 5 a)) and it is called f stop. The aperture can be seen as a circular LPF (low pass filter) at the end of the lenses.



Fig. 5 a) the aperture, b) the aperture PSF

2.3 The CCD analysis

We are interested to see what happens to an image that passes to the optical part of a CCD sensor; in order to do that we must find the MTF and the PSF.

The CCD (charge coupled device) is a complex sensor which converts focalized light in to numerical signal. Actually is a dynamic analog charge shift register implemented using closely spaced MOS (metal oxide semiconductor) capacitors using 2, 3, 4 phase clocks. Capacitors operate in deep depletion regime when clock is high. CCD image sensors consists of $n \times m$ array of pixels, each pixel contains: a photodetector that converts the incident light in to photocurrent, circuits for reading out photocurrent; part of the readout circuits are in each pixel, the rest placed at the periphery of the array.

The CCD 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 spatial (and temporal) frequencies. So spatial (or temporal) frequency components higher than the respective Nyquist rate cannot be reproduced and cause alias in the image sensor, however, is not a point sampling device in space (or time), and cannot be approximated as such. Photocurrent is integrated over the photodetector area (and in time) before sampling [2, 7, 10].

Diffusion photocurrent may be collected by neighboring pixels; instead of where it is generated these effects result in low pass filtering and crosstalk before spatial sampling. Assuming a square pixel with width p, the spatial Nyquist rate in each dimension is $f_{Nyquist} = \frac{1}{2p}$ and is typically reported in line pairs (lp/mm). Signals (image) with spatial frequencies higher than $f_{Nyquist}$ cannot be faithfully reproduced, and cause aliasing. The low pass filtering caused by integration and diffusion

reproduction of frequencies

below $f_{Nyquist}$, degradation measured by the Modulation Transfer Function (MTF).



Fig. 6 the spatial sampling

We analyze that part of the CCD sensor responsible with conversion from photons to charges. We treat the modulation transfer function of the CCD, the photon shot noise, the dynamic range and the exposure time. We do not treat Bayer sampling, interpolation and blooming.

2.3.1 The CCD modulation transfer function

The CCD optical part is characterized by its afferent MTF. The contrast in an image can be characterized by the modulation [2, 7, 10]

$$M = \frac{s_{\max} - s_{\min}}{s_{\max} + s_{\min}}$$
(26)

where s_{max} and s_{min} are the maximum and minimum pixel values over the image. Note that: $0 \le M \le 1$.

Let the input signal to an image sensor be a 1-D sinusoidal monochromatic photon flux:

$$F(x, f) = F_0 [1 + \cos(2\pi f x)]$$
(27)

for $0 \le f \le f_{Nyquist}$.

The sensor modulation transfer function is defined as:

$$MTF(f) = \frac{M_{out}(f)}{M_{in}(f)}$$
(28)

from the definition of the input signal, $M_{in} = 1$. MTF is difficult to find analytically and is typically determined experimentally. For the beginning we made a 1-D analysis for simplicity and at the end we generalize the results to 2-D model which we will use in our analyses.

By making several simplifying assumptions, the sensor can be modeled as a 1-D linear space-invariant system with impulse response h(x) that is real, nonnegative, and even. In this case the transfer function:

$$H(f) = F[h(x)] \tag{29}$$

is real and even, and the signal at *x* is:

degrades

the

$$S(x) = F(x, f) * h(x)$$

$$S(x) = F_0 [1 + \cos(2\pi f x)] * h(x)$$

$$S(x) = F_0 [H(0) + H(f) \cos(2\pi f x)]$$
(30)

therefore:

$$S_{\max} = F_0 [H(0) + |H(f)|]$$

$$S_{\min} = F_0 [H(0) - |H(f)|]$$

and the sensor MTF is given by:

$$MTF(f) = \frac{|H(f)|}{H(0)}$$
(31)

We consider a 1-D doubly infinite image sensor.



Fig. 7 the CCD sensor model

L- quasi neutral region

 L_d - depletion depth

w- aperture length

p- pixel size

To model the sensor's response as a linear spaceinvariant 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 (from $-\infty$ to ∞). The monochromatic input photon flux F(x) to the pixel current iph(x) can be represented by the linear space invariant system (Fig. 8). iph(x) is sampled at regular intervals p to get the pixels photocurrents.



integration

$$r\left(\frac{x}{w}\right) = \begin{cases} 1 & |x| \le \frac{w}{2} \\ 0 & otherwise \end{cases}$$
(32)

d(x) is the (spatial) impulse response corresponding to the conversion from photon flux to photocurrent density, and we assume a square photodetector. The impulse response of the system is thus given by its Fourier transform (transfer function) [2, 7, 10]

 $r \rightarrow$

$$h(x) = d(x) * \omega r\left(\frac{x}{\omega}\right)$$
(33)

and its Fourier transform (transfer function) is given by:

$$H(f) = D(f)\omega^2 \sin c(\omega f)$$
(34)

note that: $D(0) = n(\lambda)$.

 $n(\lambda)$ - spectral response.

By definition: the spectral response is a fraction of photon flux that contributes to photocurrents as a function of wave length. So D(f) can be viewed as a generalized spectral response (function of spatial frequency as well as wavelength).

After long calculus we get D(f) as:

$$D(f) = \frac{q(1 + \alpha L_f - e^{\alpha L_d})}{1 + \alpha L_f} - \frac{qL_f \alpha e^{\alpha L_d} \left(e^{\alpha L} - e^{\frac{L}{L_f}}\right)}{\left(1 - (\alpha L_f)^2\right) \sinh\left(\frac{L}{L_f}\right)}$$
$$H(f) = D(f) w^2 \sin c(wf)$$
(35)

the modulation transfer functions for $|f| \le \frac{1}{2p}$ is:

$$MTF(f) = \frac{|H(f)|}{H(0)} = \frac{D_f}{D_0} w^2 \sin c(wf)$$
(36)

 $\frac{D_f}{D_0}$ is called the diffusion MTF and $\sin c(wf)$ is

called the geometric MTF.

So we have:

$$MTF_{CCD} = MTF_{diffusin} \cdot MTF_{geometric}$$
 (37)

But in our analyses we use 2D signals (image) so we must generalize 1D case to 2D case. We know that we have square aperture at each photodiode with length w; so the analyses is made in Cartesian coordinate and we must generalize in x-y coordinate MTF(f) and we have:

$$MTF(f_x, f_y) = \frac{|H(f_x, f_y)|}{H(0)}$$
(38)
$$MTF(f_x, f_y) = \frac{D_{(f_x, f_y)}}{D_0} w^2 \sin c(wf_x) \sin c(wf_y)$$

 f_x - spatial frequency on x direction

 f_{y} - spatial frequency on y direction

Spatial frequency (lines/mm) is defined as the rate of repetition of a particular pattern in unit distance. It is indispensable in quantitatively describing the resolution power of a lens. The first level in a CCD image sensor is a lens which focuses the light on each pixel photodiode.



Fig. 9 a) the CCD MTF, b) the CCD PSF

2.3.2 The photon shot noise

Image noise is a random, usually unwanted, variation in brightness or color information in an image. In a CCD sensor image noise can originate in electronic noise in the input device sensor and circuitry, or in the unavoidable shot noise of an ideal photon detector. Image noise is most apparent in image regions with low signal level, such as shadow regions or underexposed images. In this paper we focus our attention on the photon shot noise produced by the input captured photons which are transformed in to charges. The photon shot noise has the Poisson distribution [1-3, 5-7, 20]

$$P(k,\lambda) = \frac{e^{-\lambda}\lambda^k}{k!}$$
(39)

 $k = 1 \div n$, λ is a positive real number.

The photon shot noise relative strength decreases with increasing signal, shot noise is often only a problem with small currents or light intensities. The intensity of a source will yield the average number of photons collected, but knowing the average number of photons which will be collected will not give the actual number collected. The actual number collected will be more than, equal to, or less than the average, and their distribution about that average will be a Poisson distribution.

We are interested about photon shot noise effect in low illuminated image. To minimize the noise effect we try to improve the luminosity and colors in the final image. In order to do that we blend the low illuminate images which present photon shot noise with the normal and high illuminate images which don't have noise.



Fig. 10 photon shot noise in low illumination



Fig.11 photon shot noise in moderate illumination



2.3.3 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} [2, 3, 6, 7, 10]. The largest saturating photocurrent is determined by the well capacity and integration time

$$\dot{i}_{\max} = \frac{qQ_{\max}}{t} - \dot{i}_{dc} \tag{40}$$

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

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

$$= 20\log_{10} \frac{qQ_{\text{max}} - i_{dc}t_{\text{int}}}{\sqrt{qt_{\text{int}}i_{dc} + q(\sigma_{read}^2 + \sigma_{DNSU}^2)}}$$
(41)

 $q = 1.6 \times 10^{-19} C$ 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).

Integration time t_{int} tells us how much time the aperture is open and function of this how much light enter in the CCD.

2.3.4 Parameters of the optical systems

To find the maximum size of a pixel in the CCD image sensor we use equation 25. 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 25, 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:

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

1/2 inch CCD with C optical interface is selected, i.e. its back working distance is 23 ± 0.18 mm. The visual band optical system has 60° FOV, f/number 2.5. According to the relation between the FOV of object space and image height if FOV and the size of CCD are selected, the effective focal length is determined.

$$2y = 2f \tan \omega \tag{42}$$

2y is the diagonal size of CCD,

f is the effective focal length,

 2ω is the full field of view in object space.

It is important to know the pixel dimensions because in function of this, we can control the amount of light that enter in to the CCD and therefore we can control the photon shot noise [3, 7, 18].

2.4 The colors enhancements

After the image has gone trough our image capture system it is noisy blurred and its contrast is degraded. We try to improve the image colors characteristics by blending the different exposed images, histogram adjustment, sharpening and colors saturations.

2.4.1 The colors blending

This method is defined as that each pixel in the resulting image is an average of the pixels from all the exposures, but the weight for each pixel is different. This algorithm works for sets of multiple images even with just two images. With two exposures, there is a long (L) exposure and a short (S) exposure. We use the grayscale value of the long exposure as the weight of the short exposure because the bright pixels in the long exposure may

be blown out or actually a bright object. In either case, we would want to use the pixel value in the short exposure. For each pixel, the resulting pixel is a weighted average of the short and long exposure pixel where the grayscale value of the long exposure pixel is the weight for the short pixel. The pixel is scaled such that energy of the pixels is not increased. This easily extends to multiple images. First, blend the two images with the longest exposure as described. This is repeated until all the images are used. This method works well, is computationally easy and in general does a pretty good job at blending the multiple exposures [3, 6, 16]

2.4.2 The sharpening

We sharp the image, in order to eliminate the blur caused by the optical system components. Sharpness describes the clarity of detail in a photo. We use a Laplacian filter [1, 3, 5, 8]:

$$L = \begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}.$$
 (43)

2.4.3 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 [3, 5, 15]. Image color saturation is obtained by multiplying the images with matrix A:

$$A = \begin{bmatrix} 1.4333 & -0.2667 & -0.2667 \\ -0.2667 & 1.4333 & -0.2667 \\ -0.2667 & -0.2667 & 1.4333 \end{bmatrix}.$$
 (44)

3 The simulation results

In this paper we try to imagine the optics of a web cam. In order to do that we cover the web cam pipeline related to light propagations and photons to charges conversion [1-20]. We use the color checker spectral image (Fig. 13 a)), which is illuminated with an EG&G PerkinElmer lamp (Fig. 13 b)), the object reflected light is captured by an achromatic doublet (Fig. 14 a)). The captured image is strongly blurred at the doublet output. In order to eliminate the blur we sharp the image using a Laplacian filter (Fig. 14 b)). Finally we focus the light on a CCD. Function of the integration time durations, the aperture is open and we can take several light exposed pictures. The short time exposed images have visible photon shot noise and the long time exposed images tend to be saturated. We change the integration times as seen in Fig.10 to Fig.12. Then we blend together the long exposed and low exposed images (Fig. 10-12) in order to get a HDR image (Fig. 15).

In Fig. 13 a) we have the original image. The image is illuminated wit a Perkin&Elmer flash lamp (Fig. 13 b)). As a result the original image changes its hues and shines. Then the image is captured by a doublet lenses. The effect of the all optical components together deteriorates the image resolution (Fig. 14 a). In Fig. 14 b) we have the sharp image and we observe that the optics don't deteriorates the quality of the image, we have only some hues and faded little colors differences. Then the light is focalized on the optical part of the CCD and in the process of conversion from light to charges appears the photon shot noise in function of the light quantity (Fig. 10-12). If is to much light we can mask the photon shot noise but the blended image is light saturated and shine too much. If the short exposed images are preponderant then the noise is visible and the image keeps its original colors. It is a balance between the images light intensities. We need to compromise somehow and if we want to have a good contrast in the final HDR image, we need to use all the intermediate exposed pictures (Fig. 10 to Fig. 12). In Fig. 12 we have adequate illumination and the photon shot noise vanishes in these illumination conditions. Also we can have more intense illumination and light saturation conditions, but because the image do not has shadows we stop the illumination at the normal eye enough illumination and perception conditions. Finally we saturate the colors. In Fig. 15 we have the HDR image; we see that the image has an enhanced contrast comparing with the Fig. 13 b). The image presents little quantity of noise after the blending process in which we use our illuminated images Fig. 10 to Fig. 12. Better results can be obtained if we used much and more illuminated images also electrical filters can eliminate the little amount of noise that still exists.

4 Conclusion

In this paper we imagine the optical part of a web cam. From the simulated images we see the good quality of the illuminated object. The reflected light is captured bay a doublet and is focalized on a CCD. During this process the optics deteriorates the resolution and the CCD introduces the photon shot noise. Using properly the colors processing techniques we can recover the image colors characteristics at the output of the objective. Finally we make a HDR image in order to minimize the photon shot noise effect and to increase the image contrast.



Fig. 13 a) original image, b) flash illuminated image



Fig. 14 a) image at the output of the optical part, b) sharp image



Fig. 15 blended noisy images

References:

- [1] Castleman K., *Digital image processing*, Prentice Hall, 1996
- [2] www.optics.arizona.edu/detlab/
- [3] http://scien.stanford.edu/jfsite/Publications_Ima geCapture.htm
- [4] Geary J. M., Introduction to lens design with practical Zemax example, William Bell, 2002
- [5] Gaurav S., *Digital Color Imaging Handbook*, CRC Press, 200
- [6] B. Hoefflinger (Ed.), *High-Dynamic-Range* (*HDR*) *Vision*, Springher, 2007
- [7] Holst G., Lomheim, T, CMOS/CCD Sensors and Camera Systems, SPIE, 2007
- [8] http://www.engr.udayton.edu/faculty/jloomis/co urses.html

- [9] Lee H. C., *Introduction to color imaging science* Cambridge, 2005
- [10] Lutz G., *Semiconductor radiation detectors*, Springer Verlag, 2007
- [11] Mastorakis N., Singular Value Decomposition in Multidimensional Arrays, *Circuits, Systems* and *Computers* 1996. International Conference with the approval and support of the Hellenic Navy General Staff. Hellenic Naval Academy, Piraeus, GREECE, July 15-17, 1996.
- [12] Mastorakis N., A continuous model for 2-Dimensional (LSI) Systems, PLENARY LECTURE in the International Colloquium on Numerical Analysis and Applications, Plodviv, Boulgaria, August 13-17, 1998.
- [13] http://optoelectronics.perkinelmer.com/content/ whitepapers/XenontheFullSpectrum.pdf
- [14] http://spectral.joensuu.fi/index.php?page=spi
- [15] Steven K. Shevell (Ed.), The Science of Color Second Edition, Elaseiver, 2003
- [16] Toadere F., Conversion from a spectral image to a HDR image, ATOM-N 2008 The 4th international conference advanced topics in optoelectronics, microelectronics and nanotechnologies, 29-31 August, Constanta, Romania.
- [17] Toadere F., Simulation the optical part of an image capture system, ATOM-N 2008 The 4th international conference advanced topics in optoelectronics, microelectronics and nanotechnologies, 29-31 August, Constanta, Romania.
- [18] Toadere F., Mastorakis N., Imaging the Optical Part of a Web Cam, Signal Processing, Computational Geometry and Artificial Vision, ISCGAV 2009, Moscow, Russia.

[19] <u>www.zemax.com</u>

[20] <u>http://en.wikipedia.org/wiki/Shot_noise</u>