BULETINUL INSTITUTULUI POLITEHNIC DIN IAŞI
Publicat de
Universitatea Tehnică „Gheorghe Asachi” din Iaşi
Tomul LVII (LXI), Fasc. 6, 2011
Secţia
ELECTROTEHNICĂ. ENERGETICĂ. ELECTRONICĂ

SIMULATION OF THE FUNCTIONALITY FOR A CMOS IMAGE SENSOR WITH PASSIVE PIXELS

BY

FLORIN TOADERE[1,*] **and RADU ARSINTE**[2]

[1]INCDTIM–Cluj-Napoca
and
[2]Technical University of Cluj-Napoca

Received, May 31, 2011
Accepted for publication: July 6, 2011

Abstract. The goal of this paper is to simulate the functionality of a passive pixel CMOS image capture sensor. The sensor is a complex optic, analog and digital device that converts the incident light into numerical signal. The optical part is characterized by the components' transmittance and their spatial resolution. The analog circuits during the functionality introduce temporal and spatial noises. The digital part converts the signal into digits and then process signals. The simulations present the essential features of the image, during its propagations trough a passive pixel CMOS sensor, in the process of conversion from light to numeric signal.

Key words: passive pixel CMOS sensor optics; analog and digital parts.

1. Introduction

In recent years digital cameras gained a lot of popularity and replaced the classic film cameras. The photographic objectives remain the same but the incident light is converted into numeric signals, consequently the recording film is no more necessary. The visual information is captured *via* a sensor placed in the focal plane of the optics of the camera. The sensor measures the light gathered during the exposure time. To capture an image means to convert the light information contained in an image to the corresponding electrical signals that can be stored in a reproducible way. The advantage of these sensors is that

---

[*]Corresponding author: *e-mail*: florin.toadere@bel.utcluj.ro

they integrate the capture, the signals and colours processing on the same chip. The CMOS sensors can be divided in PPS (passive pixel sensors) and APS (active pixel sensor). In this paper we analyse a PPS CMOS sensor. By specifying the sensor properties, the simulations can predict their performances and functionality. We simulate the essential features of an image that passes trough the PPS CMOS sensor.

2. The Optical Part of a PPS CMOS Sensor

The PPS CMOS image capture sensors is a complex device which converts the focalized light into numerical signal. CMOS image sensors consists of a $n \times m$ array of pixels; each pixel contains a photo detector that converts the incident light into photocurrent, circuits for reading out photocurrent; part of the readout circuits are in each pixel, the rest are placed at the periphery of the array (Holst & Lomheim, 2007; Image Capture). CMOS sensors integrate on the same chip the signal capture and processing.

In our analyses we use a pixel made in 0.5 µm technology. To model the response of the sensor as a linear space invariant system, we assume *n+/p-sub* photodiode with very shallow junction depth. Consequently, we can neglect the generation in the isolated *n*+ region and we only consider the generation in the depletion and *p*-type quasi-neutral regions. We assume a uniform depletion region. The values of the pixel parameters are (Holst & Lomheim, 2007; Toader & Mastorakis, 2009, 2010): $z = 5.4$ µm, $L_d = 4$ µm, $L = 10$ µm, $w = 4$ µm, $\lambda = 550$ nm $\lambda = 550$ nm. 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 (field of view) and f/number is 2.5. z represents the distance between pixel, w is the pixels width, L – the quasi neutral region and L_d – the depletion length.

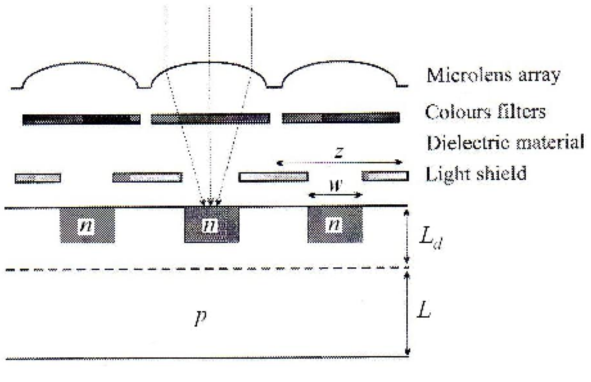


Fig.1 – The cross section view of a simplified pixel

In Fig. 1 the cross section of a pixel is represented and we can see that it is part of a periodic structure of pixels (Holst & Lomheim, 2007). The figure presents the structure of a complex device which contains the lenses, the colours filters and the analog part.

These parts are responsible with the conversion from photons to electrical charges. Supplementary (not represented in the Fig. 1) we have on the same chip the conversion from analog to digital and the numeric colours processing. The collected photons increase the voltage across the photodiode proportional with the incident photon flux. The photodiodes work by direct integration of the photocurrent and dark current.

2.1. The Light Capture

The micro lenses condense the incident light on the photodiodes. Generally speaking the lenses are made from quality materials ensuring a good visible band transmittance (98…99%). The Bayer colour filter array covers the CMOS sensor and consists of red, green and blue filters. The filters sample the incident white light. Red and blue filters are in proportion of 25% percents each and green filters are in proportion of 50% percent. Each filter covers one of the three colours waveband and it is characterized by its specific spectral power distribution (SPD). The quantum efficiency (QE) is a quantity specific for a photosensitive device such as photodiode or PPS CMOS sensors, and is expressed as the percentage of photons hitting the photoreactive surface that will produce an electron-hole pair (Holst & Lomheim, 2007; El Gamal & Eltoukhy, 2005; Janesick, 2007;Kang, 2006; Sharma, 2003).

The CMOS PPS sensor receives the result of the scalar product of the reflectance of the scene with the spectral power distribution of light, the Bayer filters and the photodiode quantum efficiency. In our simulation we take in consideration also the responses of the cones of the human eye

$$c_i = \int_{\lambda_{\min}}^{\lambda_{\max}} S_i(\lambda)F_i(\lambda)QE_i(\lambda)r(\lambda)l(\lambda)\mathrm{d}\lambda\,,\ (i = L, M, S), \tag{1}$$

where: L, M and S are, respectively, the responses of the long, medium, and short cones of the eye, $S_i(\lambda)$ – the spectral energy of the illuminant, $r(\lambda)$ – the fraction of the reflected illuminant energy, $l(\lambda)$ – the spectral distribution of light, $F_i(\lambda)$ – the filter transmittance, $QE_i(\lambda)$ – the diode quantum efficiencies.

The image obtained using eq. (1) is not appropriate for the possibilities to represent the colours in the monitor. In order to compensate this deficiency we have to arrange some compatibility between the monitor possibility of colour reproduction and how the human eye's cones perceive the radiance of the colours. We need to specify how the displayed image affects the cones photoreceptors. To make this estimate we need to know the following: 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 (detailed are further furnished by Toadere & Mastorakis, 2009)

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

Gamma represents a numerical parameter that describes the nonlinearity of the reproduction of the luminance. Phosphorous of the monitors do not react linearly with the intensity of the electron beam (Sharma, 2003). Instead the input value is effectively raised to an exponent called gamma

$$V_{\mathrm{out}} = V_{\mathrm{in}}^{\gamma}. \tag{3}$$

2.2. The Spatial Resolution

The modulation transfer function (MTF) characterizes the frequency response of the optical system (Holst & Lomheim, 2007; Toadere & Mastorakis, 2009). In a linear imaging system where $i(x,y)$ is a two-dimensional optical input and $o(x,y)$ is the output, the relationship between the input and the output is

$$o(x,y) = \int\int_{-\infty}^{+\infty} h(x-x_0, y-y_0)\mathrm{d}A, \tag{4}$$

where h is the system impulse response.

The frequency relation is obtained using the Fourier transform

$$O(f_x,f_y) = H(f_x,f_y)I(f_x,f_y), \tag{5}$$

where H is the optical transfer function (OTF).

The relation between OTF and MTF is

$$MTF(f_x,f_y) = \left| \; O(f_x,f_y) \; \right|. \tag{6}$$

After certain calculus (Toadere & Mastorakis, 2009) we obtain

$$\mathrm{MTF} = \frac{|H(f_x,f_y)|}{H(0)} = \frac{D(f_x,f_y)}{D(0)} w^2 \sin c(wf_x) \sin c(wf_y). \tag{7}$$

3. The Electrical Part of a PPS CMOS Sensor

The PPS CMOS image sensor consists of a $n \times m$ PPS array. They are based on photodiodes without internal amplification. In these devices each pixel consists of a photodiode and a transistor. After addressing the pixel by opening the row-select transistor, the pixel is reset along the bit line. The readout is performed one row at a time. At the end of integration, the charge is read out *via* the column charge to voltage amplifiers. The amplifiers and the photodiodes in the row are then reset before the next row readout begins. The main advantage of PPS is its small pixel size. In spite of the small pixel size capability and a large fill factor, they suffer from low sensitivity and high noise due to the large column's capacitance with respect to the pixel's one (Holst & Lomheim, 2007; El Gamal & Eltoukhy, 2005; Janesick, 2007; Sharma, 2003; Toadere & Mastorakis, 2010). Also during the propagations of the signal trough the bit configuration it suffers of temporal noises perturbations.

3.1 The Electrical Noises

Image noise is a random, usually unwanted, variation in brightness or colour information in an image. In a CMOS sensor, the image noise can originate in electronic noise that can be divided in temporal and FPN (fixed pattern noise), or in the unavoidable photon shot noise of an ideal photon detector.

The photon shot noise is associated with the random arrival of photons at any detector. The lower the light levels the smaller the number of photons which reach the detector per unit of time. Consequently there will not be a continuous illumination but a bombardment by single photons and the image will appear granulose. The signal intensity, *i.e.* the number of arriving photons per time unit, is stochastic and can be described by the Poisson distribution (Holst & Lomheim, 2007; Janesick, 2007; Sharma, 2003; Toadere & Mastorakis, 2009, 2010)

$$P(k,\lambda)=\frac{\mathrm{e}^{-\lambda}\lambda^{k}}{k!},\quad (k=1,\dots,n); \tag{8}$$

n is a non-negative integer and λ is a positive real number.

The readout noise of a PPS CMOS is generated by the electronics and the analog-to-digital conversion. Readout noise is usually assumed to consist of independent and identically distributed random values; this is called *white noise*. The noise is assumed to have the normal white Gaussian distribution with mean value 0 and a fixed standard deviation proportional to the amplitude of the noise. The analog to digital convertor produces quantization errors, whose effect can be approximated by uniformly distributed white noise. The standard deviation is inversely proportional to the number of bits used. The white Gaussian noise has been widely investigated in offline contexts.

3.2. The Fixed Pattern Noise

The fixed pattern noise (FPN) is defined as the pixel-to-pixel output variation under uniform illumination due to device and interconnected mismatches across the image sensor array. These variations cause two types of FPN: the offset FPN, which is independent from pixel signal, and the gain FPN or photo response non uniformity, which increases with signal level (El Gamal *et al*., 1998). Offset FPN is fixed from frame to frame but varies from one sensor array to another. The most serious additional source of FPN is the column FPN introduced by the column amplifiers (Holst & Lomheim, 2007; El Gamal & Eltoukhy, 2005; Janesick, 2007; Toadere & Mastorakis, 2010). In general PPS has FPN, because PPS has very large operational amplifier offset at each column. Such FPN can cause visually objectionable streaks in the image. Offset FPN caused by the readout devices can be reduced by CDS (correlated double sampling). Each pixel output is readout twice, once right after reset and a second time at the end of the integration. The sample after the reset is then subtracted from the sample after the integration.

3.3. The 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. 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 non-saturating photocurrent, $i_{\max}$, to its smallest detectable photocurrent, $i_{\min}$, or the ratio between full-well capacity and the noise floor (Holst & Lomheim, 2007; Janesick, 2007). 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_{reset} and $V_{\max}$; they depend on the photodiode structures and the operating conditions (Holst & Lomheim, 2007; El Gamal & Eltoukhy, 2005; Janesick, 2007; Toadere & Mastorakis, 2009, 2010).

The largest saturating photocurrent is determined by the well capacity and integration time, t_{int}

$$i_{\max} = \frac{qQ_{\max}}{t_{\text{int}}} - i_{\text{dc}} . \tag{9}$$

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

$$\mathrm{DR} = 20\log_{10}\frac{i_{\max}}{i_{\min}} = 20\log_{10}\frac{qQ_{\max} - i_{\mathrm{dc}}t_{\mathrm{int}}}{\sqrt{qt_{\mathrm{int}}i_{\mathrm{dc}} + q(\sigma_{\mathrm{read}}^2 + \sigma_{\mathrm{DNSU}}^2)}}\,, \tag{10}$$

where: $Q_{\max}$ is the effective well capacity; σ_{read}^2 – the readout circuit noise; σ_{DNSU}^2 – the offset FPN due to dark current variation, commonly referred to as DSNU (dark signal non-uniformity).

3.4. The PPS CMOS High Definition

The dynamic range of the PPS CMOS sensor is generally not wide enough to image scenes with great contrast. In order to solve this problem several images are taken in rapid succession, at different exposure levels, when the user depresses the camera shutter button halfway down (Sharma, 2003). We note that none of these images contains all the details in the scene. All the images are captured within a single normal integration time. We work with six different light exposed images. In order to obtain a high dynamic range image the captured images are blended together. Further information about the method were given by Toadere & Mastorakis (2009, 2010).

3.5. The Colour Difference Space Interpolation

The colour difference space method proposed by Yuk, Au, Li and Lam (2007) interpolates pixels in green-red and green-blue colour difference spaces as opposed to interpolating on the original red, green, and blue channels. The underlying assumption is that due to the correlation between colour channels, taking the difference between two channels yields a colour difference channel with less contrast and edges that are less sharp. Interpolating an image with less contrast yields fewer glaring errors, as sharp edges, what cause most of the interpolation errors in reconstructing an image. The colour 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 interpolated image.

3.6. The Analog-to-Digital Conversion

The analog-to-digital conversion is the last block of the analog signal processing circuits in the CMOS image sensor. In order to convert the analog signal into digital signal we compute the analog to digital curve, the voltage swing and the number of bits. The quality of the converted image is good and the image seams to be unaffected by the conversion (Holst & Lomheim, 2007; El Gamal & Eltoukhy, 2005; Janesick, 2007; Toadere & Mastorakis, 2010).

4. The Image Reconstruction

The input photon flux, in the process of the radiation capture, is deteriorated by the optics and electronics. The optics is responsible for the colours fidelity and the spatial resolution; the electronics introduce temporal and spatial electrical noises. In order to eliminate the colours cast we make the white balance. At the output of the electrical part the image is corrupted by the optical blur and the effects of the combined noises. In order to reduce the blur we use a Laplace filter, to reduce the FPN we use a frequencies amplitude filter which blocks the spikes spectrum of the FPN. We blend the different light exposed images in order to reduce the photon shot noise. Finally we reject the remains noise using a bilateral filter and we sharp with a Laplace filter.

4.1. The Laplacian Filter

In order to correct the blur and to preserve the impression of depth, clarity and fine details we have to sharp the image using a Laplacian filter (Sharma, 2003; Toadere & Mastorakis, 2009, 2010). The Laplacian filter is a 3 × 3 pixel mask

$$L = \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}. \tag{11}$$

To restore the image we subtract the Laplacian image from the original image.

4.2. The Amplitude Filter

The FPN is introduced by the sensor's column amplifiers and consists of vertical stripes with different amplitudes and periods. Such type of noise, in the Fourier plane, produces a periodic set of spikes. A procedure to remove this kind of noise is to make a transmittance mask in Fourier 2-D logarithm plane. The first step is to block the principal components of the noise pattern. This block can be done by placing a band stop filter, $H(u,v)$, in the location of each spike (Sharma, 2003; Toadere & Mastorakis, 2010). If $H(u,v)$ is constructed to block only components associated with the noise pattern, it follows that the Fourier transform of the pattern is given by the relation

$$P(u,v) = H(u,v)\log[G(u,v)], \tag{12}$$

where $G(u,v)$ is the Fourier transform of the corrupted image $g(x,y)$.

After a particular filter has been set, the corresponding pattern in the spatial domain is obtained making the inverse Fourier transform

$$p(x,y)=F\left[\exp P(u,v)\right]. \tag{13}$$

4.3. The Bilateral Filter

In order to reduce the remainder noise, after image blending and amplitude filter, we use a bilateral filter. It extends the concept of Gaussian smoothing by weighting the filter coefficients with their corresponding relative pixel intensities. Pixels that are very different in intensity from the central pixel are weighted less even though they may be in close proximity to the central pixel. This is effectively a convolution with a non-linear Gaussian filter, with weights based on pixel intensities. This is applied as two Gaussian filters at a localized neighborhood of the pixel, one in the spatial domain, named the *domain filter*, and one in the intensity domain, named the *range filter* (Toadere & Mastorakis, 2010).

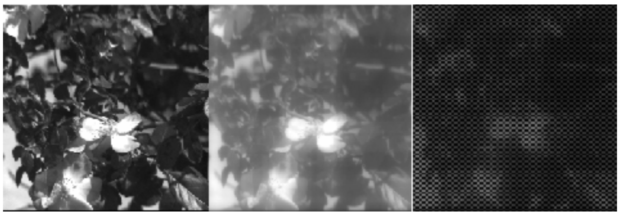
4.4. The White Balance

The process of colour balancing tries to determine what the illuminant of a particular scene is and then to adjust the intensities of the red, green, and blue channels of the image in order to recover the original colour characteristics of the scene. Since for most images we can not accurately determine the scene illuminant, we considered a Mean and Standard Deviation algorithm (Sharma, 2003). We noticed that many of the images seemed to be lacking in contrast, making them appear hazy. They also appeared to be quite dark in colour, an indication of low mean channel values. Therefore, in addition to adjusting the standard deviation values we adjust the mean of each channel as well. To compensate these shortcomings in the original image, we adjust the mean of each channel to be 0.5, and we also set the standard deviation of each to be roughly 0.27 (or 70 on a 256 value scale).

5. The Results of Simulation

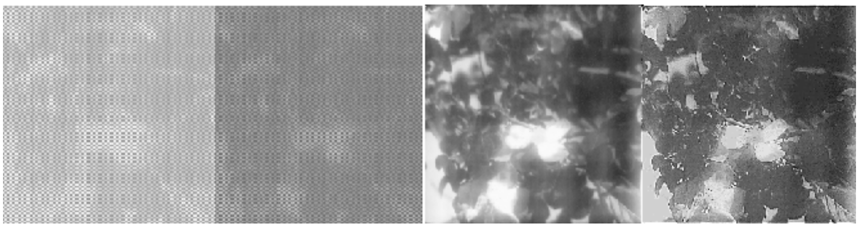
In Fig. 2 *a* the colours are obtained in function of the filters spectrums and the QE of the photodiodes array. In Fig. 2 *b* the MTF act like a low pass filter and blurs the image. In Fig. 2 *c* we have the Bayer sampled image. The sampled signal is covered by the photon shot noise. This kind of noise is inverse proportional with the intensity of the light. In function of the exposure times we obtain images with shot noise or without shot noise. In addition we have the FPN specific to PPS CMOS (Figs. 3 *a* and 3 *b*). We are speculating the inverse relation between the illumination and the photon shot noise and in Fig. 3 *c* we have the blended image, the white balance and amplitude filtered image. By blending the different exposed images we can reduce the photon shot noise.

Also we note that the FPN was rejected but remain some little shadows of noise lines. In order to obtain a good quality image, we filter the image with a Laplace and a bilateral filter (Fig. 3 *d*). Other configurations can be done using other filters and interpolations algorithms.



a *b* *c*

Fig. 2 – *a* – The image at the output of the optical part; *b* – the spatial resolution; *c* – the Bayer CFA sampling.



a *b* *c* *d*

Fig. 3 – *a* – The longest exposure, the shot noise and the FPN; *b* – the shortest exposure, the shot noise and the FPN; *c* – the blended six different times exposed images, interpolation and amplitude filter; *d* – the bilateral filter, sharp and analog to digital conversion.

6. Conclusions

The simulations realized in this paper emphasize some essential features of an image which was captured by a PPS CMOS sensor. Our model tries to assemble, all the main steps in the CMOS sensor light processing, in a coherent sequence. The optics is focusing the incident light on the photosensitive area of pixels. The photon shot noise is generated by the unavoidable photons arrivals and is added to the conversion noise. A good solution to attenuate this noise is to blend different light exposed images. This kind of sensor has an FPN component, which is removed by the amplitude filter. A bilateral filter is necessary in order to reject the remainder noises. The white balance is necessary in order to reduce the colours cast introduced by the optics. An important characteristic which characterize the performance of these sensors is the trade-off between the resolution, the light sensitivity and the exposure with the photon shot noise.

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

El Gamal A., Eltoukhy H., *CMOS Image Sensors*. IEEE Circ. and Devices Mag., **21**, *3*, (2005).

El Gamal A., Fowler B., Min H., Liu X., *Modeling and Estimation of FPN Components in CMOS Image Sensor*. Solid State Sensor Arrays: Development and Applications II, January, 26-28, 1998, San Jose, USA, Proc. SPIE, **3301**, 168-177.

Holst G.C., Lomheim T.S., *CMOS/CCD Sensors and Camera Systems*. SPIE Press, Bellingham, 2007.

Janesick J.R., *Photon Transfer*. SPIE Press, Bellingham, 2007.

Kang H.R., *Computational Color Technology*. SPIE Press, Bellingham, 2006.

Sharma G., *Digital Color Imaging Handbook*. CRC Press, Boca Raton, 2003.

Toadere F., *Conversion from Light to Digital Signal in a Digital Camera Pileine.* ATOM-N 2010, The 5th of the Internat. Conf. Adv. Topics in Optoelectron., Microelectron. a. Nanotechnol., August 26-29, 2010, Constanța, Romania, Proc. SPIE, **7821**, 124-134.

Toadere F., Mastorakis N.E., *Simulation the Functionality of a Web Cam Image Capture System*. WSES Trans. on Circ. a. Syst., **8,** *10*, 811-821 (2009).

Yuk C.K.M., Au O.C., Li R.Y.M., Lam S.-Y., *Color Demosaicking Using Direction Similarity in Color Difference Spaces.* Internat. Symp. on Circ. a. Syst., 2007, ISCAS 2007, May 27-30, 2007, New Orleans, USA, 1281-1284.

* * * http://meroli.web.cern.ch/meroli/Lecture_Particle_Detector_Noise.html

* * * http://people.csail.mit.edu/sparis/bf_course/

* * * http://personalpages.manchester.ac.uk/staff/david.foster/Research/My_PDFs/Nascimento_etal_JOSAA_02.pdf

* * * http://scien.stanford.edu/jfsite/Publications_ImageCapture.htm

SIMULAREA FUNCȚIONĂRII UNUI SENZOR CMOS CU PIXELI PASIVI

(Rezumat)

Este prezentat un model care ilustrează, prin intermediul imaginilor, funcționarea unui senzor CMOS cu pixeli pasivi. Acest tip de senzor este realizat dintr-o componentă optică, una analogică şi una digitală. Scopul lor este de a converti radiația luminoasă incidentă pe suprafața optică în semnal numeric. Partea optică este responsabilă cu transmitanța şi rezoluția componentelor. În timpul funcționării circuitelor analogice semnalul este deteriorat de zgomote temporale şi spațiale. Partea digitală converteşte semnalul analog în semnal numeric şi prelucrează culorile. Simulările din această lucrare prezintă transformările ce au loc asupra unei imagini în decursul propagării prin senzorul CMOS.