

Virtual Sensor Design

Roberto Costantini and Sabine Süssstrunk

Audiovisual Communication Laboratory, EPFL, Lausanne, Switzerland

ABSTRACT

We present a virtual digital camera sensor, whose aim is to simulate a real (physical) image capturing sensor. To accomplish this task, the virtual sensor operates in two steps. First, it accepts a physical description of a given scene and simulates the entire process of photon sensing and charge generation in the sensor device. This process is affected by noise, mostly photon noise. Second, it adds to the image the noise that results from the electronic circuitry. We present a model for the different sources of noise relative to each sensor-based image formation step, and use measurements of real digital camera images to validate the model.

Keywords: Digital cameras, noise, CCD, CMOS.

1. INTRODUCTION

A digital image is the result of three main steps: the formation of the optical image, involving the passage of light through the optics of the system, the conversion of irradiance into voltage or current, performed in the camera sensor, and some image processing operations, such as demosaicing, quantization, tone and color correction, and compression, performed on the A/D converted raw data. Each step introduces a particular type of noise. Lenses generally cause spatial/chromatic distortions, light is affected by photon shot noise, while CCD or CMOS sensors introduce random variations of the pixels values. Finally, firmware image processing algorithms introduce unwanted artifacts.

Digital imaging system manufacturers are naturally interested in understanding the origins and characteristics of digital image noise, since this knowledge would permit them to understand the impact of certain system design choices, such as type of lens, pixel size and position, color filters, etc., on the noise content of output images. For example, the knowledge of the noise characteristics of the raw image data would permit to design better and more efficient image processing algorithms, such as demosaicing and compression, which at the moment are usually optimized for either noise-free or simple Gaussian-like noisy images.

The interest in considering the entire digital image formation process has pushed researchers and industry to think about the design of virtual systems that mimic each part of the digital image formation process, as done in,¹ for example. In this way, the designers of a digital imaging system can control each part of the process and optimize single or groups of functions.

In this article, we consider the first two steps of the process described above, i.e. the image formation (excluding lens distortions) and the optical to electrical conversion operated by the sensor. We have designed a *virtual sensor* whose aim is to mimic the noise that a real sensor would produce.

The virtual sensor is software that receives as input a physical description of the scene, and computes the number of photons that strike the sensor at each pixel position. The sensor converts these photons into a current that is proportional to the light intensity and produces the relative pixel values coded with n bits. Noise relative to photon sensing, pixel non-uniformity, dark current, read-out, etc., is modeled and added to the image content, thus permitting to obtain the final noisy image.

In Section 2, we will describe the image formation principle. considered in our model. In Section 3, a detailed description of the different types of noise considered will be given, while in Section 4 we will illustrate the model of the noise used in the simulation. In Section 5, we will show the results obtained in a comparison between images produced by the simulation and some images acquired using a real digital camera. Section 6 concludes the article.

E-mail: roberto.costantini@epfl.ch, Telephone: ++41 21 693 66 09

2. SENSOR IMAGE FORMATION

We define the *sensor image* as the image that is formed on the sensor array of the digital camera by purely light-sensing mechanisms. It is created when light passes through the lenses of the camera and strikes the sensor at pixel positions, conveying a certain number of photons proportional to light intensity at those positions. Mathematically, the optical sensor image can be represented by a matrix $\mathbf{N} = \{n_{ij}\}_{(i,j \in I)}$ whose elements correspond to the number of photons that reach the sensor at a pixel position (i, j) .

In the process of image capture, the photons are further converted into electric charges by the light-sensitive element of the sensor (usually a photodiode or a CMOS capacitor). These charges create a current proportional to light intensity, which is converted into numerical values by A/D conversion: these are the pixel values of the digital image.

The process of digital image formation described can be summarized by saying that pixel values depend on the number of electrical charges generated. This is not necessarily equal to the number of photons, as usually more than one photon is needed to generate a single charge. The proportion between the number of charges produced per number of striking photons of a given wavelength λ is called *Quantum Efficiency*, and it is represented by a function $Q(\lambda)$. The pixel values then depend both on the sensor image, represented by the matrix \mathbf{N} , and the quantum efficiency of each pixel. Since this can be computed from the physical characteristics of a given light-sensing element, we consider it as known, thus concentrating on the computation of the matrix \mathbf{N} .

Our Virtual Sensor uses the physical description of a scene of Fig.1 to simulate the sensor image, i.e., to compute the number of photons striking the sensor at each pixel²:

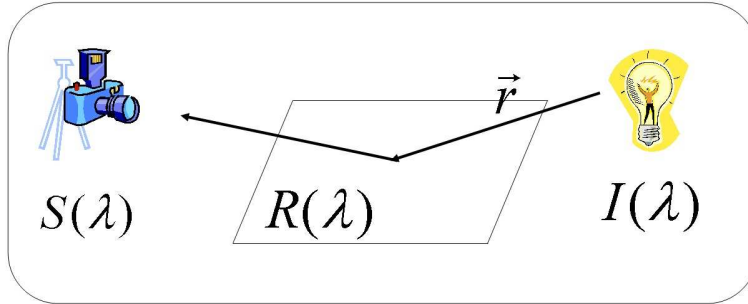


Figure 1. Image formation principle.

The term $I(\lambda)$ expresses the radiant exitance of the scene illuminant, expressed in W/m^3 ; it represents the spectral distribution of the radiation power spatial density, per unit wavelength interval. The term $R(\lambda)$ is the reflectance of the surface, which represents, for each wavelength, the percentage of radiation of that wavelength reflected by the surface. The term $S(\lambda)$ is the spectral sensitivity function of the sensor. Traditionally, it is the product between the quantum efficiency curve of the sensor and the filters sensitivity. If the spectral sensitivity curve is scaled in order to cover an area normalized to one, then it can be seen as a probability distribution function, whose integral computed in the interval $[\lambda_a, \lambda_b]$ gives the probability that a photon of wavelength between λ_a and λ_b reaches the sensor, without being absorbed before.

According to the model depicted in Fig. 1, the spectral distribution of the power density that is conveyed by ray \vec{r} is:

$$C(\lambda) = I(\lambda) \cdot R(\lambda) \cdot S(\lambda). \quad [W/m^3] \quad (1)$$

Given that a photon at wavelength ν has an energy $E = h\nu = hc/\lambda$, where $h = 6.63 \cdot 10^{-34} m^2kg/s$ is the Planck constant, and c is the speed of light, then the number of photons reaching the sensor in one second at position (i, j) is³:

$$n_{ij} = \frac{1}{hc} \int C(\lambda) \lambda d\lambda \quad (2)$$

The Virtual Sensor computes this number for each pixel position (i, j) , thus obtaining the sensor image matrix $\mathbf{N} = \{n_{ij}\}_{i,j \in I}$.

In Fig.1, the scene is very simple, since we have just considered one surface and one illuminant. Our Virtual Sensor can cope also with more complicated scenes, provided that the spectral power distribution of each single ray that strikes the sensor is given. This means that the Virtual Sensor can receive as input *multispectral images*, such as those obtained using ray tracing algorithms, for example.

3. SOURCES OF NOISE

Digital images are affected by noise of different type and origin. The first source of noise is inherent with the image formation principle, where, together with optical deformations and chromatic aberrations introduced by the lenses, photon shot noise is present. This noise is associated to the physical phenomenon of photon emission/absorption: it is not possible to know the exact number of photons that have reached a certain sensor area, since this number is a random variable that follows a certain statistic. This uncertainty corresponds to noise. The second source of noise is associated to photon-current and A/D conversion. Photons in fact generate electrical charges that constitute a current proportional to the irradiance, and this current is converted to digital values by a A/D conversion. In this process, many sources of noise affect the signal, producing a significant part of the final digital image noise content. Typical noise in this process is dark current and thermal noise, which normally affect electronic circuits. The final source of noise is that connected with processing operations that are performed on the pixel values, such as color conversions, demosaicing, and compression. The noise introduced depends on how these operations are done, which is not considered this paper, since it does not actually regard the sensor design.

Table 1 depicts some noise sources and proposes two possible classifications of noise:

Random		Fixed Pattern	
Photon Shot Noise		Amplification Gain Non Uniformity	
Dark Current Shot Noise		PRNU	
Reset Noise		Dark Current Non Uniformity	
Thermal		Column Amplification Offset	
Signal Dependent	Temperature Dependent	Exp. Time Dependent	
Photon Shot Noise	Thermal	Photon Shot Noise	
PRNU	Reset Noise	Dark Current Shot Noise	
Ampl. Gain Non Unif.	Dark Current	Thermal	

Table 1. Different sources of noise and two possible classifications; Table on top: noise classified according to its temporal properties; Table on bottom: classification according to signal, temperature, and exposure time dependence.

A first classification distinguishes between *fixed pattern noise* (FPN) and *random noise*. If we consider many digital images of the same scene, the FPN is the noise component that does not change significantly from image to image, while the random noise corresponds to time variability of pixel values. The FPN can ideally be removed by taking a certain number of images with the shutter closed, i.e. in absence of a signal, where only noise is present, and taking the mean of these images. This results in an estimation of the FPN, which can be subtracted from real images taken in the same conditions (mainly same exposure time and temperature). On the other hand, random noise cannot easily be modeled.

In Table 1, we consider a second classification where the different sources of noise are classified according to their dependency on certain parameters, such as temperature, exposure time, and presence of the signal. Some noise can belong to more than one category since it can depend both on exposure time and signal, such as the photon shot noise, for example.

Not all the sources listed in Table 1 contribute significantly to the noise content of digital images. Thus, we have chosen to consider just the more relevant ones and to describe them in detail, since the model of our virtual sensor will try to reproduce them.

3.1. Photon Noise and PRNU

Photon detection is essentially a random process obeying Poisson statistics.⁴ This means that the number of photons reaching a certain surface during a certain interval of time is a random variable distributed according

to a Poisson law. Mathematically, let T_i be the time interval in which photons arrive, A_s the surface area, and I the intensity of the light source. Indicating with Φ the number of photons striking the surface, we have⁵:

$$\Pr(\Phi = k) = \frac{\lambda^k}{k!} e^{-\lambda}, \quad (3)$$

where $\lambda = E[\Phi] = I \cdot A_s \cdot T_i$ is the mean number of photons reaching the surface. This type of noise depends on the light intensity, i.e., on the signal, and it is an important contribution to the total noise figure.

Another signal-dependent noise is the Photo Response Non Uniformity (PRNU) noise. In a sensor, each photon receptor has slightly different characteristics with respect to the other ones. This means that the same amount of photons reaching the sensor in different locations results in a different number of electric charges at these locations. This constitutes a noise that is signal dependent, since it goes to zero when no radiation arrives at the sensor.

3.2. Dark Current Noise

Dark current is the current produced by minority carriers thermally generated in the sensor well.⁴ The mean number of charges generated in the unit time interval per pixel is given by the following formula⁶:

$$\bar{D} = 2.5 \cdot 10^{15} \cdot A_s \cdot N_{dc} \cdot T^{1.5} \cdot e^{-\frac{E_g}{2kT}} \left[\frac{e^-}{s \cdot pixels} \right], \quad (4)$$

where

$$E_g = 1.1557 - \frac{7.021 \cdot 10^{-4} T^2}{1108 + T} [eV]$$

is the silicon band-gap energy, N_{dc} is the dark current at 300K supplied by the sensor manufacturer [nA/cm^2], k is the Boltzmann's constant, A_s is the pixel area [cm^2], and T is the operating temperature [K].

The dark current is not spatially uniform, since its value depends on the presence of irregularities in the silicon well, which are randomly dispersed on the sensor surface. Moreover, it is also affected by shot noise, since, as photon noise, the number of dark charges that are generated at the pixel position (i, j) varies with time according to a Poisson distribution.

3.3. Reset Noise and Amplification Gain Non Uniformity

Charges generated by the photo-electrical conversion constitute an electric voltage or current that is sensed at the read-out stage of the sensor. This is usually done according to the scheme shown in the right part of Fig. 2(a).

This scheme describes a typical readout circuit used, for instance, in a CCD. In the left part we notice the CCD line, along which charges are transported, and at the end of which they are stored on a $n+$ floating diffusion. The readout scheme is on the right, and it is the same also for CMOS sensors. The principle is the following⁴: when the reset pulse ϕ_P is deactivated, V_{out} assumes the potential V_R , whereas when it is activated, the charges are sensed by a single-stage source follower, decreasing the output voltage V_{out} by a quantity dependent on the charge content, i.e. on the signal. This operation is affected by noise, as shown in Fig. 2(b).

Noise consists in a voltage difference ΔV_R , due to the thermal noise on the resistor R_{on} . It is possible to show that the variance of this noise is $\sigma_q^2 = (CV)^2 = kTC$, where k is the Boltzmann constant.⁴ It does not depend on the resistor value, but on the stray capacity to ground and on temperature. This is called *reset noise* or also *read-out noise*, and it can have rather large values.⁴ For this reason some techniques are used to get rid of it, as the Correlated Double Sampling (CDS), where the voltage $V_R + \Delta V_R$ is sensed together with V_{out} , which is affected by the same ΔV_R . The difference between the two values is then noise-free. Actually, some noise remains, since the CDS circuitry adds itself a certain amount of noise.

For CCD sensors it is reasonable to assume that each pixel is affected by readout noise independently of the adjacent ones, since the readout stage see the same capacitance C of the CCD line. For CMOS sensors this is not true, and the reason is illustrated in Fig. 3.

In CMOS, the charges are directly converted into current by the photodiode, and this current is sensed at output when a column/row scan command is given. In CMOS sensors, there are essentially two ways to do this. In a

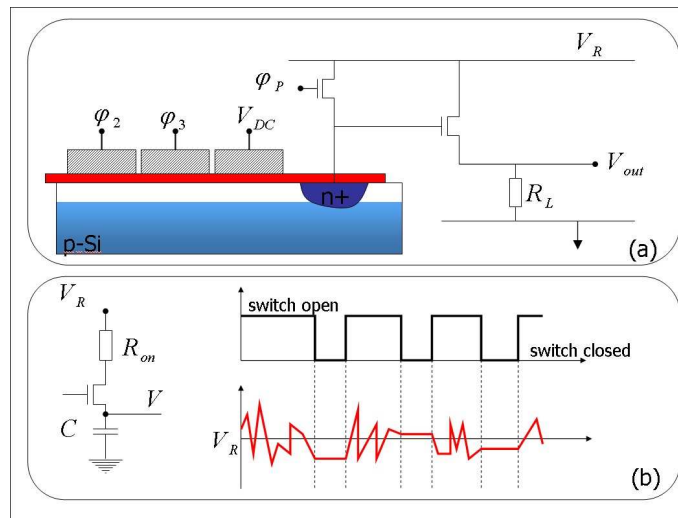


Figure 2. Readout noise in a CCD⁴; (a) readout scheme for a CCD line; (b) result of readout noise: the reference voltage V_R is affected by noise due to the reset operation of capacitor C , which is performed by the MOS-FET transistor.

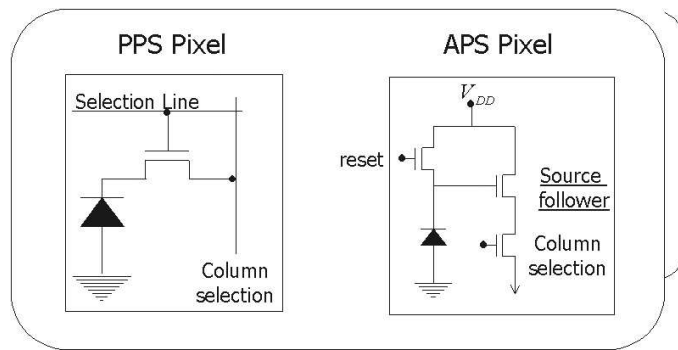


Figure 3. CMOS architecture⁷; (a) Passive Pixel Sensor architecture; (b) Active Pixel Sensor architecture: a source follower is inserted at each pixel.

Passive Pixel Sensor (PPS), the photodiode is directly connected to the column/line selection transistor, which selects the pixel for readout operations. In the first case, the column amplifier is then affected by the strong parasitic capacitance of the entire line, which causes a relevant noise figure. On the contrary, in an Active Pixel Sensor (APS), the first MOS-FET stage of the column amplifier is inserted inside the pixel, as a source follower between the photodiode and the selection transistor.⁷ The input capacitance of this transistor is minimized, and the corresponding noise figure of the amplification path is improved. In any case, both for PPS and APS architectures, the readout noise is correlated along columns: we refer to this as *column noise*.

In APS CMOS there is also a different source of noise, caused by the fact that each pixel has its own source follower. Because of technological limitations, there is a random variation of the gain of the transistor that constitutes a source of noise called amplification gain non uniformity.⁴

4. NOISY IMAGE MODEL

4.1. Notations and Assumptions

Noisy digital images can be modeled by a 2-D discrete stochastic process $\{\mathbf{X}(\vec{r}, t_n)\}_{\vec{r} \in \mathbb{N}^2, n \in \mathbb{N}}$, where $\vec{r} = (i, j)$ is the vector indicating a pixel position in the image, and t_n is the time index; for each $n \in \mathbb{N}$ we have a different realization of the process, i.e., a different image. At a certain fixed position \vec{r}_a along the time axis, the process

becomes a random variable (r.v.) that we denote as $\mathbf{X}_{\vec{r}_a} = \mathbf{X}(\vec{r}_a, t_n)$. When, on the other hand, image index t_n is fixed, the process becomes a 2-D random variable $\mathbf{X}_{t_n} = \mathbf{X}(\vec{r}, t_n)$. This means that the process has a *spatial* and *temporal* variability. We define its spatial and temporal expectations as $E_s[\mathbf{X}] = E[\mathbf{X}_{t_n}]$ and $E_t[\mathbf{X}] = E[\mathbf{X}_{\vec{r}_a}]$, which generally depend on t_n and \vec{r}_a .

If A is a discrete random variable, we indicate with $Y = P_{oiss}(A)$ the discrete r.v. distributed according to a Poisson law of parameter A . We have:

$$\Pr(Y = k) = \sum_{i=0}^{+\infty} \frac{i^k}{i!} e^{-i} \Pr(A = i). \quad (5)$$

2-D models are represented by matrices, but here we assume that pixels are spatially independent. This means that a $n \times m$ noisy image can be considered as $n \times m$ realizations of a single random variable, each corresponding to a particular pixel value. Spatial independence is generally a strong hypothesis, but it can be realistic in the case of CCD noise, where the noise content of each pixel can be considered quite independent from the noise at adjacent locations. For CMOS sensors, this hypothesis does not hold true, but in our model we will consider this case in Section 4.5. So, from now on, we consider just random variables, denoted with capital letters, which represent the value of a given pixel of a certain noisy image.

4.2. Complete Model

We model a noisy image A_{noisy} , where A denotes the noise-free image, as follows:

$$A_{noisy} = G \cdot [P_{oiss}(F) + G_{PRNU} \cdot P_{oiss}(A)] + Z. \quad (6)$$

The random variables G, F, G_{PRNU} , and Z model the amplification gain non uniformity, the dark current noise, the photo response non uniformity, and the noise introduced by circuits, respectively. The operator P_{oiss} considers the fact that dark current and the image are both also affected by shot noise: dark current shot noise and photon shot noise, respectively.

We have chosen for each of these variable a particular distribution, according to the characteristics of the noise they model. To find the correct parameters for this model, we have chosen to work separately with the noise components of A_{noisy} . More precisely, we consider separately the noise that does not depend on the signal A , and the noise that does depend.

4.3. Signal Independent Noise

We call the noise that is independent of the signal “black noise”, and denote it with X_{black} . This is the noise that is present even in the absence of a signal, and can be seen in real digital images if we take some picture with the shutter closed. This noise has a time invariant component, which corresponds to the FPN, and a time variant component. The time invariant component can be estimated taking the mean value of different black images taken under the same conditions of temperature and exposure time and is defined as:

$$X_{FPN} = E_t[X_{black}] = GF + Z, \quad (7)$$

The r.v. G, F , and Z model the pixel gain non uniformity, the temporal invariant dark current, and circuit noise. In our model, we define then as independent, and with the following characteristics:

- G is distributed as a uniform r.v. $G \sim \mathcal{U}([1 - c/2, 1 + c/2])$, where $c \in (0, 1)$ is a given constant that can be set to control its variance. We impose that $\mu_G = E_s[G] = 1$, since the mean gain over the sensor pixels can be considered to be unitary. The spatial variance of the gain is denoted with $\sigma_G^2 = Var_s[G]$ and is equal to $c^2/12$; it does not depend on exposure time;
- F is distributed as a Gamma r.v. $F \sim \Gamma(\beta_F, \lambda_F)$, where $E_s[F] = \mu_F = \beta_F \cdot \lambda_F$, and $Var_s[F] = \sigma_F^2 = \beta_F \cdot \lambda_F^2$; we impose that the dependency on exposure time W is:

$$\mu_F = \alpha W \quad , \quad \sigma_F^2 = (\beta W + \gamma W^2),$$

where α, β , and γ are the parameters of the model that have to be determined;

- Process Z models the noise introduced by the circuits of the sensor. It comprises current shot noise and read-out noise. It is modeled as a Gamma distribution $Z \sim \Gamma(\beta_Z, \lambda_Z)$. For CCD this is an appropriate model, but for CMOS, column noise has to be added, as shown in Sec. 4.5;

This model has 5 parameters to determine: $\alpha, \beta, \gamma, \mu_Z$, and σ_Z^2 . To assign them a value, we assume that the mean value of FPN varies linearly with exposure time, while the variance varies according to a quadratic law as follows:

$$\begin{aligned}\mu_{FPN} &= d_0 + d_1 \cdot W, \\ \sigma_{FPN}^2 &= b_0 + b_1 \cdot W + b_2 \cdot W^2.\end{aligned}\tag{8}$$

The terms d_0, d_1, b_0, b_1 , and b_2 are constant when the temperature is fixed, and they can be estimated from a real digital camera with measurements, or assigned a certain desired value.

Our model is set by imposing that the mean and variance of the simulated FPN have the same dependence on exposure time shown in Eq. 8. This results in the following relations (see Appendix for computations):

$$\begin{aligned}\alpha &= d_1, \\ \beta &= b_1 \left(1 + \frac{c^2}{12}\right)^{-1}, \\ \gamma &= \left(b_2 - \frac{c^2}{12}d_1^2\right) \left(1 + \frac{c^2}{12}\right)^{-1}, \\ \mu_Z &= d_0, \\ \sigma_Z^2 &= b_0.\end{aligned}\tag{9}$$

These equations determine our model, which is now able to predict the noise at different exposure times, for the reference temperature.

If this changes, we introduce a correction term that affects the dark current only, since it is the noise most influenced by temperature:

$$\begin{aligned}\mu_F(T) &= \mu_F(T_0) \left[\frac{D(T)}{D(T_0)}\right]^{0.5}, \\ \sigma_F^2(T) &= \sigma_F^2(T_0) \frac{D(T)}{D(T_0)},\end{aligned}\tag{10}$$

where $D(T), D(T_0)$ is the dark current at temperature T and T_0 respectively. With this corrections, we have obtained the complete model for the FPN, which corresponds to the time independent part of the black noise.

The time variant part of this noise is $X_{temp} = X_{black} - X_{FPN}$ and, according to Eq. 6, it has a zero temporal mean and its temporal and spatial variance are respectively:

$$\begin{aligned}Var_t[X_{temp}] &= Var_t[X_{black}] = \left(1 + \frac{c^2}{12}\right) F + 2\sigma_Z^2, \\ Var_s[X_{temp}] &= Var_s[X_{black}] = \left(1 + \frac{c^2}{12}\right) \mu_F + 2\sigma_Z^2\end{aligned}\tag{11}$$

In fact, if we fix a pixel position (i, j) , the dark current noise has a temporal mean f_{ij} , and variance f_{ij} as well, since it is a Poisson r.v. of parameter f_{ij} , which is the element of position (i, j) of matrix F . The circuit noise has variance σ_Z^2 both in space and in time. So, at a given position, the variance of the temporal noise will be $f_{ij}(1 + c^2/12) + 2\sigma_Z^2$. Analogously, if we consider the temporal noise, its spatial mean will be obviously zero and its variance will be $\mu_F(1 + c^2/12) + 2\sigma_Z^2$.

4.4. Signal Dependent Noise

In the presence of a signal A , two other contributions to the total noise are present: the photon shot noise and the photo-response non uniformity (PRNU). According to the complete model of Eq. 6, the noisy image is then:

$$X_{signal} = G \cdot G_{PRNU} P_{ois}(A) \quad (12)$$

Its temporal mean is A , since $E_t[P_{ois}(A)] = A$ and $E_t[G] = E_t[G_{PRNU}] = 1$. Its temporal variance is well approximated by $A + A^2(1 + c^2/6)$ (see Appendix for computation).

4.5. Readout Noise in CMOS Architectures

CMOS sensors have different readout noise than CCD, since they are affected by the column noise, which is not present in CCD. There exists already some models for the readout noise in CMOS, as for example in,⁸ where the column FPN noise is modeled as a first order isotropic autoregressive process. From our point of view, we have chosen to use a simple model for it. We consider that the noise affecting each column is independent of adjacent columns, and its mean value can be computed according to some parameters that can be chosen in the program or estimated from real digital camera images. We consider two contributions to the readout noise: a column and a pixel random noise, whose variances are σ_C^2 and σ_P^2 , respectively.

In a CCD the total amount of circuit noise is given by

$$\sigma_{CCD}^2 = \sigma_J^2 + \sigma_R^2, \quad (13)$$

where σ_J^2 is the noise introduced by circuits, and σ_R^2 is the read out noise variance. We suppose that for CMOS the total noise is

$$\sigma_{CMOS}^2 = \sigma_J^2 + \sigma_C^2 + \sigma_P^2, \quad (14)$$

thus CCD and CMOS share the same circuit noise, but have different readout noise. If $\sigma_R^2 = \epsilon \cdot \sigma_J^2$, where ϵ is a given constant, then σ_J^2 can be computed from from Eq. 13.

Moreover, the total amount of noise in a PPS scheme is a percentage η_{PPS} of the pixel charge capacity (CC) and η_{APS} for APS that we can estimate or give a certain value. Then:

$$\sigma_C^2 + \sigma_P^2 = \sigma_{CMOS}^2 - \sigma_J^2.$$

The ratio between σ_C^2 and σ_P^2 can be put to a certain value ξ , according to measurements or specifications. Once this ratio is known, it is possible to compute each contribution to CMOS readout noise:

$$\begin{aligned} \sigma_P^2 &= \frac{\sigma_{CMOS}^2 - \sigma_{CCD}^2 / (1 + \epsilon)}{1 + \xi} \\ \sigma_C^2 &= \xi \sigma_P^2 \end{aligned} \quad (15)$$

The column noise is then simulated as a process X_c constant on columns and variable along rows according to a Gaussian r.v. $N(0, \sigma_{X_c}^2)$, while the pixel noise is modeled as a Gaussian r.v. with zero mean and variance σ_P^2 .

5. RESULTS

To test our model, we have used a Sinar professional CCD digital camera to measure the noise content of the images obtained both in absence and presence of the signal. We have considered 4 different exposure times, of 1/30, 1/4, 1, and 4 seconds, respectively. For each of them, we have taken 3 black images at a sensor temperature of 25 degrees Celsius, and estimated the FPN by taking the mean values of the three images at each pixel position. This corresponds to the estimation of the temporal mean we have used in Eq. 7 for the black noise. By using a linear and quadratic MSE fit, we have then computed the constants d_0, d_1 , and b_0, b_1, b_2 of Eq. 8, respectively, and the parameters of our model, according to Eq. 9.

At this point, our model was set, being able to predict the mean value and the variance of the noise at different exposure times and temperatures. To test its performance, we have taken other black images at the same exposure times, but at different temperatures of 5 and 28 degrees Celsius, and compared the simulated noise to

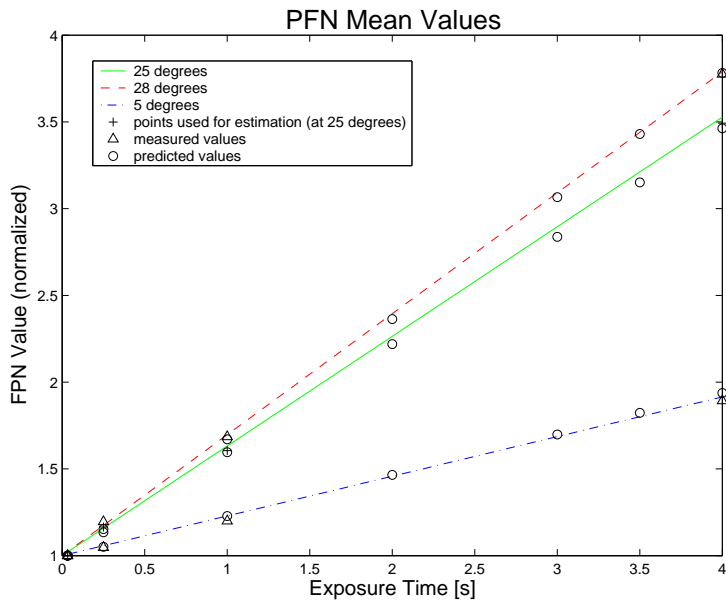


Figure 4. Spatial mean value of the FPN; circles indicate predicted data, triangles indicate measured data, and “+” indicate measured data used to fit the model.

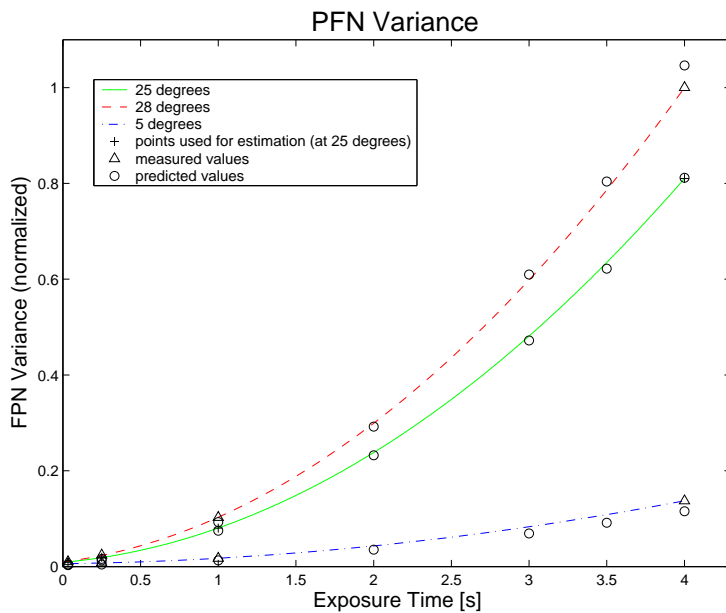


Figure 5. Spatial variance of the FPN; circles indicate predicted data, triangles indicate measured data, and “+” indicate measured data used to fit the model.

the measured one. Fig. 4 and 5 depict the result we have obtained.

We have indicated with a circle the values of noise mean (see Figure 4) and variance (see Figure 5) that are predicted by our model for different exposure times and temperatures. With a triangle Δ , we indicate the data that we have measured, while with a “+” symbol, we indicate the data that has been used to fit our model, i.e., that at 25 degrees. As said before, the fitting of the model consists in imposing that the FPN mean and

variance of the simulated noise coincide with the real data at the specified points. Our model, which was fitted to measurements at 25 degrees, should be consistent to noise measurements at all temperatures, and for each exposure time respectively. This is in fact evident in Figures 4 and 5, where the simulated data points correspond to the measured ones, both at different exposure times and temperatures. This means that our model predicts the noise mean and variance in a coherent way.

A different test has been done to consider real color images. We have taken some images of the MacBeth ColorChecker, which is a chart with 24 different colored patches. The pictures have been taken with the same Sinar digital camera used above. We have used a fixed illuminant for the scene, which has been measured with a spectrophotometer, together with the surface reflectance of each patch. These data have been used to simulate each patch with our model, adding a certain noise content. A correct model should produce simulated patches that share the same noise as the real ones. This means that the noise should be spatially similar and have the same mean and variance of the noise found in real patches. In Fig. 6 we present the results found. We show the ratio between the noise variance of the simulated and real patches obtained from the Sinar digital camera.

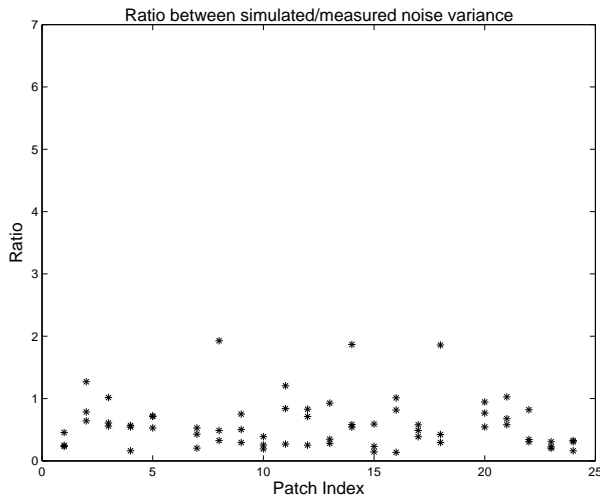


Figure 6. Ratios between the measured and estimated noise variance of the Macbeth ColorChecker; the points indicate the ratio of modeled and measured noise variance.

We notice that the ratio between noise variance in the simulated and real patches is almost constant: a part from a multiplicative constant factor, they have approximately the same behaviour. There are some outliers, but the ratio is between 1 and 2 for most patches. This means that our model overestimates the variance of the noise, but this can be corrected by a multiplicative factor. Generally, we can state that our model is capable of following the variations of noise variance along the different patches. We recall that such variations are present, since the noise is in general signal dependent, i.e., it is stronger for brighter patches. In our model this dependence is taken into account, thus permitting to obtain a coherent result.

6. CONCLUSIONS

We presented a model for digital camera noise. The model aims at simulating the noise content of real digital camera images, by describing the different physical phenomena that contribute to image formation from a statistical point of view.

We showed that it can describe quite accurately the variation of the mean value and variance of the digital images FPN, both according to temperature and exposure time variations. Moreover, in presence of signal, the variance of the simulated noise is comparable with that of a real image.

Our model does have in any case some limitations. First, we did not introduce a dependency on temperature to the readout noise and circuit noise in general, in order to maintain a simple model. Second, the supposition that noise on adjacent pixel is independent is obviously untrue in practice and a more sophisticated statistical

model should be used, such as Gibbs fields, which permits to introduce a certain spatial dependency between pixels. In addition to this, we should also note that the number of parameters that we have introduced can be too high. A study in the direction of minimizing the number of parameter could be of interest. Despite all these considerations, our model has the advantage of being analytic, i.e., it permits to make simple predictions. Our noise model can also be useful for testing and improving existing denoising algorithms. In fact, these algorithms are usually designed to remove Gaussian noise, assuming that this is the noise commonly affecting digital images. The model we propose thus permits denoising algorithms to be tested on more realistic situations and to be optimized for such situations.

APPENDIX A. DERIVATION OF FORMULAS

We consider the Fixed Pattern Noise X_{FPN} and we compute its spatial mean and variance. The spatial mean of X_{FPN} is:

$$E_s[X_{FPN}] = E_s[GF + Z] = \mu_G \cdot \mu_F + \mu_Z = d_0 + d_1 \cdot W, \quad (16)$$

where W is the exposure time, $\mu_F = \alpha W$, and $\mu_G = 1$. Thus we have:

$$\alpha W + \mu_Z = d_0 + d_1 \cdot W \Rightarrow \alpha = d_1, \quad \mu_Z = d_0.$$

The spatial variance is:

$$\begin{aligned} Var_s[X_{FPN}] &= Var_s[GF + Z] = Var_s[GF] + Var_s[Z] = \mu_G^2 \sigma_F^2 + \mu_F^2 \sigma_G^2 + \sigma_G^2 \sigma_F^2 + \sigma_Z^2 \\ &= b_0 + b_1 \cdot W + b_2 \cdot W^2, \end{aligned} \quad (17)$$

where

$$\sigma_G^2 = c^2/12 \quad \text{and} \quad \sigma_F^2 = \beta \cdot W + \gamma \cdot W^2.$$

Thus we have

$$\begin{aligned} \frac{c^2}{12} d_1^2 W^2 + \sigma_F^2 \left(1 + \frac{c^2}{12}\right) + \sigma_Z^2 &= b_0 + b_1 W + b_2 W^2 \Rightarrow \\ \Rightarrow \frac{c^2}{12} d_1^2 W^2 + (\beta W + \gamma W^2) \left(1 + \frac{c^2}{12}\right) + \sigma_Z^2 &= b_0 + b_1 W + b_2 W^2. \end{aligned}$$

From the previous equation, we find the following relationship:

$$\begin{aligned} \sigma_Z^2 &= b_0, \\ \beta \left(1 + \frac{c^2}{12}\right) &= b_1, \\ \frac{c^2}{12} d_1^2 + \gamma \left(1 + \frac{c^2}{12}\right) &= b_2, \end{aligned} \quad (18)$$

and finally we obtain:

$$\beta = b_1 \left(1 + \frac{c^2}{12}\right)^{-1} \quad \text{and} \quad \gamma = (b_2 - \frac{c^2}{12} d_1^2) \left(1 + \frac{c^2}{12}\right)^{-1}. \quad (19)$$

We have then found an expression for the 5 parameters of the model, which we resume here:

$$\alpha = d_1, \quad \beta = b_1 \left(1 + \frac{c^2}{12}\right)^{-1}, \quad \gamma = (b_2 - \frac{c^2}{12} d_1^2) \left(1 + \frac{c^2}{12}\right)^{-1}, \quad \mu_Z = d_0, \quad \sigma_Z = b_0. \quad (20)$$

The value of μ_F and σ_F^2 permit to find the parameter of the r.v. F , since it is distributed as a Gamma r.v., $F \sim \Gamma(\beta_F, \lambda_F)$ for which:

$$\lambda_F = \frac{\sigma_F^2}{\mu_F} \quad \text{and} \quad \beta_F = \frac{\mu_F^2}{\sigma_F^2}. \quad (21)$$

Now we consider the temporal noise X_{temp} . It has zero spatial and temporal mean. Its variance is:

$$\begin{aligned}
Var_t[X_{temp} = X_{Black} - X_{FPN}] &= Var_t[G(P(F) - F)] + 2\sigma_Z^2 \\
&= \underbrace{\mu_G^2 \sigma_{(P(F)-F)}^2}_{=F} + \underbrace{\mu_{(P(F)-F)}^2}_{=0} \sigma_G^2 + \sigma_{(P(F)-F)}^2 \sigma_G^2 + 2\sigma_Z^2 \\
&= (1 + c^2/12)F + 2\sigma_Z^2.
\end{aligned} \tag{22}$$

Analogously, the spatial variance is

$$Var_s[X_{temp} = X_{Black} - X_{FPN}] = Var_s[G(P(F) - F)] + 2\sigma_Z^2 = (1 + c^2/12)\mu_F + 2\sigma_Z^2. \tag{23}$$

We have a term $2\sigma_Z^2$, since the variance of the difference between two independent random variables having the same variance σ^2 is $2\sigma^2$. For noise dependent signals, we have:

$$X_{noise} = GG_{PRNU}P_{oiss}(A) - A. \tag{24}$$

Then

$$\begin{aligned}
E_t[X_{noise}] &= E_t[G]E_t[G_{PRNU}]E_t[P_{oiss}(A) - A] = 1 \cdot 1 \cdot 0 = 0, \\
Var_t[X_{noise}] &= Var_t[GG_{PRNU}P_{oiss}(A) - A] = Var_t[H \cdot P_{oiss}(A)],
\end{aligned} \tag{25}$$

where

$$H = GG_{PRNU} \text{ and } E_t[H] = 1, \quad Var_t[H] = c^2/12 + c^2/12 + (cd)^2/144 \approx c^2/6.$$

Thus

$$Var_t[HP_{oiss}(A)] \approx A + A^2c^2/6 + A^2 = A + A^2(1 + c^2/6). \tag{26}$$

ACKNOWLEDGMENTS

This work was supported by the Swiss National Science Foundation (SNF) under grant number 21-067012.01. We would also like to acknowledge Sinar AG, Switzerland for the digital camera they put at our disposal, and for their support.

REFERENCES

1. B. Wandell, “www.stanford.edu/class/ee392b/handouts/vcam.pdf,”
2. R. Hunt, *The Reproduction of Colour*. Fountain Press England, 1995.
3. G. Wyszecki and W. Stiles, *Color Science: Concepts and Methods, Quantitative Data and Formulae*, ch. 1, pp. 11–12. John Wiley and Sons, 2nd ed., 1982.
4. A. J. Theuwissen, *Solid-State Imaging with Charge-Coupled Devices*. Kluwer Academic Publisher, 1995.
5. W. Feller, *An Introduction to Probability Theory and Its Applications*, vol. I. John Wiley & Sons, third ed., 1968.
6. “www.tau.ac.il/lab3/telescope/dark_current.html,”
7. S. Tanner, *Low-Power Architectures for Single-Chip Digital Image Sensors*. PhD thesis, Université de Neuchâtel, Institut de Microtechnique, November 2000.
8. A. E. Gamal, B. Fowler, H. Min, and X. Liu, “Modeling and estimation of FPN components in CMOS image sensors,” *Proc. of SPIE*, **3301**, pp. 168–177, 1998.