Toward a quantitative visual noise evaluation of sensors and image processing pipes

Clémence Mornet, Donald Baxter, Jérôme Vaillant, Thomas Decroux, Didier Herault, Isabelle Schanen

Toward a Quantitative Visual Noise Evaluation of Sensors and Image Processing Pipes

Clémence Mornet[ac], Donald Baxter[b], Jérôme Vaillant[a], Thomas Decroux[a], Didier Hérault[a] and Isabelle Schanen[b]

[a]STMicroelectronics, Imaging Technology Line, 850 rue Jean Monnet, 38926 Crolles, France;
[b]STMicroelectronics, Imaging Division, 33 Pinkhill, Edinburgh, EH12 7BF, Scotland, UK;
[c]IMEP-LAHC, 3 parvis Louis Néel BP257, 38016 Grenoble Cedex 1, France

ABSTRACT

The evaluation of sensor's performance in terms of signal-to-noise ratio (SNR) is a big challenge for both camera phone manufacturers and customers. The first ones want to predict and assess the performance of their pixel while the seconds require being able to benchmark raw sensors and processing pipes. The Reference SNR metric is very sensitive to crosstalk whereas for low-light issue, the weight of sensitivity should be increased. To evaluate noise on final image, the analytical calculation of SNR on luminance channel has been performed by taking into account noise correlation due to the processing pipe. However, this luminance noise does not match the perception of human eye which is also sensitive to chromatic noise. Alternative metrics have been investigated to find a visual noise metric closer to the human visual system. They have been computed on five pixel technologies nodes with different sensor resolutions and viewing conditions.

Keywords: SNR, visual noise, image quality metric, colour correction, demosaicking

1. INTRODUCTION

As the resolution in camera phones increases with a fixed sensor area, the pixel size is decreasing, which limits the amount of incoming light.[1] As a result, the noise level of the sensor especially in low-light conditions becomes an important issue. This is the reason why the evaluation of CMOS sensor's performance in terms of signal to noise ratio is a big challenge for both CIS suppliers and camera phone manufacturers. The first ones want to predict and assess the performance of their pixel embedded in a sensor while the seconds' requirement is to be able to benchmark raw sensors and processing pipes. The challenge is then to find a consistent manner to evaluate the noise performance of a pixel, a RAW sensor and an imaging system.

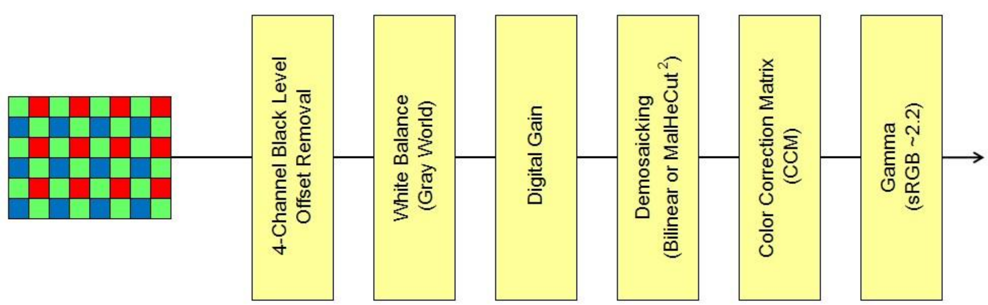


Figure 1: Standard Simple Processing Pipe

In order to simplify this objective, a simple standard processing pipe (Fig. 1) for RAW Bayer images will be used with openly available algorithms. A camera noise metric computed from different pixel technology with the same processing pipe will allow the evaluation of the pixel performance. The demosaicking method could be the bilinear interpolation or the MalHeCut[2] interpolation. An important step of the processing is the colour correction used to map the captured sensor data to the desired output colour space response as it will highly change the noise on final image. The model of full colour reconstruction transformation for a standard RGB sensor can be expressed as:

$$\begin{bmatrix} R \\ G \\ B \end{bmatrix}_{out} = \overbrace{\begin{bmatrix} c_{RR} & c_{GR} & c_{BR} \\ c_{RG} & c_{GG} & c_{BG} \\ c_{RB} & c_{GB} & c_{BB} \end{bmatrix}}^{CCM} \overbrace{\begin{bmatrix} W_R & 0 & 0 \\ 0 & W_G & 0 \\ 0 & 0 & W_B \end{bmatrix}}^{WB} \begin{bmatrix} R \\ G \\ B \end{bmatrix}_{in} \tag{1}$$

$$\Leftrightarrow v_{out} = \boldsymbol{CW} v_{in} = \boldsymbol{CW}(v_{raw} - v_{off})$$

In equation (1), $\boldsymbol{v}_{\text{in}}$ is the raw sensor data with the offsets removed to ensure a correct black level. Next, a white balance (WB) correction denoted $\boldsymbol{W}$ is applied to ensure the neutral tone. The White Balance can be computed either manually from sensitivities on colour channels or with the "gray" world method assuming that the mean of the whole image of a natural scene is gray. Finally the colour correction matrix (CCM) denoted $\boldsymbol{C}$ is applied to compensate the way the image sensor measures the spectral distributions of the incoming light. In order to avoid destroying the white balance established at the previous step, the row sums of this colour correction matrix must be equal to one. To summarize, this equation transforms the raw sensor data $\boldsymbol{v}_{\text{raw}}$ into the output corrected colours $\boldsymbol{v}_{\text{out}}$, assuming a linear relationship between the input and output colour spaces. The colour matrix is computed to produce good colour accuracy with the knowledge of the illuminant and of the coordinates of target points on the Macbeth Colorchecker in the output colour space. Usually, the colour reconstruction corrects the mixing of colours caused by the overlap of colour filters and the crosstalk. This is the reason why the off-diagonal elements of the CCM are most often negative for RGB sensors. Consequently, the diagonal coefficients will be larger than one due to the fact that the row sum must be equal to one, which causes amplification in each colour component. As a signal is amplified or combined with another signal, the noise will increase. Subtracting one signal from another is detrimental: the noise components are added while the signal is diminished, leading to a degradation of the SNR. If little colour correction is needed, the off-diagonal coefficients can stay small, and the noise degradation is limited. The choice of the colour matrix changes the noise and colour accuracy metrics on final image. This is a reason why the processing pipe should be take into account in a noise metric.

In the first part of this study, we will calculate analytically the SNR on the luminance channel for the simple processing pipe described below, the difference with the Reference SNR metric computed on RAW Bayer and the limitation of both metrics. In a second part, an overview of alternative metrics found in literature will be given. Finally, all noise metrics have been computed on four pixel technologies nodes with different sensor resolutions and viewing conditions in order to see the trade-offs between noise metrics.

2. SNR ANALYTICAL CALCULATION

The objective of this section is to calculate analytically the Signal-to-Noise Ratio (SNR) on the luminance channel for a Bayer pattern after a simple processing pipe with demosaicking and colour reconstruction but before gamma correction (Fig. 1). The calculus is possible as long as all steps are linear operations, which is not the case of the gamma correction. The first paragraph will remind the calculation of the Reference SNR metric, then the analytical calculation of the "True" SNR on luminance channel with two linear demosaicking methods will be explained. Finally, we will highlight the limits of both Reference SNR metric and analytical "True" SNR metric.

2.1 Definition of Reference SNR metric on luminance channel

The usual Signal-to-Noise Ratio (SNR) metric called "Reference" SNR is used for the evaluation of a pixel technology on a RAW image. It is defined as the scene illuminance level (in lux) for a SNR of a given value (typically 10) on the luminance channel computed on a 18% gray patch under 3200K illuminant after colour correction, with a f-number of 2.8 and a frame rate of 15fps.[3] In this metric, the luminance channel denoted Y is calculated from sensor RGB channels with the NTSC (National Television Standard Commitee) standard with the white point calculated under illuminant C (Eq. 2).

$$Y = \begin{bmatrix} \beta_R & \beta_G & \beta_B \end{bmatrix} \cdot \begin{bmatrix} R \\ G \\ B \end{bmatrix} = 0.299R + 0.587G + 0.114B \tag{2}$$

The Reference SNR of 10 on luminance channel for instance, denoted Ref SNR_{10} (in lux) is then defined as:

$$\text{Ref SNR}_{10} = xlux \text{ as SNR}(xlux) = \frac{\beta_R R' + \beta_G G' + \beta_B B'}{\sqrt{(\beta_R \sigma_{R'})^2 + (\beta_G \sigma_{G'})^2 + (\beta_B \sigma_{B'})^2}} = 10 \tag{3}$$

With $i = R', G', B'$ and $\sigma_{i'}$ respectively the signal and the noise standard deviation after colour correction on each colour plane. Eq. 3 assumes that noises on colour planes are uncorrelated unless the covariance terms are non-nil. But after colour correction matrix (and demosaicking if the colour planes are merged), noises become correlated. It implies that this metric, used on RAW image, is not representative of the real noise on luminance channel after the processing pipe, denoted "True" SNR calculated on section 2.2. However, we will see in section 4, that even if the calculus is flawed, this metric introduces a mix of luminance and chrominance noise which follows the trend of other alternative noise metrics.

2.2 True Analytical calculation of SNR on luminance channel

The method and the result of the analytical calculation of "True" SNR on the luminance channel Y defined by Eq. 2 will be given in this section for the simple processing pipe described in Fig. 1 and the Bilinear and MalHeCut[2] interpolation. The calculus are detailed in Appendix A for the Bilinear interpolation.

The SNR on Y channel can be expressed as (the terms after colour correction are denoted by a prime):

$$\text{True SNR}_\text{Y} = \frac{Y'}{\sigma_{Y'}} = \frac{\beta_R R' + \beta_G G' + \beta_B B'}{\sqrt{\text{Var}(\beta_R \sigma_{R'} + \beta_G \sigma_{G'} + \beta_B \sigma_{B'})}} \tag{4}$$

where $\beta_R = 0.299$, $\beta_G = 0.587$ and $\beta_B = 0.114$.

and $\text{Var}(\beta_R R' + \beta_G G' + \beta_B B')$

$$= (\beta_R \sigma_R^{'})^2 + (\beta_G \sigma_G^{'})^2 + (\beta_B \sigma_B^{'})^2 + 2\beta_R\beta_G \text{cov}(R', G') + 2\beta_R\beta_B \text{cov}(R', B') + 2\beta_G\beta_B \text{cov}(G', B')$$

Eq. 4 implies that both noises on colour channels and covariance terms have to be calculated. For a Bilinear interpolation and a colour correction as described in Eq. 1, all operations are linear: the covariance terms can then be developed. First, the noises on colour planes after the simple pipe can be expressed as a function of the raw noises on colour planes before corrections and the coefficients of the CCM (Eq. 5).

$$\begin{cases} \sigma_{R'} &= \sqrt{\frac{9}{16} c_{RR}^2 \sigma_R^2 + \frac{5}{16} c_{GR}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BR}^2 \sigma_B^2} \\ \sigma_{G'} &= \sqrt{\frac{9}{16} c_{RG}^2 \sigma_R^2 + \frac{5}{16} c_{GG}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BG}^2 \sigma_B^2} \\ \sigma_{B'} &= \sqrt{\frac{9}{16} c_{RB}^2 \sigma_R^2 + \frac{5}{16} c_{GB}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BB}^2 \sigma_B^2} \end{cases} \tag{5}$$

Then, the noise on luminance channel after colour correction can be calculated too. Noises on the four pixel positions (R,GR,GB,B) are four random variables on four disjunct supports, which means that the noise after interpolation on luminance channel can be written as the quadratical sum of the standard deviations on each pixel position. It follows that:

$$\sigma_{Y'} = \sqrt{\frac{9}{16} \alpha_R^2 \sigma_R^2 + \frac{5}{16} \alpha_G^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} \alpha_B^2 \sigma_B^2} \tag{6}$$

$$\text{With } \begin{cases} \alpha_R &= \beta_R c_{RR} + \beta_G C_{RG} + \beta_B C_{RB} \\ \alpha_G &= \beta_R c_{GR} + \beta_G C_{GG} + \beta_B C_{GB} \\ \alpha_B &= \beta_R c_{BR} + \beta_G C_{BG} + \beta_B C_{BB} \end{cases}$$

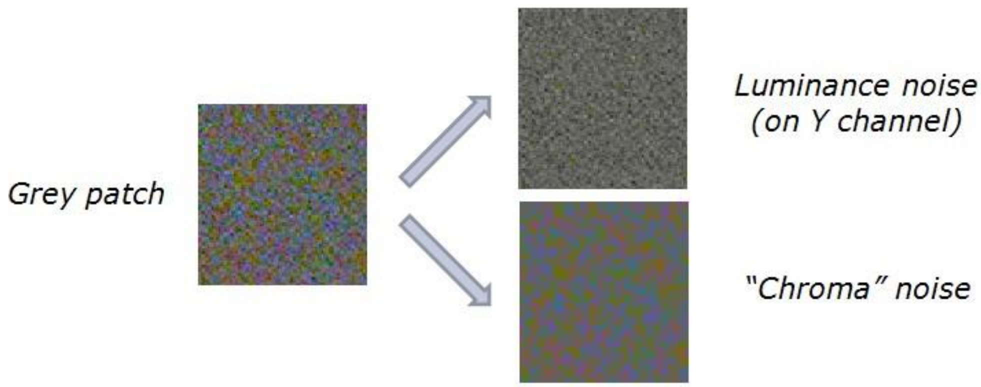


Figure 2: Decomposition of a noisy gray patch into luminance noise and chroma noise.

The same calculation can be done for the MalHeCut interpolation but with more complicated equations since this demosaicking method implies correlation between colour planes. The noise on luminance channel in this case is given by:

$$\begin{aligned}\sigma_{Y'}^2 \quad = \quad & \left(\frac{9\alpha_R^2}{16}+\frac{5\alpha_G^2}{64}+\frac{45\alpha_B^2}{256}+\frac{\alpha_R\alpha_G}{4}+\frac{3\alpha_R\alpha_B}{8}+\frac{15\alpha_R\alpha_G}{64}\right)\sigma_R^2 \\ & +\left(\frac{63\alpha_R^2}{512}+\frac{5\alpha_G^2}{16}+\frac{63\alpha_B^2}{512}+\frac{5\alpha_R\alpha_G}{16}+\frac{27\alpha_R\alpha_B}{128}+\frac{5\alpha_R\alpha_G}{16}\right)(\sigma_{GR}^2+\sigma_{GB}^2) \\ & +\left(\frac{45\alpha_R^2}{256}+\frac{5\alpha_G^2}{64}+\frac{9\alpha_B^2}{16}+\frac{15\alpha_R\alpha_G}{64}+\frac{3\alpha_R\alpha_B}{8}+\frac{5\alpha_R\alpha_G}{4}\right)\sigma_B^2\end{aligned} \tag{7}$$

With the same definition for α_R, α_G and α_B. In this section, we have demonstrated the expression of noise on the luminance channel after the colour processing pipe (but before gamma correction) for the Bilinear and MalHeCut interpolation. The "True" SNR of a gray patch after correction can then be expressed as a function of the White Balance coefficients, CCM and RAW noises on each colour channel.

2.3 The limits of True SNR on luminance channel

However, the evaluation of noise on luminance channel is not sufficient. In fact, the human eye is sensitive to the noise on luminance channel but also to chromatic noise. For instance, in the YUV space given by Rec. ITU-R BT.601 primaries (Eq. 8), it is possible to split the luminance noise in Y channel from the "chroma" noise. Fig. 2 is obtained from a noisy gray patch split into the luminance channel denoted Y and the chromatic channel (U and V).

$$\begin{bmatrix}Y\\U\\V\end{bmatrix}_{\text{Rec.601}} = \begin{bmatrix}0.299 & 0.587 & 0.114\\-0.169 & -0.331 & 0.500\\0.500 & -0.419 & -0.081\end{bmatrix}\begin{bmatrix}R\\G\\B\end{bmatrix} \tag{8}$$

The "True" SNR metric calculated in section 2.2 is not a relevant noise metric as it evaluates the noise only on the luminance channel. An analytical calculation should been done on luminance and chromatic channel once the weights between channels are determined.

The Reference SNR metric, used to assess sensor performance on RAW data, introduces a balance between the noise on luminance channel and chromatic channels. In section 4, we will see that this metric follows the trend of some other alternative noise metrics, even if the balance between luminance and chrominance noise is not necessarily the good one.

A noise metric should take into account human eye sensitivity to noise, but also display properties and viewing conditions.

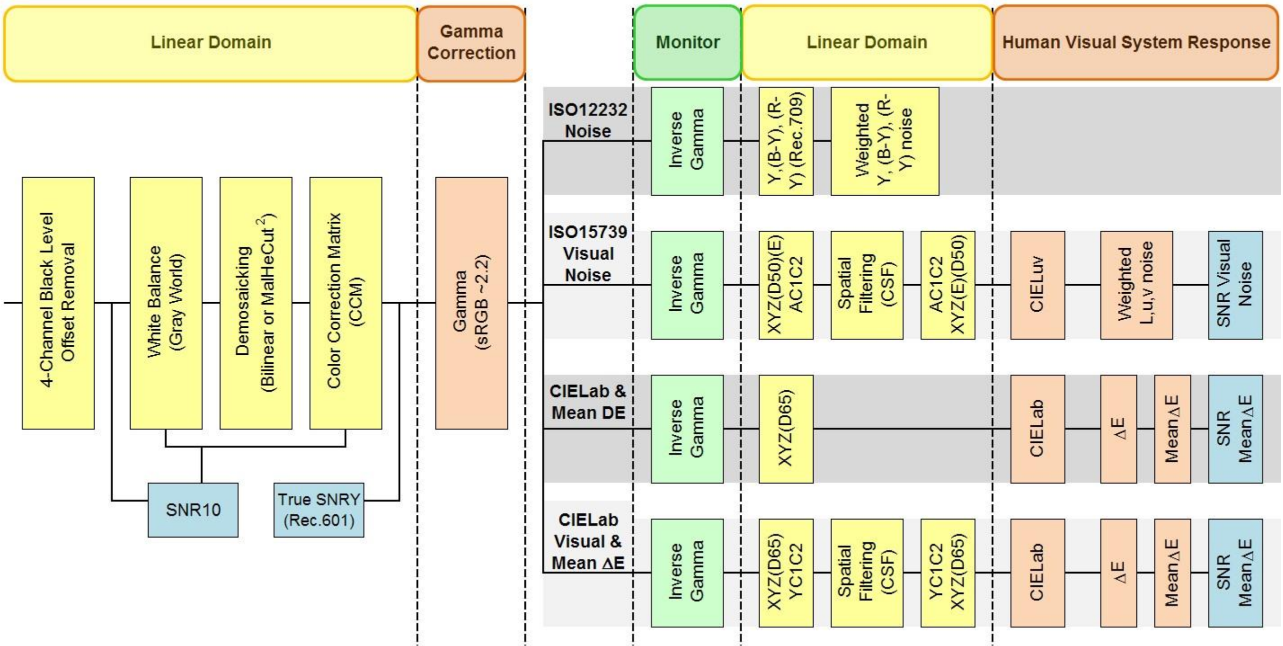


Figure 3: Overview of Alternative Camera Noise Metrics

3. OVERVIEW OF ALTERNATIVE NOISE METRICS

This section presents an overview of the following alternative camera noise metrics (Fig. 3):

- ISO12232 Noise,
- ISO15739 Visual Noise,
- CIELab Noise & Mean ΔE
- CIELab Visual Noise & Mean ΔE.

3.1 Alternative noise metrics

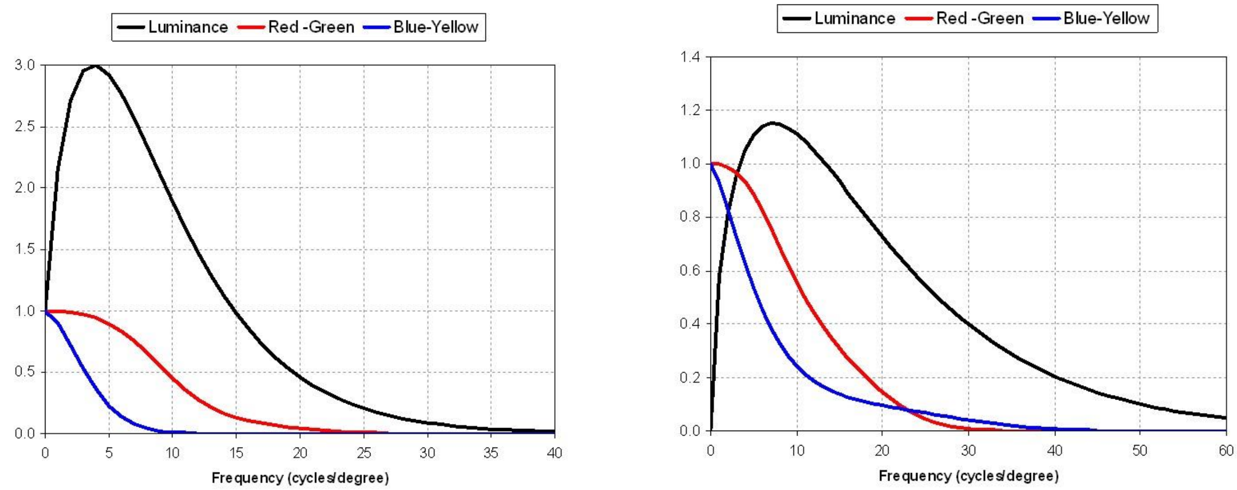
The ISO Noise described in ISO12232:1998[4] and its revision in 2006[5] is defined by weightings on luminance and chrominance noises given in ITU-R BT.709. The weights changed between the 1998 and the 2006 versions of the standard, the Eq. 9 is given in the 2006 version.

$$\text{ISO12232:2006}, \sigma(D) = \sqrt{\sigma_Y^2 + \left(\frac{291}{1000}\sigma_{R-Y}\right)^2 + \left(\frac{88}{1000}\sigma_{B-Y}\right)^2} \tag{9}$$

$$\text{Where } \begin{bmatrix} Y \\ U \\ V \end{bmatrix}_{\text{Rec.709}} = \begin{bmatrix} 0.2126 & 0.7152 & 0.0722 \\ -0.2126 & -0.7152 & 0.9278 \\ 0.7874 & -0.7152 & -0.0722 \end{bmatrix} \begin{bmatrix} R \\ G \\ B \end{bmatrix}$$

The ISO15739[6] adds a Visual Noise metric definition which models the spatial frequency response of the human eye. The noise power spectra are weighted with human spatial responses (*i.e.* CSF or Contrast Sensitivity Function) in the AC_1C_2 space consistent with specific viewing conditions. The AC_1C_2 space is defined from XYZ(E) as:

$$\begin{bmatrix} A \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.0 & 1.0 & 0.0 \\ 1.0 & -1.0 & 0.0 \\ 0.0 & 0.4 & -0.4 \end{bmatrix} \begin{bmatrix} X_E \\ Y_E \\ Z_E \end{bmatrix}$$



(a) (b)

Figure 4: Radial CSF Functions (a) ISO Visual CSF. (b) iCAM CSF.

The transformation between XYZ with white point under E illuminant and sRGB coordinates is given by the transformation between RGB and XYZ with white point under D50 illuminant (with ITU-R BT.709 primaries) and a chromatic adaptation to the illuminant E using the Von Kries adaptation model.

The use of mean of ΔE value as a noise measurement has been discussed both by Johnson and Fairchild[7] and Kleinmann and Wueller.[8] It is also used in the vSNR.[9] The use of a perceptually uniform colour space ensures a consistency with colour accuracy measurements. The Visual CIELab & ΔE is an extension of the previous CIELab & ΔE including the spatial filtering in the $Y'C_1C_2$ space. The transformation from RGB to $Y'C_1C_2$ space is given by the transformation from RGB to XYZ with white point under D65 illuminant with ITU-R

BT.709 primaries and the following transformation given by Johnson:[7]

$$\begin{bmatrix} Y' \\ C_1 \\ C_2 \end{bmatrix} = \begin{bmatrix} 0.0556 & 0.9981 & -0.0254 \\ 0.9510 & -0.9038 & 0.0000 \\ 0.0386 & 1.0822 & -1.0276 \end{bmatrix} \begin{bmatrix} X_{D65} \\ Y_{D65} \\ Z_{D65} \end{bmatrix}$$

3.2 Contrast Sensitivity Functions

Visual metrics (ISO15739 Visual noise and Visual CIELab & ΔE) include a spatial filtering in the processing by the Contrast Sensitivity Functions. Contrast Sensitivity Functions (CSF) model the spatial resolution of human visual system. The CSF is implemented as a set of 3 2-D radial frequency domain filters where the frequency is specified as cycles/degree. The CSF is applied in an opponent colour space which is aligned with human visual system response. The 3 channels in an opponent colour space are luminance, red-green chrominance and blueyellow chrominance. The visual system has different spatial resolutions for luminance, red-green chrominance and blue-yellow chrominance. The resolution is highest for luminance and lowest for blue-yellow chrominance. The number of cycles/degree varies as function of viewing condition (*e.g.* the viewing pixel size relative to the distance the pixel is being viewed at). For a given viewing condition, as the camera resolution increases, the pixel pitch the viewer sees decreases, and thus the number of cycles/degrees increases.

The use of CSF enables the cameras resolution (pixels) to be taken into account in the noise metric.

Johnson and Fairchild have reported CSF they developed for iCAM[10] which differ from the CSF in ISO15739 (Fig. 4). The effect of the chosen CSF can be significant but it is a second order importance in the visual noise metric problematic. However, a subjective work could be required to validate or select the CSF which should be used.

Table 1: Pixel generations

Resolution	Pixel size (µm)	Generation
3MP	1.75	1
5MP	1.75	2
3MP	1.75	3
5MP	1.40	1

3.3 Camera noise metric requirements

The requirement for a camera noise metric should be:

- a simple implementation for a RAW image and a processing pipe;
- the inclusion of both luminance and chrominance noise, which is the case for the 4 alternative noise metrics (at the opposite to the "True"SNR) even if the weight differs for each of them;
- the inclusion of the effects of sensor resolution, display properties and viewing conditions in the noise measurement, which is only the case for the metrics including a spatial filtering *i.e.* the ISO Visual Noise and the CIELab Visual Noise & Mean ΔE.

4. CAMERA NOISE METRICS RESULTS

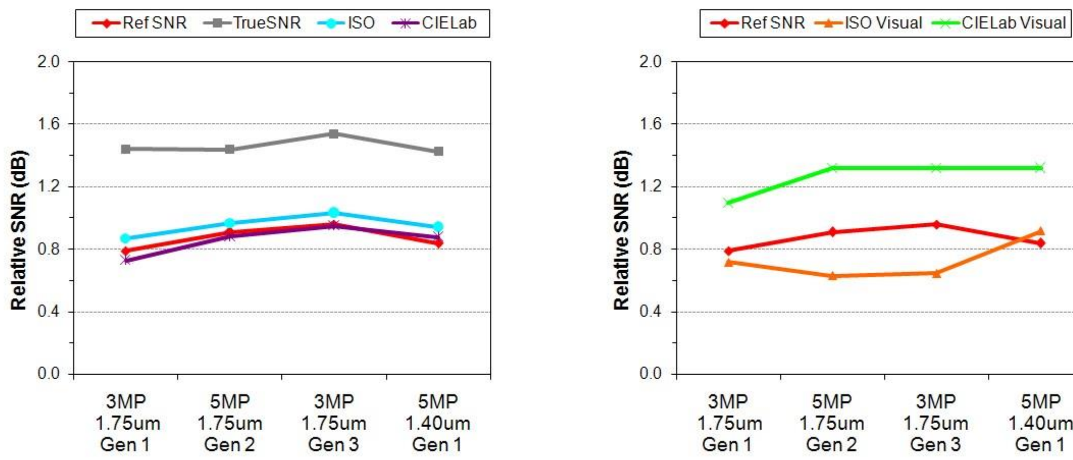
In this section, alternative noise metrics have been evaluated on four pixels technologies nodes with two pixels sizes (1.75µm and 1.4µm), two resolutions (3 and 5MPix) and with different pixel generations (three 1.75µm pixel generations and one 1.4µm pixel generation). A first study is to evaluate the noise metrics described in previous sections on these four cases under 3200K illuminant at medium light level (100 lux) in order to analyse the trends according to each noise metric. Then, we will focus on the trend versus light level and versus minimum illumination. Finally, the benefit of a visual noise metrics will be highlighted in a study of the ISO Visual noise metric on two pixels according to the sensor resolution.

4.1 Noise metrics versus pixel generation

To illustrate the behaviour of the noise metrics, a set of 18% gray patches have been captured from the four pixel technology nodes described in Tab. 1, under 3200K illuminant at medium light level (*i.e.* 100lux). The images were RAW Bayer image and were processed by the standard pipe defined in the previous section. The gray patches used are at the center of the image of a 18% gray scene in order to avoid colour shading effect. The six noise metrics described in section 2 and 3 have been evaluated on image or analytically. They are computed in order to match a SNR metric denoted:

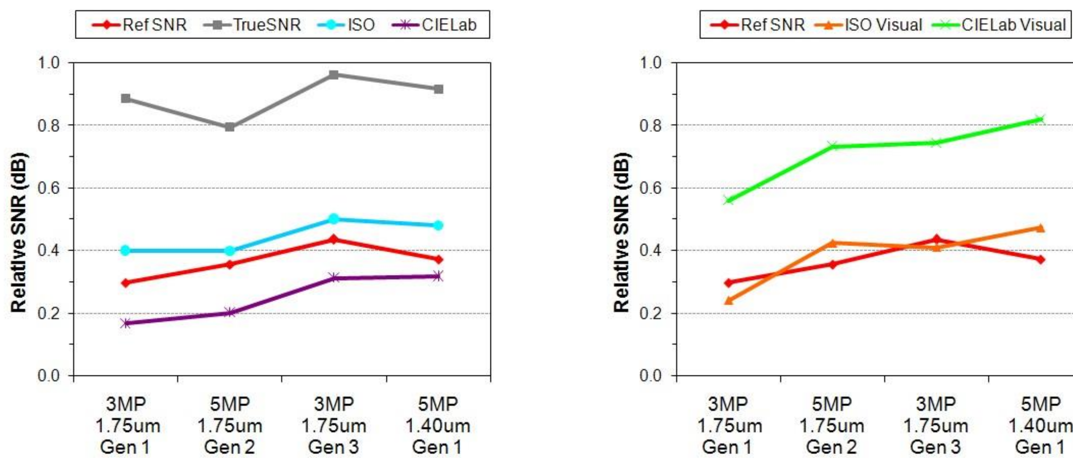
- Reference SNR metric,
- "True"SNR on luminance channel (ITU-R BT.601),
- ISO SNR,
- CIELab SNR, ISO Visual SNR,
- CIELab Visual SNR.

For the ISO and ISO Visual metric, the magnitude of the vector defined by the mean L*,u*,v* data is treated as a signal to convert the L*,u*,v* and visual noise into SNR values. For the CIELab and CIELab Visual metric, the magnitude of the vector defined by the mean of L*,a*,b* data is used to convert the L*,a*,b* and the mean ΔE into SNR.



(a) (b)

Figure 5: Results over pixel generations under 3200K at 100lux for (a) Non-visual metrics. (b) Visual metrics and Reference SNR.



(a) (b)

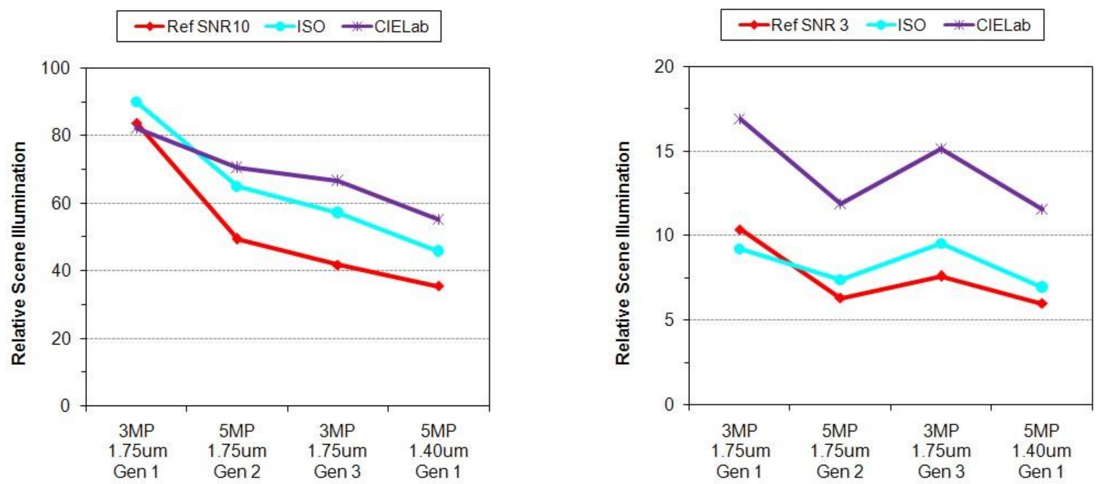
Figure 6: Results over pixel generations under 3200K at 15lux for (a) Non-Visual metrics. (b) Visual metrics.

Fig. 5-a illustrates the results of the non visual noise metrics under 3200K at 100lux whereas Fig. 5-b illustrates the results of the two visual noise metrics with the Reference SNR for the comparison. The SNR values are normalised. Compared to the Reference SNR, true luminance SNR does not follow the same trend whereas other non-visual noises (ISO noise and CIELab noise) exhibit similar trends. However, the trends of visual noise metrics (ISO Visual noise and CIELab noise) are different. For the CIELab Visual noise, 5MPix sensors show an improvement while 3MPix sensor is de-graded. ISO Visual noise doesn't follow the same trend, probably as the CSF used in these two metrics are different. For all these cases, a subjective evaluation is required to validate the numbers.

4.2 Noise metrics versus light level

The same study as in section 4.1 has been computed under 3200K illuminant for a low light condition (15lux). Results are illustrated in Fig. 6 with normalised values.

At low-light level, true luminance SNR still does not follow the Reference SNR order whereas other non-visual noises (ISO noise and CIELab noise) correlate well over the two illumination conditions. However, the results of ISO Visual noise and CIELab Visual noise metrics are consistent with expectations *i.e.* the performance of



(a) (b)

Figure 7: Results over pixel generations under 3200K for (a) SNR 10:1 illumination. (b) SNR 3:1 illumination.

the 5MPix sensors improves and the performance of the 3MPix decreases relative to the 5MPix sensors. We can also see that the trends for some noise metrics are not necessarily the same at 100lux and 15lux. This effect is highlighted in the following section.

4.3 Noise metrics versus minimum illumination

This study analyses the minimum scene illumination to reach a SNR of respectively 10 and 3 on a 18% gray patch under 3200K for a frame rate of $15fps$, a lens transmission of 80% and a f-number of 2.8 for the three metrics: Reference SNR, ISO SNR and CIELab SNR. Results are illustrated in Fig. 7 with normalised values.

For all these cases, a subjective evaluation is required but a camera which performs well at medium light level (typically 100lux) may not necessarily perform well at low-light level (typically 15lux) as other noise sources become more significant in low-light conditions. For instance, generation 2 is weak at low light level whereas it behaves well at medium light level.

4.4 Noise metrics versus resolution

In this section, the ISO Visual noise is computed on two pixel technology nodes *versus* sensor resolution. The ISO Visual noise is computed on gray patches under 3200K illuminant with a professional print viewing condition, *i.e.* a 60x40cm display view at 75cm.

The first pixel technology node is the Generation 3 (Tab. 1) with 1.75µm pixel size, available for 3MPix and 5MPix. The second pixel technology node is the Generation 1 with 1.40µm pixel size, available for 5MPix and 8MPix. Other resolutions are inferred from these available products.

Fig. 8 illustrates the ISO Visual SNR (normalised) versus the camera resolution for these two pixel technologies at 30lux under 3200K illuminant for the professional print viewing conditions. The ISO Visual SNR metric implies the following equivalences: the 3MPix 1.75µm is approximatively equivalent to 5MP 1.40µm and the 5MPix 1.75µm is approximatively equivalent to 8MP 1.40µm.

This results show that the impact of resolution has to be taken into account in the noise evaluation and visual metrics (ISO Visual and CIELab Visual) with spatial filtering are the only ones responding differently according to sensor resolution.

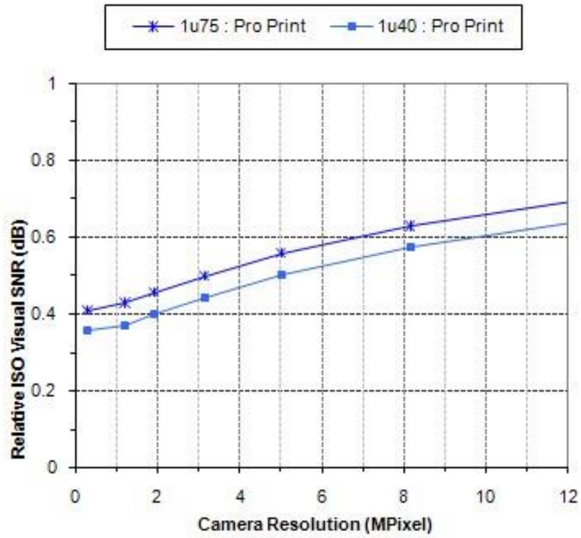


Figure 8: ISO Visual SNR metric versus Camera Resolution for two pixel technology nodes at 30lux under 3200K illuminant

5. CONCLUSION

This paper presents an overview of the problematics of a visual noise metric. The visual noise metric should include both luminance and chrominance noise with weights matching the human eye response but also take into account the effect of sensor resolution and viewing conditions. It also need to be computed either on RAW or final images. On this purpose, a simple standard image pipe has been defined. Results on four pixel generations show that a noise metric should take into account the effect of sensor resolution, scene light level, display properties and viewing conditions. A psychophysical experiment is required to correlate the results of the noise metrics with the preference score. For this experiment, the Image Quality Evaluation Tool (IQE Tool[11]) is useful as it can generate easily simulated images with a lot of different conditions such as resolution, scene illuminance, illuminant, Quantum Efficiency curves and noise.

APPENDIX A. ANALYTICAL CALCULUS OF NOISE ON LUMINANCE CHANNEL

In this section, the analytical calculus of the noise on luminance channel after a Bilinear interpolation and a colour reconstruction will be detailed. The SNR on Y channel can be expressed as (the terms after colour correction are denoted by a prime):

$$True\ \mathrm{SNR_Y} = \frac{Y'}{\sigma_{Y'}} = \frac{\beta_R R' + \beta_G G' + \beta_B B'}{\sqrt{\mathrm{Var}(\beta_R \sigma_{R'} + \beta_G \sigma_{G'} + \beta_B \sigma_{B'})}}$$

where β_R = 0.299, β_G = 0.587 and β_B = 0.114.

and $\mathrm{Var}(\beta_R R' + \beta_G G' + \beta_B B')$

$$= (\beta_R \sigma_{R'})^2 + (\beta_G \sigma_{G'})^2 + (\beta_B \sigma_{B'})^2 + 2\beta_R\beta_G \mathrm{cov}(R', G') + 2\beta_R\beta_B \mathrm{cov}(R', B') + 2\beta_G\beta_B \mathrm{cov}(G', B')$$

For a Bayer pattern as described below, the value for instance of a Red component at a GR pixel position is denoted R@GR.

G_{R1}	R_1	G_{R2}	R_2
B_1	G_{B1}	B_2	G_{B2}
G_{R3}	R_3	G_{R4}	R_4
B_3	G_{B3}	B_4	G_{B4}

To simplify the equations, the R, GR, GB and B signals are given after white balance (which is done before demosaicking and doesn't imply correlation between colour planes). The model of full colour reconstruction (Eq. 1) after white balance gives the following equation:

$$\begin{cases} R' &= c_{RR}R_{interp} + c_{GR}G_{interp} + c_{BR}B_{interp} \\ G' &= c_{RG}R_{interp} + c_{GG}G_{interp} + c_{BG}B_{interp} \\ B' &= c_{RB}R_{interp} + c_{GB}G_{interp} + c_{BB}B_{interp} \end{cases}$$

For the Red channel: $$\begin{cases} R'@R &= c_{RR}R_{interp}@R + c_{GR}G_{interp}@R + c_{BR}B_{interp}@R \\ R'@GR &= c_{RG}R_{interp}@GR + c_{GG}G_{interp}@GR + c_{BG}B_{interp}@GR \\ R'@GB &= c_{RG}R_{interp}@GB + c_{GG}G_{interp}@GB + c_{BG}B_{interp}@GB \\ R'@B &= c_{RB}R_{interp}@B + c_{GB}G_{interp}@B + c_{BB}B_{interp}@B \end{cases}$$

For a Bilinear interpolation: $$\begin{cases} R'@R &= c_{RR}R_3 + \frac{c_{GR}}{4}(G_{R3} + G_{R4} + G_{B1} + G_{B3}) + \frac{c_{BR}}{4}(B_1 + B_2 + B_3 + B\_4) \\ R'@GR &= \frac{c_{RR}}{2}(R_3 + R_4) + c_{GR}G_{R4} + \frac{c_{BR}}{2}(B_2 + B_4) \\ R'@GB &= \frac{c_{RR}}{2}(R_1 + R_3) + c_{GR}G_{B1} + \frac{c_{BR}}{2}(B_1 + B_2) \\ R'@B &= \frac{c_{RR}}{4}(R_1 + R_2 + R_3 + R_4) + \frac{c_{GR}}{4}(G_{R2} + G_{R4} + G_{B1} + G_{B2}) + c_{BR}B_2 \end{cases}$$

$$\begin{cases} \sigma_{R'}@R &= \sqrt{c_{RR}^2\sigma_R^2 + 2\left(\frac{c_{GR}}{4}\right)^2\sigma_{GR}^2 + 2\left(\frac{c_{GR}}{4}\right)^2\sigma_{GB}^2 + 4\left(\frac{c_{BR}}{4}\right)^2\sigma_B^2} \\ \sigma_{R'}@GR &= \sqrt{2\left(\frac{c_{RR}}{2}\right)^2\sigma_R^2 + c_{GR}^2\sigma_{GR}^2 + 2\left(\frac{c_{BR}}{2}\right)^2\sigma_B^2} \\ \sigma_{R'}@GB &= \sqrt{2\left(\frac{c_{RR}}{2}\right)^2\sigma_R^2 + c_{GR}^2\sigma_{GB}^2 + 2\left(\frac{c_{BR}}{2}\right)^2\sigma_B^2} \\ \sigma_{B'} &= \sqrt{4\left(\frac{c_{RR}}{4}\right)^2\sigma_R^2 + 2\left(\frac{c_{GR}}{4}\right)^2\sigma_{GR}^2 + 2\left(\frac{c_{GR}}{4}\right)^2\sigma_{GB}^2 + c_{BR}^2\sigma_B^2} \end{cases}$$

Red noises on the four pixel positions are four random variables on four disjunct supports:

$$\sigma_{R'} = \sqrt{\frac{\sigma_{R'@R}^2 + \sigma_{R'@GR}^2 + \sigma_{R'@GB}^2 + \sigma_{R'@B}^2}{4}}$$

The same calculus is made for the three colours. It follows that:

$$\begin{cases} \sigma_{R'} &= \sqrt{\frac{9}{16} c_{RR}^2 \sigma_R^2 + \frac{5}{16} c_{GR}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BR}^2 \sigma_B^2} \\ \sigma_{G'} &= \sqrt{\frac{9}{16} c_{RG}^2 \sigma_R^2 + \frac{5}{16} c_{GG}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BG}^2 \sigma_B^2} \\ \sigma_{B'} &= \sqrt{\frac{9}{16} c_{RB}^2 \sigma_R^2 + \frac{5}{16} c_{GB}^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16} c_{BB}^2 \sigma_B^2} \end{cases}$$

Luminance channel: Noises on the four pixel positions are four random variables on four disjunct supports, which means that the noise satandard deviation after interpolation on luminance channel can be written as the quadratical sum of the standard deviations on each pixel position.

$$\sigma_{Y'} = \sqrt{\frac{\sigma_{Y'@R}^2 + \sigma_{Y'@GR}^2 + \sigma_{Y'@GB}^2 + \sigma_{Y'@B}^2}{4}}$$

The same method as for colour planes gives:

$$\sigma_{Y'} = \sqrt{\frac{9}{16}\alpha_R^2 \sigma_R^2 + \frac{5}{16}\alpha_G^2 (\sigma_{GR}^2 + \sigma_{GB}^2) + \frac{9}{16}\alpha_B^2 \sigma_B^2} \tag{10}$$

$$\text{With } \begin{cases} \alpha_R &= \beta_R c_{RR} + \beta_G c_{RG} + \beta_B c_{RB} \\ \alpha_G &= \beta_R c_{GR} + \beta_G c_{GG} + \beta_B c_{GB} \\ \alpha_B &= \beta_R c_{BR} + \beta_G c_{BG} + \beta_B c_{BB} \end{cases}$$

The same method can be used for the MalHeCut interpolation.

REFERENCES

[1] Xiao, F., Farell, J., Catrysse, P., and Wandell, B., “Mobile imaging : The big challenge of the small pixel,” *Proc. SPIE* (2009).

[2] Malvar, H. S., He, L.-w., and Cutler, R., “High-quality linear interpolation for demosaicing of bayerpatterned color images,” *IEEE Int. Conf. Acoustics, Speech, Signal Proc.* **3**, 485–488 (2004).

[3] Alakarhu, J., “Image sensors and image quality in mobile phones,” *Int. Image Sens. Workshop* (2007).

[4] ISO12232:1998, “Photography - electronic still-picture cameras - determination of iso speed,” (1998).

[5] ISO12232:2006, “Photography - digital still cameras - determination of exposure index, iso speed ratings, standard output sensitivity, and recommended exposure index,” (2006).

[6] ISO15739:2003, “Photography - electronic still-picture imaging - noise measurements,” (2003).

[7] Johnson, G. and Fairchild, M., “A top down description of s-cielab and ciede2000,” *Color Research & Application* **28**(6), 425–435 (2003).

[8] Kleinmann, J. and Wueller, D., “Investigation of two methods to quantify noise in digital images based on the perception of the human eye,” *Proc. SPIE* **6494** (2007).

[9] Farrell, J., Okincha, M., Parmar, M., and Wandell, B., “Using visible snr (vsnr) to compare the image quality of pixel binning and digital resizing,” *Proc. SPIE* **7537** (2010).

[10] Johnson, G. and Fairchild, M., “Colorimetry - understanding the cie system, chapter 12. image appearance modeling,” 316–317 (2007).

[11] Mornet, C., Vaillant, J., Decroux, T., Herault, D., and Schanen, I., “Evaluation of color error and noise on simulated images,” *Proc. SPIE* **7537** (2010).