Release 4 of the EMVA 1288 Standard: Adaption and Extension to Modern Image Sensors

Bernd Jähne

Heidelberg University, HCI at IWR
Berliner Straße 43, 69120 Heidelberg

Abstract The well established and worldwide used EMVA Standard 1288 for objective camera characterization is still limited to linear monochrome or color cameras without preprocessing. This paper previews the upcoming Release 4.0 which can characterize a much wider range of imaging sensors. This includes sensors with an extended spectral range — especially into the short-wave infrared (SWIR) —, multispectral sensors with more than three color channels, polarization sensors, time-of-flight sensors, high-dynamic range image sensors and any other sensor with a non-linear characteristic curve, and sensors with preprocessing in the camera in order to optimize image quality.

Keywords Image sensor, cameras, standards, EMVA 1288

1 Introduction

The standard 1288 of the European Machine Vision Association (EMVA) is used worldwide for objective characterization of the quality parameters for industrial cameras [1–5]. It is the oldest standard activity of the EMVA. The standard has been elaborated by a consortium of the industry leading sensor and camera manufacturers, distributors, and research institutes. Work on the 1288 standard started in February 2004. A first version was published in 2005 [6] and the current release 3.1 went into effect end of 2016 [7]. This release can only be applied to cameras with a linear characteristic

DOI: 10.58895/ksp/1000124383-2 erschienen in:
Forum Bildverarbeitung 2020
DOI: 10.5445/KSP/1000124383 | https://www.ksp.kit.edu/site/books/m/10.58895/ksp/1000124383/

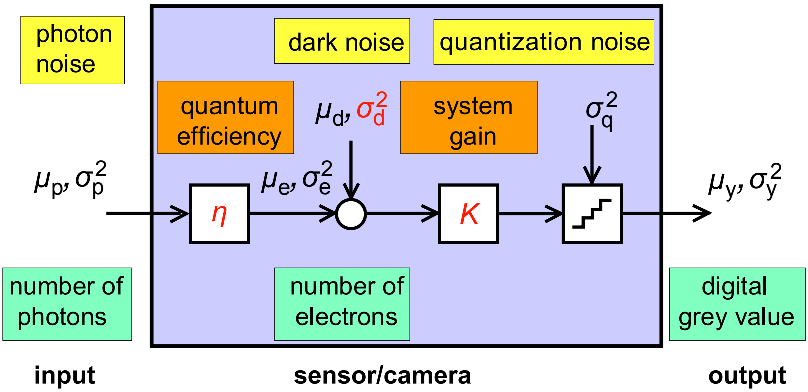

Figure 2.1: Linear model of a camera according to the EMVA 1288 standard.

curve. Furthermore, no preprocessing was possible which changes the temporal noise, except for simple operations such as binning or time-delayed-integration (TDI).

2 Linear model

The standard 1288 is generally based on a system theoretical concept which requires no measurements from within a camera. It is sufficient to measure the input signal, the mean number of photons μ_p hitting each pixel during the exposure time with a variance $\sigma_p^2 = \mu_p$ (Poisson process), and the output signal, the digital signal y (units DN) with mean μ_y and variance σ_y^2. No other measurements are required. With the current release 3.1 [7] a linear camera model is used with three unknown parameters (Fig. 2.1): the variance of the temporal dark noise σ_d^2 — subsuming *all* noise sources within the camera except for the quantization noise σ_q^2 —, the quantum efficiency η and the system gain K.

These three parameters can be determined from an irradiation series covering the whole range from dark to saturation measuring the linear characteristic curve and the linear photon transfer curve (temporal noise variance versus mean of the digital camera signal) [7,8]:

$$\begin{aligned} &\text{Characteristic curve:} && \mu_y = \mu_{y.\text{dark}} && + K\eta\,\mu_p \\ &\text{Photon transfer curve:} && \sigma_y^2 = K^2\sigma_d^2 + \sigma_q^2 && + K(\mu_y - \mu_{y.\text{dark}}) \end{aligned} \tag{2.1}$$

The most important quality parameter of any measuring system it the signal-to-noise ratio SNR. From this quantity most application-oriented camera parameters such as the absolute sensitivity threshold, the dynamic range, and the maximum SNR can be derived [7]. For a linear system the input and output SNR are equal and can be computed from (2.1) resulting in

$$\text{SNR}(\mu_p) = \frac{\mu_y}{\sigma_y} = \frac{\eta\mu_p}{\sqrt{\sigma_d^2 + \sigma_q^2/K^2 + \eta\mu_p}} \tag{2.2}$$

Except for the influence of the quantization noise and provided that the temporal dark noise σ_d^2 does not depend on the system K, the SNR is — as expected for a linear system — independent of the system gain K.

The definition (2.2) does not yet include the signal degradation by spatial variations from pixel to pixel, which can also be described by a variance. For a linear system each pixel can have a different offset (dark signal nonuniformity DSNU) and slope (photo response nonuniformity PRNU). Therefore the spatial variance s_y^2 in units e^- can be expressed by

$$s_y^2 = \text{DSNU}_{1288}^2 + \text{PRNU}_{1288}^2\left(\eta\mu_p\right)^2 \tag{2.3}$$

This spatial variances can be added to the temporal variances resulting in the total SNR

$$\text{SNR}_{\text{total}}(\mu_p) = \frac{\eta\mu_p}{\sqrt{\sigma_d^2 + \text{DSNU}_{1288}^2 + \sigma_q^2/K^2 + \eta\mu_p + \text{PRNU}_{1288}^2\left(\eta\mu_p\right)^2}}. \tag{2.4}$$

An interesting aspect of this approach is that the performance of a real sensor can directly be compared with an ideal (the best possible)

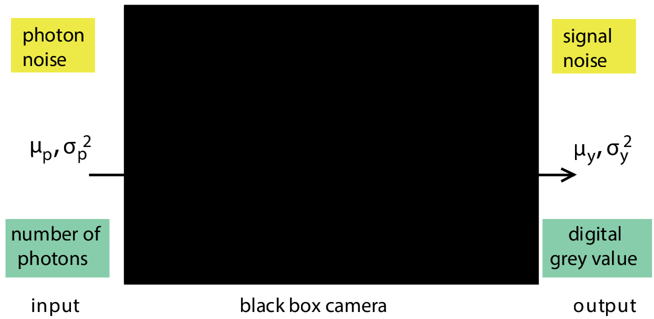

Figure 3.1: General black-box model of a camera according to the EMVA 1288 standard.

sensor. For an ideal sensor, the temporal dark noise, the quantization noise, DSNU and PRNU are zero and the quantum efficiency is one. Then (2.4) reduces to

$$\mathrm{SNR}_{\mathrm{ideal}}(\mu_p) = \sqrt{\mu_p} \tag{2.5}$$

3 General black-box model

For a camera with a non-linear characteristic curve, the linear model of EMVA 1288 release 3.1 cannot be applied. However, a camera with an arbitrary non-linear characteristic curve or a camera with preprocessing modifying the noise characteristics can be characterized by a true black-box model without *any* assumptions (Fig. 3.1). Even with this relaxed assumptions, the output SNR can be computed directly from the mean digital output signal and its temporal variance. It is also still possible to measure the characteristic curve $\mu_y(\mu_p)$ because it is the direct relation between the mean input and output signals.

For a general system the input SNR is different from the output SNR. The input SNR is the really important parameter for an image sensor. It gives the certainty with which the pixel irradiance can be measured. It is possible to compute the input SNR from the output

SNR because these two quantities are related to each other by the slope of the characteristic curve (laws of error propagation, [8]):

$$\mathrm{SNR}_{\mathrm{in}} = \frac{\mu_p}{\sigma_p} = \frac{\mu_p}{\sigma_y}\frac{\partial\mu_y}{\partial\mu_p} = \frac{\mu_p}{\mu_y}\frac{\partial\mu_y}{\partial\mu_p}\mathrm{SNR}_{\mathrm{out}} \tag{3.1}$$

It is important to note that the standard deviation σ_p does not only include the temporal noise of the incoming stream of photons (shot noise) but also all other noise sources within the non-linear camera — back-projected to the input signal.

It is also easy to specify the input SNR for an ideal general image sensor. Then there are no other noise sources and only the photon noise remains. Therefore the ideal input SNR is given — as for a linear camera (2.5) — by

$$\mathrm{SNR}_{\mathrm{in.ideal}}(\mu_p) = \sqrt{\mu_p}. \tag{3.2}$$

In this way, it is possible to specify how much worse a real camera (3.1) is in comparison with an ideal one (3.2) also in the case of a general true black-box model. Without a more detailed camera model, it is not possible to determine the quantum efficiency[1] of the sensor. However, this is not a significant disadvantage. As with a linear camera (Sect. 2) the camera performance parameters really of importance for applications such as the absolute sensitivity threshold, the dynamic range, and the maximum SNR can be derived from the input SNR *without* knowing the quantum efficiency.

4 Fast and more detailed nonuniformity characterization

In order to analyze the spatial patterns of nonuniformities by the rich set of tools from the EMVA 1288 standard such as profiles, histograms and spectrograms [7], it is required to suppress the temporal noise. Because the spatial and temporal variances are roughly of the same order of magnitude, this requires averaging over hundreds of images. This is no real problem for a linear camera because it is

[1] The quantum efficiency relative to a maximum response can still be measured by performing measurements over the whole range of wavelengths.

sufficient to analyze the nonuniformities with just two parameters, the DSNU and PRNU (Sect. 2). Thus averaging over many images is only required for the dark images and images at 50% saturation [7].

With the general black-box model (Sect. 3), averaging at just two irradiation levels is generally not sufficient. The best would be to estimate the spatial nonuniformity at all irradiation levels where the temporal noise is measured. These are at least 50 levels [7]. Therefore this approach is not feasible. Thus the question arises whether it is possible to determine at least some significant parameters of the spatial nonuniformity with much fewer images.

4.1 Temporal and spatial variances from just two images

In the following, a new approach is detailed which works with much fewer images — as few as two images are sufficient — and still provides a detailed statistical analysis of the spatial nonunifomities.

The starting point is the observation that the stationary nonuniformities can entirely be eliminated by computing the temporal noise from the difference of two images taken with the same irradiation. This is the approach taken in the EMVA standard 1288 to compute the variance of the temporal noise [7]. The mean from two images $\boldsymbol{y}[0]$ and $\boldsymbol{y}[1]$ is

$$\mu = \frac{1}{2NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (y[0][m][n] + y[1][m][n]) \tag{4.1}$$

and the temporal variance computed from the difference image is

$$\sigma^2 = \frac{1}{2NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (y[0][m][n] - y[1][m][n])^2. \tag{4.2}$$

The variance computed in this way must be divided by two because the variance of the difference image is two times higher than the variance of a single image.

The key point is now that single images contains both the temporal noise and the spatial nonuniformity. In this way, the spatial nonuniformity can be computed by subtraction. However, it must be ensured that the subtraction never results in negative variances. In the following, it is shown under which conditions this is possible.

In order to simply the equations, the following abbreviations are introduced for the mean value and the variance of image $\boldsymbol{y}[l]$[2]

$$
\begin{aligned}
\mu[l] &= \frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} y[l][m][n] = \overline{\boldsymbol{y}[l]} \\
\sigma^2[l] &= \frac{1}{NM} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} (y[l][m][n] - \mu_y[l])^2 = \overline{(\boldsymbol{y}[l] - \mu_l)^2}
\end{aligned}
\tag{4.3}
$$

In the ideal case mean values of all images at the same irradiation level have the same mean value, but they may be different and it is required to check whether this effect can result in negative variances.

The image is assumed to be composed of a mean value $\mu[l]$, differing from image to image, a zero-mean temporal noise signal $\boldsymbol{n}[l]$ with a variance σ^2 and a zero-mean and stationary spatially varying signal $\boldsymbol{s}$ with a variance s^2:

$$
\boldsymbol{y}[l] = \mu_l + \boldsymbol{n}[l] + \boldsymbol{s}, \quad \overline{\boldsymbol{y}[l]} = \mu_l, \quad \overline{(\boldsymbol{y}[l] - \mu_l)^2} = \sigma^2 + s^2. \tag{4.4}
$$

Furthermore it is assumed that the temporal noise from different images and the temporal noise and spatial nonuniformities are statistically independent, i. e. $\overline{\boldsymbol{n}[l]\boldsymbol{n}[k]} = 0$ (if $k \neq l$) and $\overline{\boldsymbol{n}[l]\boldsymbol{s}} = 0$.

Now two terms are evaluated. Firstly, the temporal variance from the difference image (4.2) needs to be corrected for possible different mean values of the two images

$$
\begin{aligned}
A &= \overline{[(\boldsymbol{y}[0] - \mu_0) - (\boldsymbol{y}[1] - \mu_1)]^2} \\
&= \overline{\boldsymbol{y}^2[0]} + \overline{\boldsymbol{y}^2[1]} - 2\overline{\boldsymbol{y}[0]\boldsymbol{y}[1]} - (\mu_0 - \mu_1)^2 \stackrel{!}{=} 2\sigma_y^2
\end{aligned}
\tag{4.5}
$$

Secondly, the variances of the two images are added up, which include both the variances of the temporal noise and the spatial nonuniformity:

$$
\begin{aligned}
B &= \overline{[(\boldsymbol{y}[0] - \mu_0)^2} + \overline{[(\boldsymbol{y}[1] - \mu_1)^2} \\
&= \overline{\boldsymbol{y}^2[0]} + \overline{\boldsymbol{y}^2[1]} - \mu_0^2 - \mu_1^2 \stackrel{!}{=} 2\sigma^2 + 2s^2
\end{aligned}
\tag{4.6}
$$

[2] Both sums are divided by the total number of pixels NM, although for a bias-free estimate of the variance the divisor should be one less ($NM - 1$). This approach is necessary to have the same averaging scheme for means and variances. The error introduced is very small and can even be corrected at the end by multipling the estimated variances with $NM/(NM - 1)$.

The difference of the two terms should be equal to s^2 and therefore always be positive:

$$B - A = 2\overline{\boldsymbol{y}[0]\boldsymbol{y}[1]} - 2\mu_0\mu_1 \stackrel{!}{=} 2s^2. \tag{4.7}$$

This can be verified by inserting the image signal (4.4) into (4.7).

It is essential to include the possibly slightly different mean values of the two images in term A (4.5). If this is not done, the term $B - A$ could become negative and too high temporal variances and too low spatial variances are computed:

$$2\overline{\boldsymbol{y}[0]\boldsymbol{y}[1]} - \mu_0^2 - \mu_1^2 = 2s^2 - (\mu_0 - \mu_1)^2 \neq 2s^2. \tag{4.8}$$

4.2 Split into row, column, and pixel nonuniformities

Modern CMOS sensors may exhibit not only pixel-to-pixel nonuniformities, but also row-to-row and/or column-to-column nonuniformities. Therefore it is important to decompose the spatial variance into row, column, and pixel variances:

$$s^2 = s^2_{\text{row}} + s^2_{\text{col}} + s^2_{\text{pixel}}. \tag{4.9}$$

All three unknowns can still be estimated by computing additional spatial variances from a rows and columns averaged over the whole image. The mean row and column of a single image are given by

$$\mu[n] = \frac{1}{M}\sum_{m=0}^{M-1} y[m][n], \quad \mu[m] = \frac{1}{N}\sum_{n=0}^{N-1} y[m][n]. \tag{4.10}$$

The column spatial variance computed from the average row

$$s^2_{\text{col}} = \frac{1}{N-1}\sum_{n=0}^{N-1}(\mu[n] - \mu)^2 \quad - \quad s^2_{\text{row}}/N - s^2_{\text{pixel}}/N - \sigma^2/(N) \tag{4.11}$$

still contains a residual row spatial variance, pixel spatial variance and temporal variance. Averaging over N rows does not completely suppress theses variances. Therefore the three terms on the right hand need to be subtracted. Likewise, the *row spatial variance* computed from the average column

$$s^2_{\text{row}} = \frac{1}{M-1}\sum_{m=0}^{M-1}(\mu[m]-\mu)^2 \quad - \quad s^2_{\text{col}}/M - s^2_{\text{pixel}}/M - \sigma^2/(M) \tag{4.12}$$

contains residual column spatial variance, pixel spatial variance and temporal variance.

The three equations (4.9), (4.11), and (4.12) form a linear equation system from which all three components of the spatial variance can be computed. With the two abbreviations

$$\begin{aligned} s^2_{\text{cav}} &= \frac{1}{N-1}\sum_{n=0}^{N-1}(\mu[n]-\mu)^2 \quad - \sigma^2/(N), \\ s^2_{\text{rav}} &= \frac{1}{M-1}\sum_{m=0}^{M-1}(\mu[m]-\mu)^2 - \sigma^2/(M), \end{aligned} \tag{4.13}$$

the linear equation system reduces to

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 1/M & 1/M \\ 1/N & 1 & 1/N \end{bmatrix} \begin{bmatrix} s^2_{\text{col}} \\ s^2_{\text{row}} \\ s^2_{\text{pixel}} \end{bmatrix} = \begin{bmatrix} s^2 \\ s^2_{\text{rav}} \\ s^2_{\text{cav}} \end{bmatrix}. \tag{4.14}$$

The solution of this linear equation systems is

$$\begin{aligned} s^2_{\text{col}} &= \frac{M}{M-1}s^2_{y.\text{rav}} - \frac{1}{M-1}s^2, \\ s^2_{\text{row}} &= \frac{N}{N-1}s^2_{y.\text{cav}} - \frac{1}{N-1}s^2, \\ s^2_{\text{pixel}} &= \frac{MN-1}{(M-1)(N-1)}s^2 - \frac{N}{N-1}s^2_{\text{cav}} - \frac{M}{M-1}s^2_{\text{rav}}. \end{aligned} \tag{4.15}$$

Negative variances cannot result from this split up of the variances, because they are calculated from single images. Therefore changes in the mean values from image to image do not influence the computation. However it is important to avoid numerical rounding errors by choosing an appropriate high-accuracy arithmetic and suitable algorithms.

4.3 Variances of temporal noise and nonuniformity; signal stability

The two new schemes detailed in the previous two sections enable the computation of the variances of both the temporal noise and the spatial nonuniformity from just two images at one irradiation level. It is even possible to split the latter into variations from row to row, column to column, and pixel to pixel.

The difference in the mean values between two images at the same irradiation level carries also an important additional information, namely how stable the irradiation measurement is from image to image.

For any camera the spatial nonuniformity is analyzed at all irradiation levels. For a camera with an arbitrary non-linear characteristic curve, then the most critical irradiation levels can be chosen from these measurements, where it is useful to average over hundreds of images to apply further tool of the EMVA standard 1288 such as profiles, histograms and spectrograms [7] for a more detailed analysis of the spatial patterns.

5 Further comprehensive extensions

In order to cope with modern image sensors release 4.0 includes many further extensions. In this section the most important of them are briefly described.

- The wavelength range is extended from deep UV to SWIR. In the deep UV, when more than one charge unit is produced by a single photon, the simple linear model can no longer be used even for a linear sensor, because a new noise source arises.

- With the general model also image intensifiers including emCCDs can be characterized.

- Raw data of *any* image modality can be characterized according to the standard

- Characterization of the polarisation angle and degree of polarization of a polarization image sensor is an example for the characterization of parameters derived from multiple channels.

The rich set of tool of the standard can also be applied to such parameters.

- Optionally, cameras with lenses or an illumination corresponding to a given exit pupil can be measured. In this way it is possible to measure also image sensors with micro lenses that are shifted towards the edge of the sensor.
- The new version includes a better measure for the linearity of the characteristic curve than in release 3.1. Because the slope of the characteristic curve is evaluated according to the general model (Sect. 3), also the differential non-linearity is known.

6 Conclusions

The new Release 4.0 adequately considers the rapid progress of imaging sensors. It will be possible to characterize a much wider spectrum of cameras/sensors: UV and SWIR-sensitive, multispectral, polarization, intensified (such as EM-CCDs), multilinear and highdynamic range. Also, cameras with lenses and preprocessing to enhance the image quality can be characterized. Despite the diversity, the quality of cameras can still be described with a minimum set of application-oriented quality parameters. It is planned to publish a release candidate of Release 4.0 in the fall of 2020.

The rich tool set of the EMVA 1288 standard to characterize temporal noise, nonuniformity and defect pixels can also be applied to any parameters derived from several channels of a multimodal image sensor. As a prime example the analysis of the degree of polarization and polarization angle computed from a polarization image sensors is contained.

The new release 4.0 does not yet cover an entirely different class of image sensors, so-called event-based or neuromorphic sensors. Research to extend the EMVA standard 1288 also for this class of sensors has already started [9].

7 Acknowledgments

The author gratefully acknowledges financial support for this research through his senior professorship, jointly funded by the Rector of Heidelberg University, HCI and IWR. The discussions within the EMVA 1288 working group were also very helpful in developing the new general model for cameras with an arbitrary characteristic curve and/or preprocessing.

References

1. A. Darmont, "Using the EMVA 1288 standard to select an image sensor or camera," in *Sensors, Cameras, and Systems for Industrial/Scientific Applications XI*, ser. Proc. SPIE, E. Bodegom and V. Nguyen, Eds., vol. 7536, 2010, p. 753609.
2. B. Jähne, "EMVA 1288 standard for machine vision – objective specification of vital camera data," *Optik & Photonik*, vol. 5, pp. 53–54, 2010.
3. A. Darmont, J. Chahiba, J. F. Lemaitre, M. Pirson, and D. Dethier, "Implementing and using the EMVA1288 standard," in *Sensors, Cameras, and Systems for Industrial/Scientific Applications XIII*, ser. Proc. SPIE, R. Widenhorn, V. Nguyen, and A. Dupret, Eds., vol. 8298, 2012, p. 82980H.
4. M. Rosenberger, C. Zhang, P. Votyakov, M. Preißler, R. Celestre, and G. Notni, "EMVA 1288 camera characterisation and the influences of radiometric camera characteristics on geometric measurements," *Acta IMEKO*, vol. 5, pp. 81–87, 2016.
5. A. Darmont, *High Dynamic Range Imaging: Sensors and Architectures*, 2nd ed. SPIE, 2019.
6. EMVA 1288 Working Group, "EMVA Standard 1288 - standard for characterization of image sensors and cameras, release A1.00," European Machine Vision Association, open standard, 2005.
7. ——, "EMVA Standard 1288 - standard for characterization of image sensors and cameras, release 3.1," European Machine Vision Association, open standard, 2016.
8. B. Jähne, *Digitale Bildverarbeitung und Bildgewinnung*, 7th ed. Berlin: Springer Vieweg, 2012.
9. A. Manakov and B. Jähne, "Characterization of event-based image sensors in extent of the EMVA 1288 standard," in *this volume*, 2020.