PROCEEDINGS OF SPIE

Specifying radiometric performance of hyperspectral and conventional cameras: a minimal set of independent characteristics

Skauli, Torbjørn

SPIE.

Specifying radiometric performance of hyperspectral and conventional cameras – a minimal set of independent characteristics

Torbjørn Skauli*[ab]

[a] University of Oslo, Department of Technology Systems, Gunnar Randers vei 19, 2007 Kjeller, Norway; [b] Norwegian Defence Research Establishment (FFI), Instituttveien 20, 2007 Kjeller, Norway

ABSTRACT

The radiometric performance of cameras is customarily characterized and specified in terms of component properties such as F-number and readout noise. This paper considers the camera as an integrated unit and derives characteristics suitable for specifying noise, throughput and saturation as determined from the overall input-to-output performance. A conventional model of signal and noise is reformulated into a simpler "equivalent camera" model with the same radiometric performance, constrained to have a lossless lens with detector-side pupil subtending 1 steradian, and a detector with a peak quantum efficiency (QE) of 1. The small parameter set of this model can then be determined with the camera treated as a "black box", relevant for verification of camera specifications. The net light collection of the real camera is expressed by

the detector area of the equivalent camera, denoted A^*, as well as the wavelength dependence of its QE, denoted $\lambda^* \eta$ ( ).

The noise floor due to readout noise can be expressed for a particular camera as a noise equivalent spectral radiance (NESR). For comparison of cameras with different bandwidths, it is shown that a comparative figure of merit, which is also independent of integration time, is the "noise equivalent radiance dose" (NERD). For a hyperspectral camera, the model parameters can be determined with a simple broadband source, while cameras with broad spectral response require measurements with tunable monochromatic light. The treatment also applies to spectrometers. Reference is made to D^*, a well-established figure of merit for detectors, and it is argued that A^*, $\lambda^* \eta($ $)$ and NERD are analogous figures of merit

for camera properties.

Keywords: Aperture, Calibration, Detectivity, Étendue, F-number, Hyperspectral, Multispectral, NEP, NESR, Noise, Radiometry, Remote sensing, Saturation, Signal model, Specification, Spectrometer, T-number, Throughput

1. INTRODUCTION

The radiometric performance of cameras is well understood, and can be accurately modeled based on the properties of the camera components. However, the customary ways in which cameras are specified do not necessarily contain full information needed in a model of radiometric performance. For conventional photographic cameras, full technical performance characteristics are not provided by the manufacturers, and must be sought on third-party sites [1]. In the field of machine vision, the EMVA1288 standard [2] specifies a robust way to express the performance of the image sensor part of the camera. Hyperspectral imaging differs markedly from conventional imaging in the need for radiometrically accurate image data, particularly in remote sensing applications. Despite these requirements, the current way of specifying hyperspectral cameras leaves much to be desired [3]. These shortcomings are being addressed by the IEEE P4001 working group for standardization in hyperspectral imaging. This paper arises from texts that have been submitted to the P4001 group for discussions which are still ongoing at the time of writing. The opinions here are those of the author.

A basic premise for any performance specification is that it should be testable by observing the actual performance. Radiometric performance of cameras is traditionally quantified in terms of a large number of characteristics such as F-number, quantum efficiency (QE), optics transmission, readout noise, bit width, well capacity, and more. Many such characteristics can only be properly tested on the component level. If a camera performs internal data processing to output radiance images, as illustrated in Figure 1, it is desirable to be able to characterize the whole camera as a "black box", at least for performance verification. The same is true to a large degree for conventional photographic cameras. The camera characteristics should therefore describe properties observable "outside the box", which can be verified without knowledge

*torbjorn.skauli@its.uio.no

of its internal components. Another desirable property of a performance specification is that the specified quantities should facilitate comparison between different camera types, as well as comparison to application requirements. Commonly reported characteristics such as readout noise or F-number carry important information about camera quality, but do not carry full information for comparison purposes. Hence the allusion to "thinking outside the box" above, since the current conventions appear to fall short of what is needed for a full specification of camera radiometric performance, indicating a need for new characteristics beyond current conventions.

An interesting reference case can be found among the characteristics used to describe the performance of detectors. The detector noise floor can be characterized by the noise equivalent power (NEP). This quantity is easy to interpret, but not a good figure of merit: For a given detector, the NEP will typically vary proportionally to the square root of signal bandwidth Δf (assuming white noise). Also, the NEP tends to vary proportionally to the square root of detector area A_{det}, which scales the noise-generating volume. A useful figure of merit for detectors is therefore the *specific detectivity* [4].

$$D^* = \frac{\sqrt{A_{\mathrm{det}}\Delta f}}{NEP}.$$

This quantity is comparable between detectors with different areas, and between applications with different signal bandwidths. With its customary units of $\mathrm{cm}\sqrt{\mathrm{Hz}}$ / W (termed Jones), D^* expresses the signal to noise ratio of a detector with area 1 cm^2 receiving a light signal with power of 1 watt in a bandwidth of 1 Hz. Such a case is unlikely to occur in practice, but D^* is nevertheless recognized as a valid and useful figure of merit.

This paper derives a set of characteristics for light collection, noise and saturation in a camera. The starting point is the need for a minimal, but sufficient set of characteristics for the radiometric performance of a hyperspectral camera. The treatment here derives characteristics that are directly comparable between cameras, and also simplify comparison to application requirements. The characteristics are testable on the camera as a whole, and have potential for use as figures of merit for radiometric performance. It turns out that the characteristics can also quantify the radiometric performance of broadband cameras. Furthermore, it can be noted that the treatment applies to spectrometers as well.

Of course, the characteristics discussed here represent just a small subset of what is needed to fully characterize a camera, since only the radiometric properties of a single camera pixel are considered. Also, the treatment is arguably only a rehashing of textbook concepts. Still, the results appear to be potentially useful for providing better specification of cameras.
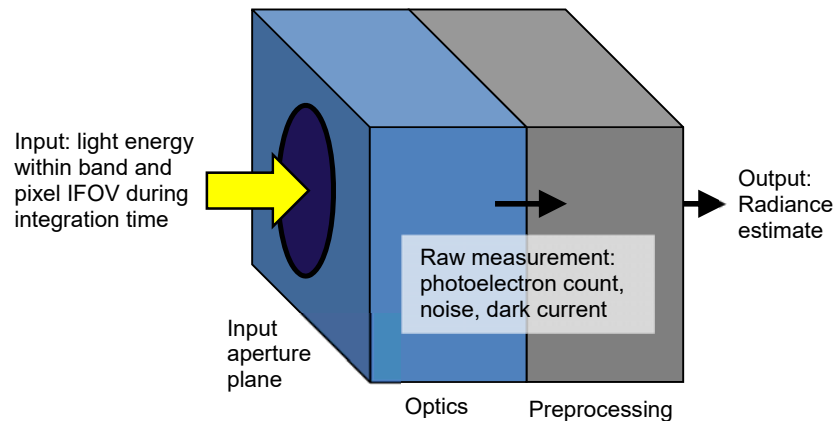

Figure 1. The external "black box" view of a radiometric camera, as considered in this paper.

2. CONVENTIONAL SIGNAL MODEL

Consider the measurement of radiance in a single band in a single pixel in an output image from a hyperspectral camera. The output data is assumed to consist of sampled values of spectral radiance in each pixel. To simplify the treatment, it is assumed that the image pixel corresponds to a single detector element in the image sensor of the camera. Possible generalizations are considered in the discussion section.

Practically all hyperspectral cameras employ photon detectors, where the basic recorded signal is a charge packet of photoexcited electrons, with some contribution from dark current. The quantity to be measured is the incoming spectral radiance $L_\lambda(\lambda)$. The basic recorded signal can be expressed as an expectation value for the electron count given by

$$N_e = N_s + N_d = \underbrace{A_{det}\Omega_{det}}_{collection} \; \underbrace{T(\lambda)\eta(\lambda)}_{loss} \; \underbrace{t_i\Delta\lambda}_{\substack{camera \\ settings}} \; \underbrace{\frac{\lambda}{hc}}_{\substack{photon \\ energy}} \; \underbrace{L_\lambda(\lambda)}_{\substack{input \\ signal}} + \; \underbrace{i_d t_i}_{\substack{dark \\ current}} \tag{1}$$

Here N_s is the photosignal due to light collected by the lens, whose exit pupil subtends a solid angle Ω_{det} seen from the detector with area A_{det} (neglecting here the possibility of optical immersion of the detector). Light loss is described by the optics transmission $T(\lambda)$ and the detector quantum efficiency $\eta(\lambda)$ (taken to include the fill factor of the image sensor, if relevant). Light is collected over an integration time t_i within a spectral bandwidth $\Delta\lambda$, and hc/λ is the photon energy at wavelength λ, taken here to be the band center. (This factor can be omitted if the quantity to be measured is taken to be the photon spectral radiance, $L_{ph,\lambda}(\lambda)$.) N_d is the contribution due to dark current i_d (in electrons per unit time).

A full model of the pixel signal would include a weighted averaging of the scene radiance over the pixel area, according to a point spread function (PSF), and a weighted averaging over the radiance spectrum according to the spectral response function (SRF). For the purpose of the radiometric signal model here, this spatial and spectral averaging need not be taken into account explicitly. The spatial and spectral averaging is represented in (1) by multiplication by the pixel area and bandwidth, just like multiplication by t_i represents an average over time. This approach is valid for hyperspectral imaging with a small relative bandwidth $\Delta\lambda/\lambda$ as long as the light source has a smooth spectral and spatial variation, as is the case for commonly used calibration sources.

The noise in the measured value has contributions from several noise mechanisms. The readout noise σ_r is customarily represented as an RMS fluctuation of the electron signal due to noise contributions from the camera electronics. Here, σ_r is taken to include digitization noise as well. The readout noise sets a lower limit on the measurable light levels. The other noise contributions in the signal model here are the Poisson noise of the photoelectrons, and of the dark current electrons, with variances equal to their mean values, $\sigma_s^2 = N_s$ and $\sigma_d^2 = N_d$ respectively. The total RMS noise in the photoelectron count is then

$$\sigma_e = \sqrt{N_s + N_d + \sigma_r^2} \tag{2}$$

Keep in mind that this includes digitization noise, so that σ_r represents the noise floor of the output measured radiance, and σ_e the total output noise, referenced back to an equivalent noise in the electron count.

Saturation occurs when the charge capacity of the detector element is reached, or when the signal reaches the maximum value in subsequent analog circuits or in the digitization. In any case, the saturation level can be referenced back to a maximum count of electrons at the detector, denoted N_{sat}.

The electron signal N_e, with noise σ_e, is digitized to a value

$$D = \text{round}\left[G(N_e \pm \sigma_e) + D_0\right]$$

where G is a gain factor and D_0 represents any internal signal offset. The output value from the camera is an estimated spectral radiance value

$$\tilde{L}_\lambda = \frac{(D - \tilde{D}_0)}{\tilde{K} t_i \Delta\lambda} \frac{hc}{\lambda}$$

Here, $\tilde{K}$ and $\tilde{D}_0$ are parameters for overall gain and offset estimated from camera calibration. The discussion here does not consider systematic errors resulting from inaccuracies in these parameters. Therefore most of the discussion will be in terms of the electron signal N_e. Thus the discussion is valid for any camera that outputs light samples proportional to radiance, such as raw images from photographic cameras.
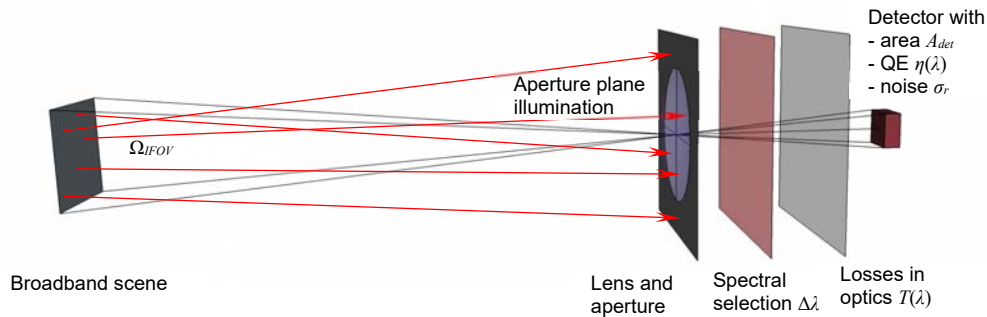
Figure 2. Schematic illustration of a real camera. Broadband light from a pixel area in the scene (red arrows) illuminates the aperture plane of a camera. A lens collects the part of the light falling within its aperture area, and images it onto a detector. Spectral optics filters out a spectral band $\Delta\lambda$. Some light is lost in the optics and in the detector, as described by the lens transmission $T(\lambda)$ and the detector quantum efficiency $\eta(\lambda)$. The detector has an area A_{det}, which together with the focal length determines an instantaneous field of view Ω_{IFOV} in the scene. Detector noise σ_e is represented as an RMS electron count for the integration time t_i.
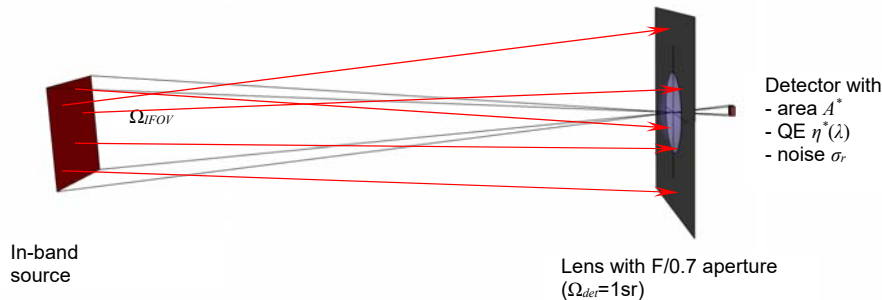

Figure 3. Equivalent camera model representing radiometric performance with a minimum of parameters. The scene is assumed to emit light within the band from within the IFOV onto the camera aperture plane. The camera consists of an ideal lens and an equivalent pixel detector. The size scaling is defined by requiring $\Omega_{det} = 1\,\mathrm{sr}$ (implying F/0.7 aperture) and a detector with peak QE of 1 and an area A^* giving the same light collection as the real camera. The QE $\eta^*(\lambda)$ accounts for wavelength-dependent losses. The focal length is adapted to maintain the same FOV as the real camera. Thereby, the detector properties become a standardized representation of the camera radiometric performance, as discussed in the text.

3. CHARACTERIZATION VIA AN EQUIVALENT CAMERA

3.1 Transformation to an equivalent camera

The photosignal in (1) is a product of many factors which all have the effect of scaling the photoelectron signal N_s. When characterizing the camera as a "black box", without access to individual components, these factors cannot be observed independently. It is therefore useful to establish a simpler model of the camera, described by a minimal number of parameters sufficient to capture the quantities of interest for performance evaluation. For that, the real camera, illustrated in Figure 2, can be transformed into a radiometrically equivalent camera with fewer parameters, illustrated Figure 3. The transformation can be envisaged as a sequence of steps, each maintaining the throughput, noise and saturation level of the real camera:

1. As discussed above, assume that the entire scene consists of a uniform radiance equal to the pixel radiance signal that results from PSF and SRF weighting. This allows us to disregard the actual PSF and SRF shapes.

2. The losses in the optics and the detector are combined into an overall QE for the detector, and the lens is assumed lossless.

3. The lens aperture is changed so that the focused light cone from the lens subtends a solid angle of 1 steradian seen from the detector (corresponding to a F/0.7 circular aperture), and the detector area is scaled by an inverse amount to maintain the same throughput. (If the camera employs optical immersion of the detector, this is included in the rescaling, so that the rescaled detector area is for a detector without immersion.) The readout noise and dark current are not rescaled.

4. The camera is scaled down in size, and the QE is scaled up accordingly to maintain overall throughput, so that the overall QE is 1 at the wavelength where it has its maximum value.

5. The focal length of the lens is changed, while maintaining its F/0.7 aperture, so that the pixel IFOV returns to its original value for the new detector area.

The result of this transformation is a camera with the same light collection, noise, and saturation level as the real camera, but with fewer parameters. The next sections express this model mathematically.

3.2 Characterizing throughput

In the transformation to an equivalent camera, the scaling factor for detector area in step 3 above is $A_{det}(\Omega_{det}/1\,\mathrm{sr})$. In step 4, the detector is further scaled by a factor equal to the peak value of $T(\lambda)\eta(\lambda)$ over all wavelengths. Thus the detector area of the equivalent camera can be written

$$A^* = A_{det}\Omega_{det}\max_{\lambda}\left[T(\lambda)\eta(\lambda)\right] \tag{3}$$

The proposed notation here parallels D^*, because A^* can be seen as a figure of merit for the overall light collection of the hyperspectral camera. Note that A^* can also be taken to express the peak light collection for a camera with broadband response receiving monochromatic light at the wavelength of peak overall quantum efficiency, as discussed below.

Observe that according to (3), the physical dimension of A^* is area times solid angle ($\mu m^2 sr$). In the equivalent camera model, with Ω_{det} fixed at 1 steradian, A^* can be interpreted as an equivalent pixel area. However when A^* is used in expressions for other quantities below, the appropriate unit is area times solid angle.

The quantum efficiency of the equivalent camera can be found from

$$\eta^*(\lambda) = \frac{T(\lambda)\eta(\lambda)}{\max_{\lambda}\left[T(\lambda)\eta(\lambda)\right]}$$

This function gives the relative variation of overall throughput across the spectral range. For hyperspectral imaging, $\eta^*(\lambda)$ is strictly defined only at sampling points at the center wavelength of each band, but it will normally form a smooth graph representing the wavelength dependence of the overall light loss. By normalizing to a peak value of 1, $\eta^*(\lambda)$ provides information about relative wavelength-dependent loss in a way which is comparable between cameras, independently of the throughput. It is thus a useful "graph of merit". Numerical figures of merit for the spectral variation can be defined based on $\eta^*(\lambda)$. Notably, the minimum of $\eta^*(\lambda)$, which can be denoted η^*_{min} expresses the relative amount of variation in quantum efficiency across the spectral range of the camera. At the same time, η^*_{min} together with A^* gives a guaranteed minimum performance level for light collection. The average value of $\eta^*(\lambda)$ over all wavelengths also carries information about the overall losses within the specified spectral range for the camera. However this quantity may be less suited for specification purposes, since it does not guarantee a certain level of performance at any given wavelength.

3.3 Equivalent camera signal model

With A^* representing nominal throughput and $\eta^*(\lambda)$ representing losses, the signal model (1) simplifies to

$$N_e = N_s + N_d = A^*\eta^*(\lambda)\,t_i\Delta\lambda\,\frac{\lambda}{hc}\,L_\lambda(\lambda) + \mathrm{i}_d t_i \tag{4}$$

Note that this still has the form of a physical camera model, whose performance can be expressed in conventional terms such as in the signal and noise models of EMVA1288.

3.4 Characterizing noise

With risk of stating the obvious, consider this: We are normally interested in measuring radiance, therefore we would like to express noise as an equivalent radiance quantity. From energy conservation, we have the well-known relation

$$L_\lambda A_{det} \Omega_{det} = L_\lambda A_{entrance} \Omega_{IFOV} . \tag{5}$$

showing that the radiance collected in the entrance pupil $A_{entrance}$ within the pixel instantaneous field of view Ω_{IFOV} is equal to the radiance presented to the detector from the exit pupil of the lens (neglecting transmission losses). This has a useful interpretation: The effect of the lens can be seen as to present to the detector the pixel radiance we wish to measure, in a well-defined geometry. Thus, the detector signal simply needs to be scaled by a calibration factor to yield a correct radiance estimate. Therefore, if we can express noise and saturation in terms of the radiance arriving at the detector, we have at the same time a quantity that relates noise and saturation to the pixel radiance to be measured.

The noise floor of a hyperspectral camera can be expressed as a noise equivalent spectral radiance (NESR), representing a signal level equal to the internally generated noise. NESR is then a signal satisfying

$$N_s = \sqrt{\sigma_r^2 + N_d} \tag{6}$$

The NESR can be found from (4) and (6) by solving for L_λ (neglecting for simplicity the offset due to dark current):

$$\sqrt{\sigma_r^2 + N_d} = A^* \eta^*(\lambda)\, t_i \Delta\lambda\; \frac{\lambda}{hc}\, L_\lambda(\lambda)$$

$$\Rightarrow NESR(\lambda) = \frac{hc}{\lambda}\, \frac{\sqrt{\sigma_r^2 + i_d t_i}}{A^* \eta^*(\lambda)\, t_i \Delta\lambda} \tag{7}$$

This quantity is useful for relating a particular camera to the sensitivity requirements of a particular application. However, NESR has a disadvantage similar to NEP mentioned in the introduction, in that it depends on a particular choice of integration time and spectral bandwidth. A given application will typically accommodate a range of bandwidths and integration times, but NESR is not well suited for comparing cameras with different bandwidths within this range, since a camera with larger bandwidth will tend to have a lower NESR simply due to the bandwidth difference.

To derive quantities better suited for camera benchmarking, note that the contribution of dark current to noise can often be neglected, since in the majority of hyperspectral applications, the light level is such that the photocurrent is significantly larger than the dark current. The readout noise then sets the lower limit on the measurable amount of light entering the camera within the integration time, limiting the allowable frame rate for a given light level and SNR requirement. We can obtain a characteristic for noise floor which is independent of integration time by considering the integral of spectral radiance over the integration time, approximated here by multiplication by t_i. The signal is equal to the readout noise for a value of this integral which can be termed the "noise equivalent spectral radiance dose", NESRD:

$$NESRD(\lambda) = NESR(\lambda) t_i = \frac{hc}{\lambda}\, \frac{\sigma_r}{A^* \eta^*(\lambda)\, \Delta\lambda}$$

For the purpose of camera specification, NESRD still has the disadvantage of a direct dependence on bandwidth. Thus two cameras with different spectral resolution cannot be compared head to head by comparing either NESR or NESRD if they have different bandwidths within the range acceptable for the application. Consider for example a case where two cameras have bandwidths that differ by a factor 2, but with the same detector readout noise. The camera with the best spectral resolution will appear from NESR or NESRD to have a noise floor higher by a factor 2, even if its data can be binned spectrally to produce an image comparable to the camera with poorer spectral resolution. This issue can be resolved by characterizing noise by the total amount of light, within the integration time and within the band, which is needed to give a signal equal to the readout noise. This quantity is obtained by integrating NESRD over the bandwidth, approximated here by multiplication by $\Delta\lambda$. The resulting measure of noise floor can be termed the "noise equivalent radiance dose", NERD:

$$NERD(\lambda) = NESRD(\lambda)\Delta\lambda = \frac{hc}{\lambda}\frac{\sigma_r}{A^*\eta^*(\lambda)}$$

This quantity has the same kind of application-independent invariance for comparison of cameras as D^* has for detectors. Note that just like NESR, NERD is also independent of IFOV. NERD expresses the amount of light within the band and integration time, per area in the entrance aperture plane and per solid angle within the IFOV, required for the signal to equal the noise. It can be interpreted as a mapping of readout noise into the input radiance space according to (5).

To express performance in a single numerical value, which also may make the interpretation clearer, we consider the wavelength of peak QE, where $\eta^*(\lambda) = 1$, and consider the photon radiance dose, *i.e.* the number of photons per area arriving at the aperture plane where signal equal to the noise. Then NERD can be expressed as a photon quantity:

$$NERD_{ph,\min} = \frac{\sigma_r}{A^*}$$

This is the number of photons per solid angle and per unit area which must arrive at the camera entrance aperture plane within the pixel IFOV and within the integration time, at the wavelength of lowest loss, in order to generate a signal equal to the readout noise. As a photon quantity, rather than an energy quantity, it can be seen to give a better representation of readout noise, whose origin is independent of wavelength. Observe that $NERD_{ph,\min}$ is a clear-cut representation of the technological tradeoff that determines the noise floor in the input radiance space: To improve performance, either the detector readout noise must be reduced, the light collection must be increased, or the losses must be reduced. $NERD_{ph,\min}$ is therefore an application-independent benchmark for the noise floor, which represents the performance of the complete camera rather than an internal component. This quantity is also well defined for multispectral and broadband cameras, representing noise floor at the wavelength of minimum loss.

The notation $NERD_{ph,\min}$ is not particularly elegant. In keeping with other proposals here, the symbol N^* could be used to denote this quantity. Alternatively, a proper figure of merit, in the sense that a higher value is better, could be obtained from the inverse quantity, which might be called "specific photon sensitivity" and denoted $S^* = 1/NERD_{ph,\min}$.

3.5 Characterizing saturation
The saturation level is an electron count just like the readout noise, therefore quantities representing saturation can be derived analogously to the noise floor. Using the form of (7), saturation occurs when the input radiance reaches the saturation spectral radiance, SSR:

$$SSR(\lambda) = \frac{hc}{\lambda}\frac{N_{sat} - i_d t_i}{A^*\eta^*(\lambda)\ t_i\Delta\lambda}$$

If the contribution from dark current is small, the SSR becomes

$$SSR(\lambda) \approx \frac{hc}{\lambda}\frac{N_{sat}}{A^*\eta^*(\lambda)\ t_i\Delta\lambda} \tag{8}$$

Since the expressions for SSR and NESR are essentially the same, apart from a scaling factor difference, these two quantities can conveniently be represented by the same graph with two different vertical scales in a specification sheet.

As for the noise floor, it would be possible to define an application-independent comparative measure of saturation level, a "saturation equivalent radiance dose". However the characterization of saturation level in itself is not very important since problems with saturation can be resolved easily by reducing integration time in most applications. Arguably, a better figure of merit is the signal to noise level at saturation. In the absence of significant dark current, assuming that photoelectron noise is dominating, and ignoring clipping of noise peaks near saturation, the maximum SNR is

$$SNR_{\max} = \sqrt{N_{sat}}$$

This is again not a conservative quantity, but rather a best case, like A^* and $NERD_{ph,\min}$. On the other hand, it is a useful parameter describing the charge capacity of the detector. It is also a well understood quantity which is widely reported, for example in EMVA1288-compliant datasheets.

3.6 Characterizing dark current

In some cases, such as for SWIR cameras, the dark current (including the effect of internal thermal radiation) may fill a significant part of the well capacity. The treatment here assumes that the camera performs internal dark current subtraction in order to output radiance estimates, in which case the dark current is not directly observable to the user. A significant dark current will still have three different effects: 1) The noise floor increases due to the Poisson noise of dark current, 2) the optical saturation level is reduced due to the partial well fill by dark current, and 3) the peak SNR at optical saturation will be reduced. These effects are captured by the equivalent camera model here. As discussed below, these effects therefore enable estimation of dark current based on radiance images output by the camera, for example for the purpose of verifying a specification.

If a camera has significant dark current, this should be brought out in the camera characteristics. In addition to, or instead of, specifying the dark current directly, it can be informative to present two derived characteristics which are easier to interpret for non-expert users: Firstly, the smallest amount of dark current that may affect the signal occurs at the integration time where the Poisson noise of dark current is equal to the readout noise:

$$t_{knee} = \frac{\sigma_r^2}{i_d}$$

At this integration time, there will be a "knee" in a graph of detector dark noise as a function of integration time. This quantity can be compared to a chosen integration time in order to check if the noise floor is affected by dark current noise. Secondly, the dark current can be expressed as the integration time needed for the dark current to fill the well capacity entirely:

$$t_{dark} = \frac{N_{sat}}{i_d}$$

This quantity is helpful because it can be compared directly to a chosen integration time to check if the saturation level will be significantly reduced due to dark current. The quantities t_{knee} and t_{dark} are thus potentially useful figures of merit, though complicated by the fact that dark current is strongly temperature dependent. Specification of this temperature dependence is not discussed here. See for example the treatment of temperature dependence in EMVA1288.

It can be noted that dark current could be specified as a "dark current equivalent spectral radiance" (DCESR) given by

$$DCESR(\lambda) = \frac{hc}{\lambda} \frac{i_d}{A^* \eta^*(\lambda) \, \Delta\lambda}$$

For a camera specification, it may not normally be justifiable to include DCESR as a separate graph due to the added complexity and space required. Note, however, that DCESR can be represented by the same graph as NESR and SSR, on a separate vertical scale. Still, this quantity might be considered of only marginal interest to users in most cases.

4. MEASUREMENT OF CAMERA CHARACTERISTICS

The following is a very brief outline of a possible procedure to determine values for the proposed characteristics for the case of a "black box" camera outputting radiance estimates. The procedure uses only these radiance values, without any knowledge of the inner construction of the camera. Only a single pixel is considered here. Characterization of variations in response across the camera field of view (due to vignetting etc.) is a separate issue not discussed here.

- Integration times are assumed to be very close to their set values, relying on the accuracy of electronic clocks and switching.

- Spectral bandwidths are assumed to be measured separately. Provided that the test spectrum is smooth, the actual SRF can be represented by the bandwidth of an equivalent box-shaped SRF, as assumed here for simplicity.

- The measurement is based on a stable, spatially uniform broadband source, such as a halogen lamp in an integrating sphere, with a known radiance spectrum. It will be helpful if the source light level can be varied without changing the spectral distribution.

- A sequence of image frames is recorded for several different settings of integration time, spanning from very short, at the noise floor, to very long, in saturation. The measurement is made with a relatively high light level in order to have a negligible contribution from dark current.

- $SNR_{\max}$ is determined from a photon transfer analysis [5] based on these data, and squared to find N_{sat}.

- $SSR(\lambda)$ is determined from the saturation level for each band, as observed in the photon transfer characteristic.

- The product $A^* \eta^*(\lambda)$ is determined for each band from N_{sat} and $SNR_{\max}(\lambda)$ using (8).

- A^* is determined as the peak value of of $A^* \eta^*(\lambda)$ over all bands.

- $\eta^*(\lambda)$ is found from the relative variation of $A^* \eta^*(\lambda)$ over all bands, scaled to a peak value of 1.

- The measurement of SSR is repeated for lower light levels and correspondingly longer integration times. Dark current is estimated from the reduction of SSR with increasing integration time.

- $NESR(\lambda)$ is determined directly from the noise in radiance images recorded in the dark[1] at a short integration time. Results can be compared to the expected dark current noise to verify that the observed NESR is dominated by readout noise.

- σ_r can be determined from the photon transfer measurement, or from NESR using (7) , providing a consistency check. In principle, the same value should be obtained by both methods and over all bands. (In some cases, σ_r may vary from band to band, for example if detector gain settings vary between groups of bands, or if the internal processing in the camera involves spectral resampling, see the discussion.)

- Other noise characteristics can now be calculated from σ_r and other measured quantities.

Thus, even for a camera that performs internal "black box" preprocessing to output radiance spectra, the proposed characteristics can be determined "outside the box" from dark frames and images of a known, uniform broadband source, which can be readily measured. Of course, characterization can often be made using simpler and more direct methods, in cases where raw data or component characteristics are available.

5. DISCUSSION

5.1 The equivalent camera representation

The equivalent camera model constrains the lens exit pupil to subtend a solid angle of 1 steradian, seen from the detector. This is an arbitrary choice. It is arguably a natural choice from the mathematical point of view, since solid angle has no physical unit. D^* sets a solid precedent for constraining the definition to a mathematically convenient value. On the other hand, this solid angle corresponds to a lens aperture of F/0.7 which is only rarely encountered in practice. Therefore, values for A^* may be significantly smaller than the actual pixel area.

The physical dimension of A^* is a potential source of confusion. It is derived as a product of area and solid angle, and this is how it must be used in calculations. At the same time, A^* is the area of the detector element in the equivalent camera. Thus when it is used as a figure of merit, it is fair to report A^* in units of μm^2.

Another potentially confusing aspect is in the definition of NERD, which counts the number of photons per solid angle within the IFOV and per area in the entrance pupil plane, integrated over the band and integration time. The reason for not integrating over IFOV is basically the relation (5), which shows that the detector in the equivalent camera receives the signal radiance to be measured over a solid angle of 1 steradian, making NERD invariant of with respect to IFOV on the object side. The reason for not integrating over the entrance pupil area is that the NERD figure of merit (as well as N^*) measures the ability of the camera to collect the light that is available at its aperture. Thus NERD is a relevant figure of

---

[1] Note that about half of the radiance values should be negative in darkness, in cases where a non-negligible dark current or offset is subtracted internally in the camera. If negative radiance values are clipped to zero by the camera, the noise estimate will have to be corrected by assuming a symmetrical distribution of values around zero.

merit for normal use cases of cameras, where only a small fraction of the light from the object is collected by the camera aperture. Notably, this excludes microscopy, where light collection is characterized customarily and adequately by the numerical aperture of the objective.

The equivalent camera model could alternatively have been formulated by replacing the relative QE $\eta^*(\lambda)$ by a relative responsivity peaking at 1 for a somewhat different wavelength than the peak of $\eta^*(\lambda)$. Given the practical fact that photoelectron statistics is a dominating noise source in the detectors normally used, and also considering that the QE has 1 as a natural peak value, a formulation based on photons is quite clearly favorable. The equivalent camera model is then essentially the same as the EMVA1288 signal model.

Examples or experiments illustrating the use of the equivalent camera model in practical cases are missing from this paper, regrettably, but may be presented in follow-up work as time permits.

5.2 Simplifications made without loss of generality

This treatment has been simplified in several ways in order to separate out the characteristics related to throughput, noise and saturation. Mostly, the simplifications do not lead to loss of generality: The shapes of PSF or SRF, which are important for resolution and coregistration, do not enter into the analysis of radiometry, as discussed in Section 2, because the radiometric characteristics are measurable using uniform sources with slowly varying spectra. The treatment here is valid for a single pixel, but radiometric characteristics may vary across the camera field of view due to vignetting etc. This variation is arguably a separate issue to be characterized by other quantities, and the treatment here can be taken to represent an average pixel. Similarly, the absolute accuracy of the output radiance values is essentially a separate issue. Of course, the absolute accuracy will affect the accuracy of SSR and NERD, but the accuracy requirement for these characteristics is arguably relaxed compared to normal accuracy requirements for absolute radiometric calibration. Therefore, in the absence of gross errors in radiometric calibration, the measurement procedure outlined above can determine parameters discussed here to a sufficient accuracy. Finally, the treatment has been simplified by assuming that the detector is a photon detector where the important noise mechanisms are readout noise and Poisson noise of the generated electrons. It can be noted that in the case of hyperspectral imaging, other detector types, such as thermal detectors (microbolometers) are not in common use, and there is no sign that this situation will change in the foreseeable future.

5.3 Cameras employing resampling or reconstruction

A simplification that sacrifices some generality is the assumption that a radiance value in the output image is based on the signal from a single detector element. Many camera types employ some form of resampling or reconstruction to generate an output radiance image where each radiance value may depend on the signal from multiple detector elements. Since detectors generally have a highly linear response, an output radiance value will usually be a linear combination of the signals from a set of detector elements, with a correspondingly weighted summation of the variances. For each output pixel, it will then be possible to approximate the radiometric behavior as an equivalent single detector using the parameters of the single-detector model discussed here. However, it is likely that the parameters will vary from band to band, and from pixel to pixel in the image, due to different sets of weights in the linear combination of signals from detector elements. The radiometric characteristics can then still be characterized by giving mean or worst-case values for the parameters of the equivalent camera model. Even if the actual operation of the real camera is different, and some accuracy may be lost, the set of parameters of the equivalent camera has advantages by being simple and widely understood, as well as by being comparable between cameras.

5.4 Application beyond hyperspectral imaging

As mentioned above, the quantities proposed here as figures of merit for hyperspectral cameras are also potentially useful in other fields. The model for a single band in a hyperspectral camera can readily be reinterpreted as a model for a camera with broadband response, by letting $\eta^*(\lambda)$ represent the wavelength-dependent QE, instead of the QE in each band. The quantities, A^*, $\eta^*(\lambda)$, $NERD_{ph,\min}$ and $SNR_{\max}$, supplemented by t_{knee} when needed, are then relevant figures of merit even for conventional cameras, and suitable for characterization of the camera as a "black box". It is clear, however, that outside the field of hyperspectral imaging, radiometric performance is somewhat less of a priority, with more emphasis on characteristics such as resolution, spatial noise, and color balance. It is possible that the quantities derived here, or similarly defined characteristics, have been proposed by other workers in the past, but not gained widespread use. Another reason why "black box" characteristics for camera radiometry are not common may be that lenses have traditionally been exchangeable, leading to the present convention where specifications are given for lenses and image sensors separately. Notably, the T-number of a lens is an equivalent F-number taking into account lens transmission, but still not including

losses in the detector. Today, it is common to have cameras where the lens and the image sensor form an integrated unit, such as mobile phone cameras. Standards exist for evaluation of the performance of such cameras, but the emphasis is then on the visual quality of the output images [6,7], and not on the quantitative radiometric characteristics. It is possible to establish radiometric and spectral calibration for such cameras[8], but the resulting performance cannot be well predicted by camera specifications, and the calibration factors are not a good way to specify the camera. There is thus a clear rationale for considering new figures of merit such as those discussed here, not only for hyperspectral imaging, but even for conventional cameras, particularly for scientific and technical applications where radiometric quality is important.

In the case of conventional imaging, cameras often apply noise reduction or other image manipulations aimed at producing a visually pleasing image. Such processing steps violate the assumptions of the equivalent camera model, but the concepts discussed here can still be applied to raw images from such cameras.

It can be noted that since the radiometric model represents a single pixel in a hyperspectral image, the treatment here is also directly applicable to spectrometers.

6. SUMMARY AND CONCLUSION

This paper has considered radiometric characterization of a complete hyperspectral camera, including optics and image sensor as well as preprocessing for dark signal subtraction and radiometric calibration. Thus it is assumed that the camera outputs images in the form of estimated spectral radiance for each pixel. It has been outlined how characteristics for a complete camera can be determined from output radiance images recorded from known calibration sources, without any knowledge of internal component properties such as F-number or readout noise. Results then represent the total effect of the optics, the image sensor, and the internal preprocessing in the camera.

The conventional signal model for a camera is basically a product of many factors: Aperture, transmission, detector area, quantum efficiency, responsivity etc. These factors cannot be determined individually from measurements on the output from the complete camera. Therefore, the signal model has been reformulated to a simpler model of a radiometrically equivalent camera where the lens aperture on the detector side is constrained to fill a solid angle of 1 steradian seen from the detector element. The camera light collection capability is then represented by a single parameter which is the detector pixel area of this equivalent camera, A^*. Wavelength-dependent losses are represented by an equivalent quantum efficiency $\eta^*(\lambda)$ with a peak value of 1.

Noise in the measured radiance values is ideally dominated by the Poisson noise of photoelectrons in the detector. The camera noise floor is determined by the additional readout noise, customarily expressed as an RMS fluctuation in the electron count. This quantity is not directly observable in the output data, but can be determined from a photon transfer measurement. However the noise floor in terms of measurable radiance levels also depends on the net light collection of the camera. Therefore, the readout noise by itself does not represent the overall noise performance of the camera. NESR can be calculated from the readout noise using the equivalent camera model, and can relate the noise floor to application requirements. Due to its dependence on bandwidth, NESR is not suitable for camera comparison, however. For benchmarking purposes, it is proposed to use a figure of merit for noise floor invariant to bandwidth and integration time, the noise equivalent radiance dose, NERD. Different ways to express and denote this quantity have been discussed.

Saturation can be described for the complete camera by the saturation spectral radiance, SSR, and the peak SNR, which are relatively conventional quantities. For presentation of camera characteristics, it is convenient that SSR and NESR can be represented by the same graph using two different axis scales.

Dark current can be specified in a conventional way as electrons per second, but the discussion here includes suggestions for figures of merit representing the integration times where dark current begins to affect the noise floor and the saturation level. The "dark current equivalent radiance" has also been mentioned as a possible way to express dark current.

In conclusion, the set of characteristics discussed here arise out of the signal model for a hyperspectral camera, treated as a "black box" which outputs images in the form of radiance estimates. Importantly, conventional cameras can also be meaningfully characterized by the quantities discussed here, by assuming that the signal consists of narrowband light, instead of the narrow bandwidth of a hyperspectral camera. In particular, the quantities A^*, $\eta^*(\lambda)$, and $NERD_{ph,\min}$ are arguably useful characteristics for many types of cameras. These quantities arise from the equivalent camera model referenced to a particular image-side lens aperture, in a way similar to how D^* is referenced to a particular detector area and bandwidth.

ACKNOWLEDGMENTS

REFERENCES

[1] See for example https://www.dxomark.com/ and https://www.photonstophotos.net/

[2] "Standard for Characterization of Image Sensors and Cameras Release 3.1", EMVA standard 1288, European Machine Vision Association (2016)

[3] Torbjørn Skauli, "Feasibility of a standard for full specification of spectral imager performance," *Proc. SPIE* **10213**, 102130H (2017)

[4] R. C. Jones, "Quantum efficiency of photoconductors," *Proc. IRIS* **2**, 9 (1957)

[5] J. R. Janesick, K. P. Klaasen, and T. Elliott "Charge-Coupled-Device Charge-Collection Efficiency And The Photon-Transfer Technique," *Optical Engineering* **26**(10), 261072

[6] Robert L. Nicol, "Image sensor characterization using SMIA standard," Proc. SPIE **6196** 61960E (2006)

[7] Dietmar Wueller, Akira Matsui, Naoya Katoh, "Visual Noise Revision for ISO 15739," Proc. Electronic Imaging, Image Quality and System Performance XVI, pp. 315-1-315-7 (2019)

[8] Olivier Burggraaff, Norbert Schmidt, Jaime Zamorano, Klaas Pauly, Sergio Pascual, Carlos Tapia, Evangelos Spyrakos, and Frans Snik, "Standardized spectral and radiometric calibration of consumer cameras," Optics Express **27**(14) 19075-19101 (2019)