

Sensor noise informed representation of hyperspectral data, with benefits for image storage and processing

Torbjørn Skauli*

Norwegian defence research establishment (FFI), P O Box 25, 2027 Kjeller, Norway

*torbjorn.skauli@ffi.no

Abstract: Many types of hyperspectral image processing can benefit from knowledge of noise levels in the data, which can be derived from sensor physics. Surprisingly, such information is rarely provided or exploited. Usually, the image data are represented as radiance values, but this representation can lead to suboptimal results, for example in spectral difference metrics. Also, radiance data do not provide an appropriate baseline for calculation of image compression ratios. This paper defines two alternative representations of hyperspectral image data, aiming to make sensor noise accessible to image processing. A “corrected raw data” representation is proportional to the photoelectron count and can be processed like radiance data, while also offering simpler estimation of noise and somewhat more compact storage. A variance-stabilized representation is obtained by square-root transformation of the photodetector signal to make the noise signal-independent and constant across all bands while also reducing data volume by almost a factor 2. Then the data size is comparable to the fundamental information capacity of the sensor, giving a more appropriate measure of uncompressed data size. It is noted that the variance-stabilized representation has parallels in other fields of imaging. The alternative data representations provide an opportunity to reformulate hyperspectral processing algorithms to take actual sensor noise into account.

©2011 Optical Society of America

OCIS codes: (110.4280) Noise in imaging systems; (110.4234) Multispectral and hyperspectral imaging; (100.4145) Motion, hyperspectral image processing.

References and links

1. B. Penna, T. Tillo, E. Magli, and G. Olmo, “Transform coding techniques for lossy hyperspectral data compression,” *IEEE Trans. Geosci. Rem. Sens.* **45**(5), 1408–1421 (2007).
2. Q. Du and J. E. Fowler, “Hyperspectral image compression using JPEG2000 and principal component analysis,” *IEEE Geosci. Remote Sens. Lett.* **4**(2), 201–205 (2007).
3. J. Wang and C. I. Chang, “Independent component analysis-based dimensionality reduction with applications in hyperspectral image analysis,” *IEEE Trans. Geosci. Rem. Sens.* **44**(6), 1586–1600 (2006).
4. B. Penna, T. Tillo, E. Magli, and G. Olmo, “Progressive 3-D coding of hyperspectral images based on JPEG 2000,” *IEEE Geosci. Remote Sens. Lett.* **3**(1), 125–129 (2006).
5. J. Mielikainen and P. Toivanen, “Clustered DPCM for the lossless compression of hyperspectral images,” *IEEE Trans. Geosci. Rem. Sens.* **41**(12), 2943–2946 (2003).
6. B. Aiuzzi, L. Alparone, and S. Baronti, “Near-lossless compression of 3-D optical data,” *IEEE Trans. Geosci. Rem. Sens.* **39**(11), 2547–2557 (2001).
7. M. J. Ryan and J. F. Arnold, “Lossy compression of hyperspectral data using vector quantization,” *Remote Sens. Environ.* **61**(3), 419–436 (1997).
8. M. J. Ryan and J. F. Arnold, “The lossless compression of AVIRIS images by vector quantization,” *IEEE Trans. Geosci. Rem. Sens.* **35**(3), 546–550 (1997).
9. S. E. Qian, A. B. Hollinger, D. Williams, and D. Manak, “Fast three-dimensional data compression of hyperspectral imagery using vector quantization with spectral-feature-based binary coding,” *Opt. Eng.* **35**(11), 3242–3249 (1996).

10. R. E. Roger and M. C. Cavenor, "Lossless compression of AVIRIS images," *IEEE Trans. Image Process.* **5**(5), 713–719 (1996).
11. G. P. Abouseelman, M. W. Marcellin, and B. R. Hunt, "Compression of hyperspectral imagery using the 3-D DCT and hybrid DPCM/DCT," *IEEE Trans. Geosci. Rem. Sens.* **33**(1), 26–34 (1995).
12. Recent papers on hyperspectral compression were retrieved from the ISI Web of science database in January 2011, but not listed here for space reasons. References are available from the author.
13. J. W. Boardman, "Using dark current data to estimate AVIRIS noise covariance and improve spectral processing" in *Summaries of the Fifth Annual JPL Airborne Geoscience Workshop*, Pasadena, CA 1995.
14. T. Skauli, T. V. Haavardsholm, I. Kåsen, G. Arisholm, A. Kavara, T. O. Opsahl, and A. Skaugen, "An airborne real-time hyperspectral target detection system," *Proc. SPIE* **7695**, 76950A, 76950A-6 (2010).
15. A. S. Norsk Elektro Optikk, <http://www.neo.no/products/hyperspectral.html>.
16. T. Skauli, "Sensor-informed representation of hyperspectral images," *Proc. SPIE* **7334**, 733418, 733418-8 (2009).
17. E. L. Dereniak and G. D. Boreman, *Infrared Detectors and Systems* (Wiley, New York 1996).
18. P. C. D. Hobbs, *Building Electro-Optical Systems: Making It All Work, 2nd ed.* (Wiley, 2009).
19. Spectral Imaging, Ltd., <http://www.specim.fi>.
20. R. O. Green, M. L. Eastwood, C. M. Sarture, T. G. Chrien, M. Aronsson, B. J. Chippendale, J. A. Faust, B. E. Pavri, C. J. Chovit, M. Solis, M. R. Olah, and O. Williams, "Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS)," *Remote Sens. Environ.* **65**(3), 227–248 (1998).
21. J. Hyneczek, "Impactron - A new solid state image intensifier," *IEEE Trans. Electron. Dev.* **48**(10), 2238–2241 (2001).
22. B. Fowler, C. Liu, S. Mims, J. Balicki, W. Li, H. Do, J. Appelbaum, and P. Vu, "A 5.5Mpixel 100 frames/sec wide dynamic range low noise CMOS image sensor for scientific applications," *Proc. SPIE* **7536**, 753607, 753607-12 (2010).
23. J. D. Beck, C.-F. Wan, M. A. Kinch, J. E. Robinson, P. Mitra, R. E. Scritchfield, F. Ma, and J. C. Campbell, "The HgCdTe electron avalanche photodiode," *J. Electron. Mater.* **35**(6), 1166–1173 (2006).
24. A. Martinez, "Capacity bounds for the Einstein radiation channel," *Proc. Int. Symp. Inf. Theory, ISIT 2006*, 9–14 July 2006, Seattle (USA), pp. 366–370 (2006).
25. F. J. Anscombe, "The transformation of Poisson, binomial and negative-binomial data," *Biometrika* **35**, 246–254 (1948) (Curiously, Anscombe credits the "Anscombe transform" to A. H. L. Johnson.).
26. J. A. Rice, *Mathematical Statistics and Data Analysis*, (Duxbury press, 1995) p. 321.
27. R. L. White and J. W. Percival, "Compression and progressive transmission of astronomical images," *Proc. SPIE* **2199**, 703–713 (1994).
28. M. A. Nieto-Santisteban, D. J. Fixsen, J. D. Offenberg, R. J. Hanisch, and H. S. Stockman, "Data compression for NGST" in *Proceedings of Astronomical data analysis software and systems VIII*, ASP Conference series 172, 137–140 (1999).
29. R. A. Gowen and A. Smith, "Square root data compression," *Rev. Sci. Instrum.* **74**(8), 3853–3861 (2003).
30. R. L. Seaman, R. L. White, and W. D. Pence, "Optimal DN encoding for CCD detectors" in *Proceedings of Astronomical Data Analysis Software and Systems XVII*, ASP Conference Series, Vol. 411, 101 (2009).
31. G. M. Bernstein, C. Bebek, J. Rhodes, C. Stoughton, R. A. Vanderveld, and P. Yeh, "Noise and bias in square-root compression schemes," *Proc. Astron. Soc. Pacific* **122**(889), 336–346 (2010).
32. F. Murtagh, J.-L. Starck, and A. Bijaoui, "Image restoration with noise suppression using a multiresolution support," *Astron. Astrophys. Suppl. Ser.* **112**, 179–189 (1995).
33. J.-L. Starck and F. Murtagh, "Astronomical image and signal processing: looking at noise, information, and scale," *IEEE Signal Process. Mag.* **18**(2), 30–40 (2001).
34. B. Zhang, M. J. Fadili, J.-L. Starck, and J.-C. Olivo-Marin, "Multiscale variance-stabilizing transform for mixed-poisson-gaussian processes and its Applications in Bioimaging" in *Proceedings of IEEE International conference on image processing* (Institute of Electrical and Electronics Engineers, New York, 2007) pp. VI233–VI236.
35. C. Poynton, *A Technical Introduction to Digital Video* (Wiley, 1996), Chap. 6.
36. M. Stokes, M. Anderson, S. Chandrasekar, and R. Motta, "A standard default color space for the internet - sRGB", <http://www.w3.org/Graphics/Color/sRGB> (1996).

1. Introduction

Hyperspectral imaging depends on image processing to extract information from image data. A hyperspectral data set should provide accurate information about light levels for the processing, but the hyperspectral sensor inevitably adds noise to the signal. The noise level is an important input parameter to many image processing algorithms, and therefore also a desirable data product from the sensor. Furthermore, in a practical system it is important that the sensor data can be represented in a way that is as compact as possible, since the data volumes produced by hyperspectral sensors are large even for today's computers.

Image sensor technology has made very significant progress since the first hyperspectral sensors appeared in the late 80's. In the current generation of hyperspectral sensors, the dominating noise contribution is a fundamental quantum mechanical effect, namely the random arrival and absorption of photons. This *photon noise* is signal-dependent, but well characterized and understood by the sensor designers. Therefore, good estimates of noise amplitude can be provided to image processing, based on the sensor physics. As will be shown here, information about sensor noise can also lead to reductions in data volume.

Arguably, the sensor can provide two images per recording, one representing measurements of incoming radiance and another representing estimates of noise. However, hyperspectral data formats tend not to make the "noise image" easily accessible for the data user. There is a large body of literature on methods for noise estimation from hyperspectral image data, but there appears to be very few works that make use of noise information derived from sensor physics.

To substantiate this impression, consider the literature on hyperspectral image compression: Clearly, compression is a type of image processing for which noise information is highly relevant, since an important objective must be to avoid wasting storage space on noise. In the 10 most highly cited papers on hyperspectral image compression [1–11], as well as more than 10 of the most recent papers on this topic [12], there is no mention of noise estimation based on sensor properties. Also, the metrics for reconstruction error used by these and many other papers all operate on radiance values. It will be pointed out here that error metrics based on radiance are suboptimal when considering the sensor physics. About four of the papers concern lossless compression, where noise estimation and error metric are not directly relevant. However it is well known, and will be illustrated here, that lossless compression is not necessarily desirable since it implies preserving random noise in full detail.

In other fields, particularly in astronomy, the use of sensor noise information is well established in methods for storage and processing of images. Some of these results are reviewed in section 3.4 below, and found to be highly relevant for hyperspectral imaging.

Although the value of sensor noise information in hyperspectral imaging has been pointed out before [13], the above literature review suggest that there is a need for a closer connection between hyperspectral sensor physics and image processing. The aim of this paper is to contribute to this connection, partly based on experience from development of an airborne real-time processing system for hyperspectral images [14] and collaboration with a manufacturer of hyperspectral sensors [15].

This paper first formulates a simple signal model which applies to many hyperspectral sensors. The main part of the paper describes and discusses different representations of the image data. Two of these, raw data and radiance data, are conventional and widely used. The paper then introduces two alternative representations, "corrected raw data" and "variance stabilized data". For each representation, it is discussed how noise levels can be estimated from the data, for exploitation in image processing. Also, the data volume requirements for the different representations are compared. Consideration of sensor noise characteristics is shown to lead to efficient formats for data storage, with implications for further compression of hyperspectral data. It is pointed out that a variance-stabilized representation has significant benefits for processing, with parallels to standard representations of conventional imagery and to the information-theoretic channel capacity of the sensor.

This paper is based on Ref. 16, with substantial revision and some corrections.

2. Hyperspectral image recording process and information capacity

The signal chain of a hyperspectral sensor is illustrated schematically in Fig. 1 for a single light sample. The signal from the scene passes through optics, photodetector, electronics and software. The interdisciplinary field of optical sensing is well described in textbooks, for example Refs. 17 and 18. Here, a model of the recording process will be formulated with the

aim of being as simple as possible while capturing characteristics important to hyperspectral imaging.

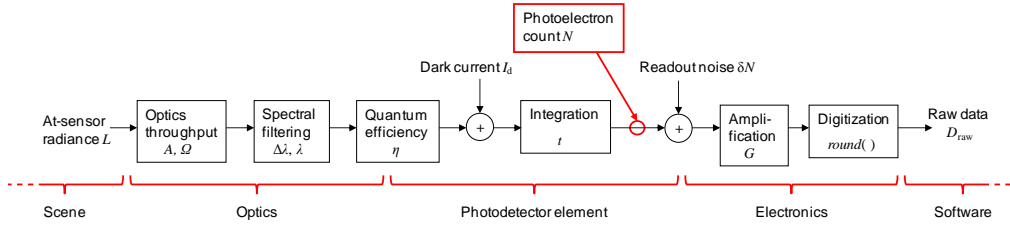


Fig. 1. Model of the light sampling process, common to several important types of hyperspectral sensors in which each light sample corresponds to a single photodetector readout. For the sensor model in this paper, the quantum efficiency is defined to account for losses in the optics and filtering, as well as the photodetector. Due to the random arrival and absorption of photons, the photoelectron count has a Poisson-distributed random variation. This “photon noise” is the dominating noise mechanism in most hyperspectral sensors.

2.1 Basic signal model

The sensor model formulated here applies to hyperspectral sensor types where each radiance sample in the image is derived from a single photodetector reading. This is the case for the commonly used pushbroom-scanning imaging spectrometers, exemplified by AISA [19] or HySpex [15]. The model also applies to whiskbroom-scanning sensors like AVIRIS [20], as well as framing sensors based on filter wheels or tunable filters. Although these different sensor types sample the hyperspectral image cube differently, the characteristics of the photodetector impact the elements of the cube in a similar manner. The sensor model used here does not apply directly to sensors based on interferometers or resampling. In these cases a single light sample in the cube has contributions from several photodetector readings and a somewhat more complex model will be needed to describe sensor noise.

The photodetector is usually an element in a one- or two-dimensional array of photodetectors. For the sensor types considered here, each photodetector element corresponds to a band index i and spatial pixel index j . For a pushbroom-scanning sensor, the pixel index j indicates the position along the line-shaped instantaneous field of view. For a frame imaging sensor, j indicates the pixel position in a two-dimensional field of view.

Consider a spectral band centered at a wavelength $\lambda[i]$ with width $\Delta\lambda[i]$, and an image pixel where the spectral radiance from the scene in band i is $L[i,j]$ (power per unit wavelength, area and solid angle). In the notation here, index $[i]$ refers to a vector with one element for each band. Indices $[i,j]$ refer to a matrix with elements corresponding to all bands and all spatial pixels seen simultaneously by the sensor. Sometimes the indices are omitted for simplicity.

Light enters the sensor aperture with area A . For each spatial pixel j , the sensor receives light over an instantaneous field of view Ω (of course pointing in different directions for different pixels). Then for each spectral band the sensor selects light within a spectral range $\Delta\lambda[i]$ centered on the wavelength $\lambda[i]$. Light is collected over a time t , known as the integration time (or exposure time). The amount of light that corresponds to this spatial, spectral and temporal selection can be expressed as the number of photons entering the sensor:

$$N_{ph}[i,j] = L[i,j]tA\Omega \Delta\lambda[i] \frac{\lambda[i]}{hc} \quad (1)$$

where h is Planck's constant and c the speed of light. Note that Eq. (1) is a simplifying notation where dense spectral sampling is assumed, so that scene radiance and photon energy

can be assumed to be constant within the band. Otherwise the multiplication by $\Delta\lambda$ must be replaced by an appropriate integral over wavelength.

Photons are absorbed in the photodetector by excitation of electrons. These excited electrons, known as *photoelectrons*, are collected and become the electrical signal from the photodetector. The resulting signal can be expressed as a photoelectron count

$$N[i,j] = \eta[i,j]N_{ph}[i,j] + I_d[i,j]t + \delta N. \quad (2)$$

Here the factor $\eta[i,j]$ is the *quantum efficiency*, which accounts for signal losses and sensor response nonuniformity as discussed below. The last terms in Eq. (2) account for unwanted contributions to the signal: The *dark current* $I_d[i,j]$ (electrons per unit time in a particular detector element) is apparent signal from effects such as thermally excited electrons or thermal radiation emitted within the sensor enclosure. The term δN accounts for the *readout noise*, the total signal variability in the transfer and amplification of the photoelectron signal. The readout noise is assumed to be Gaussian with zero mean and standard deviation δN .

After amplification by a gain factor G , the photoelectron signal is digitized to raw digital numbers D_{raw} :

$$D_{raw}[i,j] = \text{round}(GN[i,j] + D_0) \quad (3)$$

where $\text{round}()$ represents the rounding of analog signals to discrete values in digitization. D_0 represents a possible offset in the readout electronics. In the following, it will be assumed that D_0 is zero, or that the offset has been subtracted.

In astronomy, with specialized high-performance sensors, a data representation corresponding to $G=1$ is often used. Commercial imaging systems usually have significant readout noise, and are operated with $G < 1$ so that each increment in D_{raw} corresponds to several photoelectrons. Some newer types of commercial photodetector array can work meaningfully with $G \geq 1$ so that individual photoelectrons give a measurable signal increment. This is the case for silicon-based electron-multiplying CCDs [21] and “scientific CMOS” arrays [22] as well as avalanche photodiodes based on HgCdTe [23].

Note that the photodetector elements have a limited capacity for holding photoelectrons, known as the *full well capacity*, here denoted N_{max} . This effect sets an upper limit on the measurable signal range and on the range of D_{raw} . Stronger signals are said to be in *saturation*, and will typically produce some maximum output value D_{max} .

The set of $D_{raw}[i,j]$ values form a spectral image cube, or a part of it, depending on the sensor type. (Recall that j indexes the spatial pixels that are seen simultaneously by the sensor.) Before processing, the data are usually first corrected for sensor effects and transformed to represent for example radiance values. The data correction and representation is a central topic of this paper.

2.2 Quantum efficiency

Now consider the quantum efficiency η , which describes signal losses and is of particular interest for a hyperspectral sensor. A part of the loss occurs in the optics, due to reflections or absorption as well as to inefficiencies in the spectral selection. Another part of the loss occurs in the photodetector element: Photons may get absorbed by unwanted processes that do not produce a photoelectron, and some photoelectrons may be lost and not become part of the signal. In Eq. (2), the quantum efficiency matrix $\eta[i,j]$ is defined to account for all losses through the system. For simplicity at this point, the matrix $\eta[i,j]$ also incorporates residual nonuniformities of gain and optical throughput, assumed to be small. Conventionally, quantum efficiency is defined as a property of the detector only, or lumped with other factors into a responsivity value. For the sensor model here, however, it is relevant to define η for the entire sensor.

It should be noted that for a hyperspectral sensor, the quantum efficiency tends to have a large variation across the spectral range. This is because the limits of the spectral range tend

to be determined by the falloff of detector quantum efficiency, or the transmission of optical components. Sensor designers tend to try to make a sensor with the widest possible spectral range for a given detector type, and therefore allow designs with low quantum efficiency at the ends of the spectral range. The spectral variation of η is illustrated by the realistic example in Fig. 2. The combined effects of the grating and photodetector wavelength dependencies lead to a very large relative variation of η across the spectral range of the sensor.

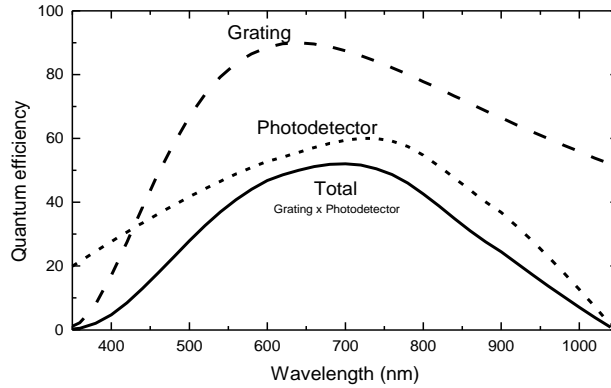


Fig. 2. Illustration of a possible spectral dependence of quantum efficiency (in %) for a grating-based imaging spectrometer sensor with a silicon CCD detector array (roughly comparable to the HySpex VNIR-1600). Lines show the diffraction efficiency of the grating and the quantum efficiency of the CCD array alone, as well as the product of these. The product approximates the quantum efficiency η for this hypothetical but realistic sensor. Spectral variation of the transmission of the optics has been neglected. Over a spectral range from 400 to 1000 nm, the total quantum efficiency varies by a factor 11.

2.3 Noise properties of raw data

It is important to realize that in optical imaging, the quantization of light into photons is not an obscure phenomenon of academic interest, but actually causes the dominating noise contribution, which will be modeled here. The arrival of photons and excitation of photoelectrons are random processes which follow a Poisson distribution. Therefore the photoelectron count N is a Poisson-distributed random variable. This random variation is the “photon noise”, which is a fundamental lower limit on the noise properties of a photodetector.

If \bar{N} denotes the mean photoelectron count, the photon noise has standard deviation

$$\Delta N = \sqrt{\bar{N}} \quad (4)$$

and the signal to noise ratio is $N / \sqrt{\bar{N}} = \sqrt{\bar{N}}$. Thus the noise is signal dependent: With increasing N the noise increases, but its relative amplitude becomes smaller.

Sensors dominated by photon noise can be said to have an optimal signal to noise ratio. Note, however, that the sensor is then only optimal given the values of A , Ω , η , t and $\Delta\lambda$ pertaining to the particular sensor, since these parameters determine N according to Eqs. (1) and (2). The noise performance can still be improved by modifying the sensor to increase the value of the $A\Omega\eta t\Delta\lambda$ product. However increasing $A\Omega\eta$ tends to be expensive and technologically difficult, while increasing t and $\Delta\lambda$ may not be permitted by the application.

Crucially, the measured values can provide good estimates of the photon noise, independently for each light sample, by replacing the mean \bar{N} in Eq. (4) by the measured value N . The typical error in ΔN estimated from the data can be found by considering a sample that is one standard deviation from the mean. The resulting error in estimation of ΔN is then:

$$\sqrt{\bar{N} \pm \Delta N} - \sqrt{\bar{N}} \approx \sqrt{\bar{N}} \left(1 \pm \frac{1}{2\sqrt{\bar{N}}}\right) - \sqrt{\bar{N}} = \pm \frac{1}{2} \ll \Delta N \text{ when } \bar{N} \gg 1$$

Therefore, for the normal case of $\bar{N} \gg 1$, the error in noise estimation from the data value is small. Image data representations proposed below seek to make this noise information easily available to the data user.

Now consider other noise sources: The contribution of dark current $I_d[i,j]$ to the signal, according to Eq. (2), can also be assumed to follow a Poisson distribution. Thus in cases where the photoelectron signal is small, dark current can become the dominating source of noise.

Finally, the *digitization noise*, expressed by the rounding in Eq. (3) has an RMS amplitude [17] of $1/\sqrt{12}$ on the D_{raw} scale, corresponding to $1/(G\sqrt{12})$ electrons. As discussed below, this rounding error is small compared to the photon noise, except at low signal levels.

In the early generations of hyperspectral sensors, the combination of dark current and readout noise were a significant or dominating contribution to the total noise. This is exemplified by the evolution in signal to noise ratio over successive upgrades of the AVIRIS sensor [20]. For the early sensors, the often-used assumption of constant additive noise was thus probably valid. Today, however, sensors exhibit signal-dependent noise according to Eq. (4). (For hyperspectral sensors at thermal infrared wavelengths, the spectral contrasts are often small relative to the mean signal level. In these cases the assumption of constant noise still holds to a good approximation even if the sensor is photon noise limited.)

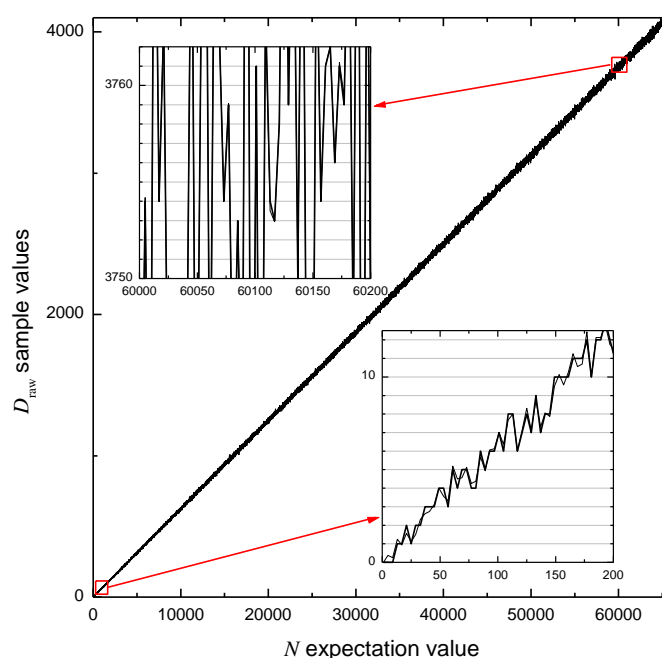


Fig. 3. Synthetic data samples for a sensor where 12-bit raw data correspond to the signal range of a photodetector with full well capacity 2^{16} electrons, resembling a regular silicon-based device. The synthetic data are made by randomly drawing Poisson data points with increasing expectation value. The signal-dependent noise level is clearly seen in the main graph, although it must be noted that the apparent width of the trace here is several times the standard deviation. Insets show magnified parts of the curve at the low and high signal ends, with horizontal lines representing each discrete output value. In the insets, the thin line represents the raw data values without digitization rounding. For low signals, the digitization causes significant rounding errors. At high signal levels, the precision is excessive relative to the noise, and storage space is wasted.

Several kinds of noise and radiometric errors are not part of the noise model here. Examples are offset drift or “hot pixel” photodetector elements with excessive noise, which can cause severe distortions and artifacts in the output of image processing. However these types of noise can generally be avoided or made negligible in state-of-the-art sensors. Therefore the treatment here focuses on the fundamental photon noise.

2.4 Example illustrating photon noise in raw data

Consider a hyperspectral sensor where the photodetector elements have a full well capacity of 2^{16} electrons and the signal is read out as D_{raw} values with 12 bits of precision. Then each step in D_{raw} corresponds to 16 electrons. Figure 3 shows a set of simulated random D_{raw} values drawn over the full range of N , illustrating how photon noise varies with signal level. For a signal level of $N=2^8$ electrons, the standard deviation of photon noise is $\sqrt{2^8} = 16$ electrons, equal to the D_{raw} step size. Below this signal level, rounding errors in D_{raw} start to become non-negligible, and higher precision of the digitization may be needed to avoid degradation of the signal to noise ratio. At the other end of the scale, near the full well signal level, the standard deviation of photon noise is 2^8 electrons, which corresponds to 16 steps in D_{raw} . We see that for high signal levels, resolution is wasted on digitizing noise. Therefore, the raw data representation of hyperspectral images may be suboptimal with respect to either accuracy or data volume or both, depending on the resolution of the digitization.

2.5 Information rate of the hyperspectral sensor

In information-theoretic terms, the sampling of light level is a communication channel with a certain information capacity, measured in bits per sample. The channel introduces a Poisson distribution around the mean signal value. It has been shown [24] that when the sample values vary from zero to an upper limit N_{max} then, in the limit of large N_{max} , an upper bound for the information capacity is

$$\frac{1}{2} \log_2 N_{\text{max}} - \frac{1}{2} \log_2 \frac{\pi e}{2} \approx \frac{1}{2} \log_2 N_{\text{max}} - 1.0 \quad (5)$$

bits per sample. In terms of imaging, N_{max} is the full well capacity. Thus if the full well capacity is 2^W electrons then the information capacity of the imaging system is $W/2-1$ bits per sample in the limit of large N_{max} . In the example of Fig. 3, the information capacity of the sensor is at most 7 bits per sample, consistent with the observation that 12-bit data have excess precision over most of the dynamic range. Thus the information-theoretic limit provides clear motivation to choose more efficient data representations.

3. Conventional and alternative representations for hyperspectral images

The data representation of a hyperspectral image should preserve information about light levels, preferably expressed as standardized quantities and units. Here it is emphasized that the image data should also preferably provide easy access to estimates of noise. Furthermore, the representation should preferably fit in a compact data format. This section discusses how several data representations conform to these requirements.

3.1 Raw data

The raw data D_{raw} can only be processed with the aid of a large set of metadata to account for sensor properties, including the pixel-to-pixel variations described by the dark current and quantum efficiency matrices $I_d[i,j]$ and $\eta[i,j]$. Therefore it is generally difficult to process raw data directly. This representation of hyperspectral images is at best only suitable for systems where processing is very tightly integrated with the sensor.

The raw data are integers that fit within a 2-byte data type, which is how they are normally represented. Of course, bit shifting may be employed to reduce data volume. In the example

above, two 12-bit values can be stored in 3 bytes. Either way, the raw data can be stored in a reasonably compact format, and of course without loss of information. However the data volume is significantly larger than the information capacity of the sensor, as illustrated in the example above.

3.2 Radiance representation

Conventionally, the raw data D_{raw} are converted to estimated values for incoming radiance in a pre-processing step, based on sensor calibration data. In principle, the radiance estimates are found from Eqs. (1) to (3), but in practice the conversion to radiance is based on gain and offset correction matrices $C_1[i,j]$ and $C_2[i,j]$ measured during sensor calibration:

$$\hat{L}[i, j] = \left(\frac{1}{G} D_{\text{raw}}[i, j] - I_d[i, j]t \right) \frac{hc}{\eta[i, j]tA\Omega\Delta\lambda[i]\lambda[i]} \equiv C_1[i, j](D_{\text{raw}}[i, j] - C_2[i, j]t) \quad (6)$$

where \hat{L} denotes estimates of incoming radiance. Thus in terms of the physical sensor parameters, the two calibration matrices can be expressed as

$$C_1[i, j] = \frac{hc}{G\eta[i, j]tA\Omega\Delta\lambda[i]\lambda[i]} \text{ and } C_2[i, j] = GI_d[i, j].$$

The estimated radiance values \hat{L} are possibly the most common representation of hyperspectral images. Represented in this way, the image data can be processed without knowing the details of the sensor.

Signal-dependent photon noise can be estimated from each radiance sample individually by first reconstructing the photoelectron count

$$N[i, j] \approx \hat{L}[i, j]\eta[i, j]tA\Omega\Delta\lambda[i] \frac{\lambda[i]}{hc} + I_d[i, j]t = \frac{1}{G} \left(\frac{\hat{L}[i, j]}{C_1[i, j]} + C_2[i, j]t \right)$$

From this, the photon noise can be estimated using Eq. (6) and expressed as a radiance noise:

$$\Delta L[i, j] = GC_1[i, j]\sqrt{N[i, j]} = \sqrt{GC_1[i, j] \left(\hat{L}[i, j] + GC_1[i, j]C_2[i, j]t \right)} \quad (7)$$

Thus, with knowledge of the calibration matrices $C_1[i,j]$ and $C_2[i,j]$ and the gain G , a data user can estimate the photon noise from image data represented as radiance. A somewhat inelegant aspect of Eq. (7) is that the data user needs the full calibration matrices of the sensor to estimate noise. The next section discusses ways to simplify the metadata needed for noise estimation.

In principle, radiance data should be stored as floating point values (usually 32 bits), both to avoid additional rounding error and because values in standard units rarely fit an integer format. To save storage space, radiance data are usually scaled and rounded to fit a more compact integer format (usually 16 bits), with the scaling factor supplied as part of the image metadata. The scaling sacrifices the (not so important) requirement of standardized data units. The rounding introduces an additional error, which will be discussed in the next section.

It should be noted that the possible range of radiance values from the sensor is not well bounded, because of the large spectral variation of quantum efficiency discussed above. For bands with low quantum efficiency, the saturation level D_{max} corresponds to a very large radiance, as expressed by Eq. (6). Therefore the radiance representation must allow a larger range of values than for the raw data. As an example, consider the case in Fig. 2, where the quantum efficiency at the blue end of the spectrum is about 10% of its peak value. Then, roughly speaking, the blue band can measure a 10 times higher radiance value without saturating. If the data are represented as integers then the bit width must be increased by $\log_2 10 = 3.3$ bits, compared to the raw data, to preserve the full dynamic range of the sensor.

The example sensor with 12-bit raw data would then need a data size of 16 bits to store the data as radiance values without loss of precision. An obvious way around the increase in data size is to use a different scaling factor for each band for conversion of radiance values to integers. The following section introduces such a scheme, with the added benefit of easy noise estimation.

3.3 Radiance-linear “corrected raw data” representation

In an application data flow, the sensor should provide relevant data for processing, but hide uninteresting sensor details. For example, the dark current of individual sensor elements, $I_d[i,j]$, is of little interest in data processing, but the average amount of dark current could be of interest for noise estimation. For the quantum efficiency $\eta[i,j]$, spatial variation within a band is usually small to moderate, while the spectral variation can be large, as discussed above. In the following, a transformation of the raw data is developed that compensates for small and uninteresting sensor effects. The transformation produces “corrected raw data” that can be seen as the output of an idealized sensor, for which the sensor physics is more easily accessible.

First, the overall quantum efficiency $\eta[i,j]$ is split into two parts. Let $\eta[i]$ (with only a band index) represent the average quantum efficiency over all spatial pixels for band i . Thus for a sensor with P spatial pixels,

$$\eta[i] = \frac{1}{P} \sum_{j=1}^P \eta[i,j]$$

Also let $F[i,j]$ be the residual responsivity nonuniformity error defined by

$$F[i,j] = \frac{\eta[i,j]}{\eta[i]}$$

Defined in this way, the mean of the elements in F is 1. Here, F accounts for small differences in loss and gain between individual photodetector elements, as well as for residual variation in the optics throughput AQ which was taken to be constant in Eq. (1). The range of $F[i,j]$ values depends on the details of the sensor, but should be well within $0.5 < F[i,j] < 2$ in practice. With D_{\max} being the full scale of raw data, let $C_{\max} > D_{\max}$ be the new maximum data value for the corrected raw data. Define a scaling factor S so that

$$S = \frac{C_{\max}}{D_{\max}} G \quad (8)$$

By using $F[i,j]$ and $I_d[i,j]$ to correct the data for pixel-to-pixel nonuniformities and then scaling by S , we obtain the corrected raw data representation D_C :

$$D_C[i,j] = \frac{S}{GF[i,j]} (D_{\text{raw}}[i,j] - GI_d[i,j]t) \quad (9)$$

This is proposed here as one alternative representation for hyperspectral image data. $D_C[i,j]$ can be considered to be scaled photoelectron counts of an idealized sensor without pixel-to-pixel nonuniformities and with gain S (noting that $D_C[i,j] \approx SN[i,j]$). Processing algorithms for radiance can be applied directly to the D_C data, with appropriate rescaling of spectrally varying algorithm parameters, since D_C is proportional to \hat{L} .

When needed, the $D_C[i,j]$ values are readily converted to radiance estimates:

$$\hat{L}[i,j] = \frac{hc}{S\eta[i]tA\Omega\lambda[i]\lambda[i]} D_C[i,j] \equiv K[i]D_C[i,j] \quad (10)$$

where

$$K[i] = \frac{hc}{S\eta[i]tA\Omega\Delta\lambda[i]\lambda[i]}$$

Comparing with Eq. (6), we see that conversion to radiance is reduced to applying a band-specific scaling factor $K[i]$.

The noise level varies between photodetector elements as described by the responsivity variations $F[i,j]$ and dark current matrix $I_d[i,j]$. However, it can be noted that estimation of noise can normally be permitted to be less accurate than estimation of radiance. Therefore, it is generally acceptable to neglect the pixel-to-pixel variations of $F[i,j]$ (whose mean is 1) in estimation of photon noise. Also, variations in gain G between photodetector elements is usually small and therefore neglected here. To estimate the total noise, dark current and readout noise should also be considered. Unless there are large nonuniformities of dark current $I_d[i,j]$ across the detector array, the noise contribution from dark current can be represented by an average value. The variance contributions from dark current and readout noise δN can be combined in a parameter N_0 :

$$N_0 = \bar{I}_d t + (\delta N)^2$$

where \bar{I}_d is the mean of the dark current matrix $I_d[i,j]$. We can then calculate an effective photoelectron count which is actually an approximation to the total signal variance:

$$N_{eff}[i, j] = \frac{D_C[i, j]}{S} + N_0 \quad (11)$$

where the nonuniformities described by $F[i,j]$ and $I_d[i,j]$ have been neglected. Under this approximation, an estimate of the noise amplitude, in photoelectrons, is obtained from the estimated standard deviation

$$\Delta\hat{N}[i, j] = \sqrt{N_{eff}[i, j]} \quad (12)$$

The equivalent radiance noise can then be estimated from

$$\Delta L[i, j] = K[i]S\Delta\hat{N}[i, j] = K[i]S\sqrt{\frac{D_C[i, j]}{S} + N_0} \approx K[i]\sqrt{SD_C[i, j]}$$

where the approximation is valid when the dark current contribution is small.

Thus the D_C data can be straightforwardly converted to estimates of radiance *and* noise, given the simple parameter set $K[i]$, S and, when needed, N_0 . In place of $K[i]$, the image metadata can preferably include explicit values for the factors entering in $K[i]$: $\eta[i]$, t , A and Ω . Of course, values for $\Delta\lambda[i]$ and $\lambda[i]$ are also supplied with the image. Then the image metadata not only enable calculation of radiance and noise, but also provide the user with a first-order model for the sensor.

Now consider storage of D_C values. If F_{min} is the minimum value in the responsivity correction matrix $F[i,j]$, then the possible range of D_C values is

$$0 \leq D_C[i, j] \leq \frac{SD_{max}}{F_{min}}$$

To adapt the D_C data to an n -bit format, choose $C_{max}=2^n-1$, which determines S from Eq. (8). Observe that the D_C data range is unaffected by the quantum efficiency variations described by $\eta[i]$, and that $F_{min} \approx 1$. Therefore the range of values for $D_C[i,j]$ is well bounded, to a much smaller range than for radiance values at a comparable level of rounding error.

Rounding of the D_C values introduces an error with RMS amplitude $1/\sqrt{12}$ on the D_C scale [17]. On the same scale, the digitization error in the raw data is $C_{\max}/(D_{\max}\sqrt{12})$ or, equivalently, $S/(G\sqrt{12})$. The relative increase in the total amount of rounding error is then

$$\frac{\sqrt{(1/\sqrt{12})^2 + [C_{\max}/(D_{\max}\sqrt{12})]^2}}{C_{\max}/(D_{\max}\sqrt{12})} = \sqrt{1 + \frac{D_{\max}^2}{C_{\max}^2}} \approx 1 + \frac{1}{2} \frac{D_{\max}^2}{C_{\max}^2}. \quad (13)$$

If, for example, the D_C data are stored in a format 1 bit wider than D_{raw} then $C_{\max}=2D_{\text{raw}}$ and the rounding of D_C values increases the total rounding error by $\approx 12\%$. This slight increase should be unproblematic, assuming that digitization noise in the raw data is already small compared to the photon noise.

Equation (13) disregards the variations in quantum efficiency between photodetector elements described by $F[i,j]$. The element with the highest value of $F[i,j]$, denoted F_{\max} , experiences the largest additional rounding error, as seen from Eq. (9). The error due to rounding of D_C values remains less than the digitization noise in D_{raw} as long as $C_{\max} > F_{\max} D_{\max}$. Note that as long as this condition holds, it is possible to reconstruct the raw data exactly, using the full calibration data, since the steps in D_C are smaller than the steps of D_{raw} on a radiance scale. Therefore, sensor data may be stored directly as D_C data during recording *without any loss of information*, even if the actual raw data are not stored. For the example sensor above with 12-bit raw data, the D_C data could be stored losslessly in 13 bits even if F_{\max} were as large as 2.

In summary, the corrected raw data D_C can be processed directly and are easily converted to radiance and to estimates of sensor noise. The required metadata are simple parameters of a first-order sensor model, in itself informative for the data user. With a very modest increase in data size over the raw data, on the order of 1 bit per sample, the D_C representation is fully lossless. This data representation has been implemented as an option for data storage in the HySpex line of commercial hyperspectral sensors [15]. The sensor software also provides streaming D_C data in real time. This data stream is used for data processing in FFI's airborne real-time processing system [14].

3.4 Image data representation with variance stabilization

Figure 3 illustrates that any representation proportional to radiance, including the D_C representation, will be inefficient with respect to storage space, due to the signal-dependent noise. By applying a *variance stabilizing transform* to the data, the noise can be made approximately signal-independent. A well known variance-stabilizing transform is the Anscombe transform [25]. For the Poisson-distributed quantity N , the Anscombe transform is $N \rightarrow \sqrt{N + \frac{3}{8}}$. The resulting transformed variable has variance close to 1/4 for values of data mean down to about 2. In the case of hyperspectral data, the photoelectron counts are usually hundreds or more. With the simpler transform [26] $N \rightarrow \sqrt{N}$, the variance is close to 1/4 for a data mean down to about 10, covering the relevant range of values for hyperspectral imaging.

Variance-stabilizing transforms based on the square root have been extensively used in Astronomy. One main area of application is data compression [27–31]. Another application of variance-stabilizing transforms in astronomy is as a step in image restoration for Poisson images [32–34]. By appropriate scaling and rounding of variance-stabilized data, a predetermined fraction of the noise variance can be retained. Such a scheme is outlined here. Note that the resulting integer data can then be compressed with standard lossless techniques, resulting in a significant degree of data compression. Some further discussion of variance-stabilizing data transformation is given in 5.3.

Following Refs. 30 and 32, a variance-stabilized representation of the hyperspectral image, which also takes into account dark current and read noise, is obtained from

$$R[i, j] = S_R \sqrt{N_{\text{eff}}[i, j]} = S_R \sqrt{\frac{D_C[i, j]}{S} + N_0}. \quad (14)$$

This representation is equal to the noise estimate in Eq. (12) apart from the scaling factor S_R , which will be used to rescale the data for storage, as S does for D_C above. The statistical distribution of the transformed data R is very close to Gaussian with variance $S_R^2/4$, as can be easily verified by numerical tests. This is a very interesting property of the R representation: As opposed to radiance-linear representations, the nonlinearly transformed R data conform to the often-used assumption of signal-independent additive Gaussian noise.

The R data can be readily converted to radiance and to estimates of radiance noise. Following Eqs. (10) and (14),

$$\hat{L}[i, j] = \frac{hc}{\eta[i]tA\Omega\Delta\lambda[i]\lambda[i]} \left[(R[i, j]/S_R)^2 - N_0 \right] \equiv SK[i] \left[(R[i, j]/S_R)^2 - N_0 \right].$$

The noise expressed as radiance is

$$\Delta L[i, j] = SK[i] R[i, j] / S_R \quad (15)$$

so that the radiance signal to noise ratio is

$$\frac{\hat{L}[i, j]}{\Delta L[i, j]} = \frac{SK[i] \left[(R[i, j]/S_R)^2 - N_0 \right]}{SK[i] R[i, j] / S_R} \approx \frac{R[i, j]}{S_R}.$$

As for the D_C representation, the user can reconstruct radiance and noise from the R data with a simple set of metadata that also give a first-order description of the sensor.

Now consider storing the variance-stabilized data R as integers, after rounding. The scaling factor S_R should be chosen so that the rounding error is negligible. Due to the nonlinear transformation, rounding error in R cannot be referred to digitization error in D_{raw} . Instead, it must be compared to the photon noise. The relative noise contribution of the $1/\sqrt{12}$ RMS rounding error in R , here denoted r , is

$$r = \frac{\text{rounding error}}{\text{photon noise}} = \frac{\left(R + \frac{1}{\sqrt{12}}\right)^2 - R^2}{S_R R} = \frac{2}{\sqrt{12} S_R} + \frac{1}{12 S_R R} \approx \frac{2}{\sqrt{12} S_R}$$

Thus the rounding error is independent of signal level, except for very weak signals which were disregarded in our choice of variance-stabilizing transform. Choosing $S_R=1$ gives $r=0.58$. Then the RMS sum of photon noise and rounding error is 15% larger than photon noise alone, which may not be acceptable. Choosing $S_R=2$ gives $r=0.29$ and a 4% higher noise, which should be tolerable in many cases: As pointed out in [31], the added error with $S_R=2$ can be compensated by an 8% increase in exposure time. Expressed this way, the consequences of any given choice of S_R can be readily assessed. The particular choice $S_R=2$ transforms the photon noise conveniently into a constant variance and standard deviation of 1.

Note that as shown in [31], rounding of the nonlinearly transformed R values introduces a small bias in the reconstructed radiance corresponding to $1/12$ photoelectron for $S_R=2$, which is negligible in most cases but can easily be corrected for.

Consider again the CCD from the example above, with full well capacity $N_{\text{max}}=2^{16}$. We note that with $S_R=1$, the full-scale value of R is 2^8 and the image data can be fitted in a 1-byte format. With $S_R=2$, the data can be stored in 9 bits per sample with only a 4% degradation of

SNR. Thus the variance-stabilized R representation is quite efficient with respect to data storage and approaches the minimum data size of 7 bits given by the information capacity (5).

4 Possible sensor-related refinements of the data representation

4.1 Low photoelectron counts

The addition of N_0 in Eq. (14) incorporates the noise from dark current and readout so that the noise estimation in Eq. (15) is valid to very low signal levels. At the same time, this definition of R is consistent with the recommended practice to keep the zero-signal point within the data range [18] to be able to detect offset drift errors. The D_C representation, on the other hand, is defined to be proportional to radiance, and the dark current term is subtracted in Eq. (9). When the signal is comparable to or smaller than the dark current then the subtraction of a mean dark current may lead to a negative output value for some samples. In many cases these negative values can be truncated to zero for storage in an unsigned integer format. However such truncation would tend to skew the data distribution in a way that is not consistent with the sensor physics. An alternative definition of D_C would be to incorporate an offset term, for example N_0 , so that zero signal is represented by a small positive D_C value.

4.2 Large constant background level

In some cases, image contrasts are small compared to the mean photosignal. For example, this is normally the case in the thermal infrared. A similar situation arises when the dark current contribution is large. In these cases, the values of R will vary over a small range relative to their mean. Then essentially the R values have a large constant offset or pedestal value R_{min} , which is undesirable for storage in a compact format. However this offset can be subtracted and supplied as metadata, resulting in modified data $R \rightarrow R - R_{min}$ that fit in a more compact format. The modified data still maintain the useful property that the noise level is constant and known.

4.3 Saturation

In many cases there is a risk that some sensor elements will saturate due to strong signals, for example a specular reflection in the scene. In such cases, the sensor reading is only a lower bound on the light level. Then the data for the particular pixel and band cannot be correctly processed along with normal image data. Normally the raw data from a saturated sensor element is the value D_{max} . Note that if this value is converted into any of the representations discussed above, it may result in an output value falling within the range of normal unsaturated data. Therefore the saturated light samples need to be treated specially.

Saturated values can be flagged in a separate map of bad pixels supplied as part of the metadata. Alternatively, saturation can be flagged by a reserved data value. A reasonable choice, used in our real-time system, is to reserve the maximum data value in an integer data type as a saturation flag. Then the data structure is kept simple. For the D_C representation in an n -bit format, using $C_{max}=2^n-2$ guarantees that the maximum value will never be used for normal data.

4.4 Bad sensor elements

In most types of photodetector arrays, a small fraction of the elements are defective in some way, such as excessive dark current, excessive noise, or no signal at all. In conventional imaging, data from these pixels are replaced by interpolating spatially between neighboring pixels. When the array is used for hyperspectral imaging, such interpolation may produce unacceptable spectral artifacts depending on the specifics of the sensor and application. Therefore it is desirable to flag defective pixels in the image, for example in the form of a reserved data value. This reserved value may be the same as the saturation flag. However a

separate reserved value for defective pixels will preserve information about saturation, which is potentially usable in the data processing.

4.5 Known upper limit for the signal spectrum

It can be noted that a further stage of physics-based compression is possible if the maximum signals of some bands are guaranteed to be below a known fraction of the saturation level. Then the allocated bit width can be reduced accordingly for bands with smaller signal ranges. An example is remote sensing with a visible and near-infrared grating-based imaging spectrometer with spectral properties similar to Fig. 1. Consider in addition the fact that the peak of solar irradiance is in the middle of the spectral range. The integration time will be set so as to avoid saturation in the bands with the largest photoelectron count, and then the bands near the ends of the spectral range will span only a small part of their possible signal range.

5. Summary and discussion

5.1 Noise and data processing

The review of data compression papers in the introduction illustrates that published works on hyperspectral image processing for remote sensing rarely exploit noise information from basic sensor physics. There is no clear technical reason why such noise information is not being provided and used. I suspect that the situation is partly a matter of tradition in the field, dating from the first generations of hyperspectral sensors. The situation is different in other fields of optical imaging, for example in astronomy where photon noise is taken into account in data compression and image processing [27–34]. In the field of thermal imaging, photon noise is the basis of the standard “background limit” on the signal to noise ratio [17].

All radiance-linear representations are in principle equivalent and share the problem that the noise level is signal-dependent. This has implications for algorithms containing spectral difference metrics. Many such metrics, such as Euclidean distance or integrated difference, are based on band-by-band differences between spectra. For example, image compression algorithms often use radiance mean square error (MSE) or similar measures of compression quality [6] and assume additive noise. If photon noise makes a significant contribution to the reconstruction error after compression, which is a desirable situation, then the radiance MSE will tend to be dominated by the brightest parts of the image since photon noise is largest there. Thus there is a risk that systematic reconstruction errors in darker parts of the image will not be captured by the MSE compression quality metric. Similar problems could arise in other types of image processing. A possible alternative metric for spectral difference could be the mean error or mean square error in variance-stabilized data $R[i,j]$, which would tend to introduce balance between bright and dark areas. There is an obvious potential to reformulate existing algorithms to take advantage of square root transformed data. An important limitation of the variance-stabilized representation is that the image data do not conform to the assumptions of the often-used linear mixing model.

5.2 Storage and compression

It has been pointed out here that any radiance-linear representation of hyperspectral images is inefficient with respect to data size. In the limit of low dark current and large full well capacity, the data size is a factor ~ 2 larger than the information capacity of the sensor. When the data are represented as radiance, data size may increase further due to a large relative variation in quantum efficiency with wavelength. In comparison, the variance-stabilized R data can be stored in a data volume comparable to the information capacity, depending on the desired accuracy. In the example case used throughout the text, the raw data are 12 bits, D_C data would require 13 bits per sample for lossless storage and radiance data would require 16 bits. In comparison, it was shown that variance-stabilized R data could be stored in 9 bits, although with a small (4%) degradation of the signal to noise ratio. It is thus possible to

achieve a near-lossless compression on the order of 50% just by transforming the data based on basic measurement physics. Of course, a much higher compression level can be achieved by proper compression algorithms that exploit redundancies in spectral, spatial or temporal image information. However, it appears reasonable to say that square-root transformation followed by generic lossless data compression is an appropriate baseline for calculation of data compression ratios.

5.3 The many virtues of the square-root transformation

In other types of imaging, the square-root representation is well established. Its ability to stabilize photon noise variance is just one of the reasons why the square root transformation is widely used. Conventional digital photography and video employ a square-root-like transformation known as “gamma correction”. (See for example Refs. 35 and 36.) It is thanks to the gamma transformation that 8-bit precision is sufficient for digital photography. Historically, gamma correction originates from analog video formats, where the nonlinear transformation adapted the signal to the characteristics of cathode ray tubes (old TV displays). At the same time, the transformation equalizes the effect of additive noise over the dynamic range of the analog signal, just as rounding errors are equalized in a digital signal. Furthermore, as discussed in [35], there is an interesting coincidence between the square root transformed data and the human visual perception, such that the data are an approximate metric for perceived contrast. Finally, the square root transformation adapts the data for storage in a data volume comparable to the Poisson channel capacity. Thus there are several fundamental and coincidental reasons to say that a variance-stabilized representation is the natural choice for Poisson-dominated image data. Returning to hyperspectral imaging, it is clear that the gamma characteristic of display systems must be taken into account when visualizing radiance data.

6. Conclusions

Knowledge of physical sensor noise tends to remain unexploited in hyperspectral image processing. Also, representation of hyperspectral images as radiance values is inefficient with respect to data volume. Two alternative representations for hyperspectral images, “corrected raw data” D_C and variance-stabilized data R have been proposed here. These representations facilitate the use of physical noise estimates in processing of hyperspectral data. There is an obvious potential in reformulating many image processing algorithms to work on variance-stabilized data. The variance-stabilized representation also allows compact data storage approaching the information-theoretic limit, and is an appropriate baseline for evaluation of hyperspectral image data compression. If knowledge of sensor noise is not taken into account, there is an obvious risk that image processing results are suboptimal, and that data volumes are much larger than necessary.