Information capacity as a figure of merit for spectral imagers: the trade-off between resolution and coregistration

Torbjørn Skauli

Norwegian Defence Research Establishment (FFI), P.O. Box 25, Kjeller 2027, Norway (torbjorn.skauli@ffi.no)

Received 16 October 2012; revised 24 January 2013; accepted 24 January 2013; posted 24 January 2013 (Doc. ID 178165); published 11 February 2013

The performance of spectral imagers is customarily described by several characteristics including resolution, noise, and coregistration. These must be traded off against each other in a practical imager design. This paper proposes a way to use the information capacity, in an information-theoretic sense, as a figure of merit for spectral imagers. In particular, it is shown how a metric [Opt. Express **20**, 918 (2012)] can be used to incorporate coregistration performance in a definition of total noise, which in turn can be used in the definition of information capacity. As an example, it is shown how the information capacity can be used to optimize the pixel size in a simple case that can be treated analytically. Generally, the information capacity is attractive as a fundamental, application-independent figure of merit for spectral imager optimization and benchmarking. © 2013 Optical Society of America

OCIS codes: 110.4234, 220.4830, 110.3055.

1. Introduction

The performance of a spectral imager is normally specified by an extensive set of performance characteristics. There is at present no widely accepted way to combine performance characteristics into a single figure of merit for benchmarking or design optimization. The performance can in principle be characterized using end-to-end simulations, which must include details of the scene and the image processing to be representative of the system performance. However, such simulations are a complex process closely coupled to application details.

An imaging system can be considered as a communication channel in the information-theoretic sense, where information about the imaged object is transferred to the digital image data output. This channel will be characterized by the information capacity, which is the upper limit on the ability of the imager to collect information about the scene. Although it has never become a mainstream way of characterizing imaging systems, there has over the years been a fair number of papers discussing the information theory of imaging systems (see, for example, [1–6]). However, these papers have generally only considered the monochrome imaging case. References [5,6] use an information-theoretic approach to derive a single figure of merit for conventional color cameras.

Particular to spectral imaging is the need for spatial coregistration of multiple bands. There will inevitably be some degree of coregistration error resulting from imperfections in the imaging process, such as chromatic aberrations or inaccuracies in scan movement, depending on the type of spectral imager. Early works [7] stipulated that coregistration error should not affect the recorded signal by more than 5%, which is still large compared to the noise level, on the order of 1% or better for current sensors. It is common to see coregistration specifications of 10% and more, both in the scientific literature and in commercial imagers, particularly for imagers with a high pixel count. There is at present no widely standardized way to fully characterize spatial coregistration errors. Imaging spectrometers are customarily characterized in terms of "keystone" distortion,

1559-128X/13/070C58-06$15.00/0

which measures only position differences between bands and not differences in their point spread function (PSF) shape, and the corresponding “smile” distortion in the spectral dimension. Recently, however, a simple metric has been shown to express the combined effect of all types of coregistration error [8].

For a given amount of coregistration error, coregistration will tend to be improved if the detector elements that define the pixel size are made larger, but at the expense of lower spatial resolution. This is illustrated in Fig. 1 for a very simple case with two bands, whose pixel footprints in a scene are shown with red and blue rectangles. On the left is a case with two spatial pixels and a spatial coregistration error that is a large fraction of the pixel size, leading to significant errors in the recorded “spectrum” (with only two bands here). On the right is a case where the pixel footprint is twice as large, with the same amount of distortion. The latter case has a smaller amount of coregistration error in the recorded spectrum, as well as lower noise due to collection of more light, but only half as many pixels. In both cases it is assumed that the pixel spectrum is processed as if there was no coregistration error, an assumption that is tacitly made in practically all hyperspectral image processing. The issue discussed in this paper is basically how to choose between these two cases. In the design or characterization of a spectral imager, it is not immediately clear how to balance coregistration performance against resolution, unless very specific application requirements are given. This trade-off is an example of the difficulty in combining different characteristics into a single figure of merit for spectral imagers.

Fig. 1. (Color online) Illustration of the trade-off between coregistration error and pixel count. The figure illustrates the pixel footprint for two bands, shown as red and blue rectangles, for two cases. There is a fixed amount of spatial distortion, caused by the imaging optics or other imperfections in the imaging process. On the left is a case with small pixels and a spatial coregistration error that is a large fraction of the pixel size. On the right is a case with larger but fewer pixels. Then the relative amount of distortion is reduced, and also the signal-to-noise ratio improves. The scene is assumed to contain contrasts on a wide range of spatial scales, illustrated here by an urban landscape. Here it is not obvious how to make the important trade-off between coregistration and spatial resolution in spectral imaging. This paper proposes information capacity as a relevant figure of merit.

This paper proposes that information theory, in combination with the coregistration metric, can provide a figure of merit for spectral imagers, applicable to the trade-off between coregistration and pixel count. The coregistration metric is used to express the effect of coregistration error as an added noise, which can be incorporated in a model for the information capacity of the imager. To illustrate the concept, an analytic solution is derived for the simple task of optimizing the detector pixel size, keeping the imaging optics (and other image distortions) unchanged. It is argued that the treatment can be extended to include spectral coregistration error between bands, to become a single performance metric that encompasses many characteristics of spectral imagers.

2. Signal Model

Consider a single pixel in a single spectral band of a spectral imager. The basic output signal is the number of excited photoelectrons N for a given incoming spectral radiance L:

$$N = L\eta t A\omega\Delta\lambda\frac{\lambda}{hc}.$$

Here η is the quantum efficiency, t is the integration time, A is the area of the entrance pupil of the imager, ω is the solid angle subtended by the pixel in the scene, $\Delta\lambda$ is the spectral bandwidth, λ is the wavelength, and λ/hc is the photon energy. If the mean radiance level in the scene is $\bar{L}$ and the imager has P pixels in its field of view, the total number of photoelectrons for all pixels in the band is

$$N_{\mathrm{tot}} = P\bar{L}\eta t A\omega\Delta\lambda\frac{\lambda}{hc}.$$

The mean number of photoelectrons per pixel is then

$$\bar{N} = \frac{N_{\mathrm{tot}}}{P}.$$

The fundamental noise mechanism is Poisson fluctuations in N. The mean of this “photon noise” can be estimated as

$$\Delta N_{\mathrm{phot}} \approx \sqrt{\bar{N}}. \quad (1)$$

Dark current and readout noise are neglected here for simplicity but would have to be included for cases where they become significant contributions to ΔN_{phot}.

Coregistration imperfections will introduce errors in the signal, depending on the scene. In the common case where the pixel contains an inhomogeneous mixture of scene materials, the weighting of the materials in the pixel signal may differ from band to band due to spatial coregistration error. As shown in [8], the maximum weighting error between two bands i and j in a pixel p is given by the metric

$$\varepsilon_{s,ijp} = \frac{1}{2}\iint_{x,y} |f_{jp}(x,y) - f_{ip}(x,y)|\mathrm{d}x\mathrm{d}y. \qquad (2)$$

Here, f_{ip} and f_{jp} are the normalized sampling point spread functions (SPSFs) for the recording of light at pixel p in band i and j, and the integration is over the image plane. The overall coregistration performance can be expressed as an average of Eq. (2) over all bands and pixels, denoted $\bar{\varepsilon}_s$.

Imaging systems are normally designed such that the SPSF for one pixel overlaps somewhat with that of its neighbors (since nonoverlapping SPSFs would imply spatial subsampling of the image plane). Then the between-band SPSF differences characterized by Eq. (2) will introduce differences in the signal influence from scene materials present in neighboring pixels [8,9]. Therefore, the effect of coregistration errors on the signal can be approximately modeled by assuming proportionality with the amount of nearest-neighbor contrast in the image. Let the mean difference between nearest-neighbor pixels be $\alpha\bar{N}$ with $0 < \alpha < 1$. The mean amplitude of signal errors due to misregistration can be estimated by

$$\Delta N_{\mathrm{coreg}} \approx \bar{\varepsilon}_s \alpha \bar{N}. \qquad (3)$$

This approximation may be somewhat coarse, but in the limit of a uniform scene it becomes exact, since $\alpha = 0$, and it is clear from physics that there can be no signal error due to misregistration in the optics. Note that it may be reasonable to use a higher value of α than the mean neighbor difference, to account for the higher risk of signal errors near material boundaries in the scene, which typically make up only a small fraction of the image area. If a value $\alpha = 1$ is chosen, Eq. (3) expresses the largest possible error for any scene under the assumption [Eq. (3)]. (It is conceivable to have an even larger error in cases where the PSF is smaller than the pixel sampling interval, which is unlikely in a practical imager design.) The signal contamination due to coregistration will here be represented as additive Gaussian noise with zero mean and standard deviation $\bar{\varepsilon}_s \alpha \bar{N}$.

3. Estimating the Information Capacity

Consider the task of choosing the optimal pixel size for a spectral imager where at least one spatial dimension is imaged by a photodetector array. Examples include the imaging spectrometer (one spatial dimension imaged, the other scanned) or the filter wheel camera (two-dimensional spatial imaging with sequential recording of bands). The array determines the pixel size and the number of pixels. Assume that the imager has a total field of view Ω divided into P spatial pixels. Regardless of the measurement concept employed, the pixel field of view ω tends to vary as

$$\omega = \frac{\Omega}{P}.$$

Assume that the pixel size is increased by binning or by increasing the detector element size, while keeping the imaging optics unchanged. This will tend to reduce the coregistration error ε_s inversely proportional to the change in pixel size. As a result, the error in the signal is reduced, and the information collection capacity of a single pixel tends to increase. On the other hand, the reduction in the total number of pixels tends to decrease the total amount of information collected by the imager. In the following, it is shown how information capacity optimization can be used to choose the pixel size in this simple model case.

The metric $\bar{\varepsilon}_s$ specifies coregistration performance for the complete imager at a particular pixel size. The performance of the optics can be expressed independently of the pixel size by the ratio

$$P_{\mathrm{lim}} = \frac{P}{\bar{\varepsilon}_s},$$

which will tend to stay constant as the pixel size is varied. Approximately at this "limiting number of pixels," the coregistration error would be $\bar{\varepsilon}_s = 1$. The signal distortion due to coregistration error can then be expressed as

$$\Delta N_{\mathrm{coreg}} = \bar{\varepsilon}_s \alpha \bar{N} = \frac{P\alpha\bar{N}}{P_{\mathrm{lim}}} = \frac{\alpha N_{\mathrm{tot}}}{P_{\mathrm{lim}}}.$$

The ratio of coregistration error to photon noise becomes

$$\frac{\Delta N_{\mathrm{coreg}}}{\Delta N_{\mathrm{phot}}} = \frac{P}{P_{\mathrm{lim}}}\alpha\sqrt{\bar{N}} = \frac{\alpha\sqrt{PN_{\mathrm{tot}}}}{P_{\mathrm{lim}}}. \qquad (4)$$

The ratio [Eq. (4)] should be less than 1 for the sensor to approach ideal behavior. This expression summarizes several of the trade-offs in the design of a spectral imager. The relative importance of coregistration error tends to increase with increasing number of pixels. Also, the coregistration error increases linearly with the signal level, faster than the photon noise, so that the relative impact of coregistration is strongest when the signal is large. (Not captured in this expression is the coupling between coregistration error and the numerical aperture of the lens, which affects the signal level.) An important point here is that Eq. (4) does not provide a criterion for optimal choice of the pixel size in the presence of a given amount of coregistration error, since it does not account for the utility of a higher number of pixels.

Assume that the pixels are binned spatially in groups of b pixels, or that the area of the photodetector elements is changed by a factor b. Then $b < 1$ can be taken to represent a reduction in element size. After binning or element resizing, the photoelectron count changes proportionally to b. The coregistration error $\bar{\varepsilon}_s$ tends to vary inversely proportionally to b. The resulting signal error will depend on b in a way that depends on the spatial frequency distribution of the scene, so that α becomes a function of b. However, for natural scenes, contrasts tend to vary inversely proportionally to spatial frequency [10]. Thus for the purposes of formulating an illustrative model, it is not unreasonable to assume α to be independent of b as long as the size of the binned pixels is larger than the width of the PSF of the optics.

To model the imaging process as a communication channel, the pixel count is analogous to the bandwidth in a regular communication channel, and the total noise has contributions from photon noise and coregistration error. The amount of information that can be collected by a single pixel in a single band is determined only by the signal level (which also determines photon noise) and coregistration performance (represented as added noise), according to some function $C(\bar{N}, \bar{\varepsilon}_s)$, which gives the information capacity in bits. A readout of all pixels in the band then produces $PC(\bar{N}, \bar{\varepsilon}_s)$ bits of information.

If P_0, $\bar{N}_0$, and $\bar{\varepsilon}_{s0}$ are taken to represent a reference case for which $b = 1$ and the information capacity per pixel is $C(\bar{N}_0, \bar{\varepsilon}_{s0})$, then the information capacity after binning by a factor b becomes

$$C(N_0, b) = \frac{P_0}{b} C\left(b\bar{N}_0, \frac{\bar{\varepsilon}_{s0}}{b}\right). \tag{5}$$

for one band.

For information channels with stationary additive Gaussian noise, the capacity is given by the well-known Shannon theorem. This theorem is at best is an approximation for optical imaging, where the noise is signal dependent according to Eq. (1). For channels where Poisson noise is dominating or significant, a similar theorem does not exist. However, several works have provided upper and lower capacity bounds for channel models relevant to optical imaging, including [11–13]. Reference [12] gives bounds for a Poisson channel with a dark current. Reference [13] gives bounds for a Poisson channel approximated as a signal-dependent Gaussian channel, with additive Gaussian noise. The applicability of these bounds as approximations to the channel capacity needs further investigation for practical use. It may be necessary to resort to numerical calculation of channel capacity [14], due to the lack of an accurate analytical capacity model for relevant channel characteristics.

4. Analytical Optimization in an Example Case

As an example, assume a hypothetical imaging spectrometer with a spatial coregistration error of $\bar{\varepsilon}_s = 0.15$, representative of the "keystone" distortion specified for several sensors currently used for remote sensing. The value of α can be estimated from real images or from some model of image contrasts, depending on the application. For example, $\alpha \approx 0.05$ for a typical remote sensing scene with resolved shadow areas recorded by an airborne hyperspectral sensor in the visible and near infrared (VNIR) spectral range. For this value of α, the ratio [Eq. (4)] is unity for a mean signal of $\bar{N} \approx 20,000$, representative of commonly used VNIR hyperspectral imagers. With the two equal noise contributions, the image can be considered to have a total signal-to-noise ratio of about 100. At higher signal levels, the signal-to-noise ratio observed on a uniform scene area will be higher. However, in average over the image, the effect of coregistration error will then be larger than the photon noise. An increase in detector pixel size, keeping the same imaging optics, will tend to reduce both the photon noise and the coregistration error, but at the expense of lower spatial resolution and fewer pixels.

The optimization of a spectral imager can be illustrated by using published analytical capacity bounds as approximations to the channel capacity C. For the purpose of the analysis here, the most relevant case is the lower capacity bound for a Poisson channel with a constrained mean value from [12] and [13]. When inserting $\varepsilon_s \alpha \bar{N}$, respectively, as the variance of the dark current and of the additive Gaussian noise, the lower bounds for the mean-constrained case in both these papers reduce to the same expression for the channel capacity of a single pixel in a single band when $\bar{N} \gg 1$:

$$C(\bar{N}, \varepsilon_s) \geq \frac{1}{2} \log \bar{N} - \bar{\varepsilon}_s \alpha \sqrt{\frac{\pi \bar{N}}{2}}, \tag{6}$$

given in nats/pixel. (1 nat = 1.44 bits.) This bound is not tight in the limit of large added noise, where its value can become negative, but for purposes of illustration it is used as a model of the channel capacity here.

Figure 2 shows the resulting information capacity estimate for one spectral band, for particular choices of $\bar{\varepsilon}_{s0}$ and b in the example case. Here a value $\alpha = 0.15$ has been used to emphasize signal integrity at the scene edges more strongly than using the mean neighbor difference value from the example above. The figure illustrates how the information capacity reaches an optimum for a particular pixel count if the distortions of the optics are held fixed. If the distortions in the optics are large, it is clearly beneficial to improve coregistration by making the pixels larger, even if the pixel count is correspondingly reduced. The figure illustrates that coregistration error can lead to a large loss of information.

5. Extension to Spectral Coregistration and an Overall Figure of Merit

In the spectral dimension, it is important that all pixels in a given band exhibit the same spectral response. For many types of spectral imagers, a

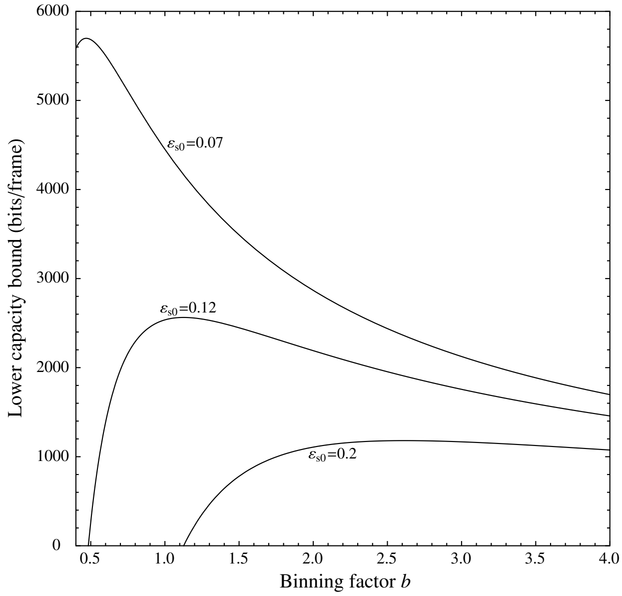

Fig. 2. Information capacity for one band of a spectral imager, with varying sizes of the detector pixels. This illustrates how information-theoretic considerations can be used to determine the optimal spatial resolution in a spectral imager by optimizing the information capacity. Different curves represent different degrees of coregistration error in the imaging optics. The difference between the curves illustrates the large information loss that can result from imperfect coregistration. The plot uses the approximate capacity model [Eq. (6)] and assumes the following parameter values at $b = 1$: $\bar{N} = 20,000$, $\alpha = 0.15$, $P = 1000$.

trade-off exists between this spectral coregistration and the number of bands. In [8], it is discussed how a metric similar to Eq. (2) can be applied in the spectral dimension. Signal errors resulting from imperfect registration of spectral bands between pixels can be estimated in an analogous way to Eq. (3). These errors can be added as an independent noise contribution in the same way as for spatial coregistration error above. Then the information capacity can be taken as a figure of merit even for the spectral resolution trade-off. In analogy with pixel size variation, let s be a spectral binning factor. Let B_0 be the number of bands for $s = 1$ and $\bar{\varepsilon}_{\lambda 0}$ be the mean spectral coregistration error [8]. $\bar{N}_0$ is the mean photoelectron count for $s = 1$ and $b = 1$, for simplicity assumed to be the same in all bands. Then, in analogy with Eq. (5), the information capacity of the imager, incorporating all bands and all pixels, will tend to vary as

$$C(N_0, b, s) = \frac{P_0}{b}\frac{B_0}{s} C\left(bs\bar{N}_0, \sqrt{\left(\frac{\bar{\varepsilon}_{s0}}{b}\right)^2 + \beta\left(\frac{\bar{\varepsilon}_{\lambda 0}}{s}\right)^2}\right), \tag{7}$$

where the spectral and spatial contributions to coregistration error are assumed independent and β is a factor representing the relative strength of spectral coregistration error. This outlines how information capacity could be used to optimize both the spatial and spectral resolution. However, the phenomenology of signal variation along the spectral dimension is very different from that of spatial contrasts and is strongly dependent on the spectral range and resolution. Therefore, the value of β as well as the validity of Eq. (7) will have to be determined for the applications of interest, and a detailed treatment is not given here.

6. Discussion and Conclusions

Note that through $\bar{N}_0$, the information capacity [Eq. (7)] also characterizes the light collection efficiency of the imager expressed by η, A, and ω. By summing over all bands, using proper wavelength dependencies of $\bar{L}$ and η, one obtains a total information collection capacity that captures the effect of signal-to-noise, spectral/spatial coregistration error, number of bands, and number of pixels in one figure of merit. The overall information capacity should also take into account image blur introduced by the SPSF. The application is represented by the simple parameters $\bar{L}$, α, and β, making the information capacity a fairly generic measure of performance. This could be useful either for comparative benchmarking of instruments for a given application or for overall design optimization. A detailed discussion of such a figure of merit will not be undertaken here.

The main point of this paper has been to argue that the combination of the coregistration metric with information theory enables estimation of the information capacity of a spectral imager. The information capacity captures many of the essential performance characteristics in a single figure of merit that is not directly tied to a particular application. Possibly, information capacity could be developed into a standardized performance metric for spectral imagers.

The analytical treatment in the example above considers the very basic case of optimizing the pixel size while keeping the imaging optics fixed. This can be employed as an element in the design of imaging optics by repeating the pixel size optimization in each iteration of the design, using the information capacity as a figure of merit for optimization.

The results illustrate that coregistration error can lead to a strong reduction in the information capacity. This degradation is most significant under conditions where the ratio of signal to photon noise is high.

References

1. P. B. Fellgett and E. H. Linfoot, "On the assessment of optical images," Phil. Trans. R. Soc. A **247**, 369–407 (1955).
2. I. J. Cox and C. J. R. Sheppard, "Information capacity and resolution in an optical system," J. Opt. Soc. Am. A **3**, 1152–1158 (1986).
3. C. L. Fales, F. O. Huck, and R. W. Samms, "Imaging system design for improved information capacity," Appl. Opt. **23**, 872–888 (1984).
4. R. Piestun and D. A. B. Miller, "Electromagnetic degrees of freedom of an optical system," J. Opt. Soc. Am. A **17**, 892–902 (2000).
5. C.-L. Tisse, F. Guichard, and F. Cao, "Does resolution really increase image quality?" Proc. SPIE **6817**, 68170Q (2008).

6. F. Cao, F. Guichard, and H. Hornung, "Information capacity: a measure of potential image quality of a digital camera," Proc. SPIE **7537**, 75370F (2010).
7. P. Mouroulis, D. A. Thomas, T. G. Chrien, V. Duval, R. O. Green, J. J. Simmonds, and A. H. Vaughan, Trade Studies in Multi/Hyperspectral Imaging Systems—Final Report (NASA Jet Propulsion Laboratory, 1998).
8. T. Skauli, "An upper-bound metric for characterizing spectral and spatial coregistration errors in spectral imaging," Opt. Express **20**, 918–933 (2012).
9. T. Skauli, "Quantifying coregistration errors in spectral imaging," Proc. SPIE **8158**, 81580A (2011).
10. G. J. Burton and I. R. Moorhead, "Color and spatial structure in natural scenes," Appl. Opt. **26**, 157–170 (1987).
11. A. Martinez, "Spectral efficiency of optical direct detection," J. Opt. Soc. Am. B **24**, 739–749 (2007).
12. A. Lapidoth and S. M. Moser, "On the capacity of the discrete-time Poisson channel," IEEE Trans. Inf. Theory **55**, 303–322 (2009).
13. A. Lapidoth, "Capacity results of an optical intensity channel with input-dependent channel noise," IEEE Trans. Inf. Theory **58**, 207–223 (2012).
14. S. Arimoto, "An algorithm for computing the capacity of arbitrary discrete memoryless channels," IEEE Trans. Inf. Theory **18**, 14–20 (1972).