Measurement of point spread function for characterization of coregistration and resolution: comparison of two commercial hyperspectral cameras

Hans Erling Torkildsen, Torbjørn Skauli
Measurement of point spread function for characterization of coregistration and resolution - comparison of two commercial hyperspectral cameras

Hans Erling Torkildsen, Torbjørn Skauli*
Norwegian Defence Research Establishment (FFI), P O Box 25, 2027 Kjeller, Norway

ABSTRACT

For spectral cameras, the spatial resolution and coregistration of bands are important performance characteristics. The paper discusses how these can be quantified in a way that is relevant for the spectroscopic processing step which is at the beginning of most spectral image analysis methods. It is argued that the most appropriate measure of resolution may be the ensquared energy of the mean point spread function (PSF) within the specified pixel field of view. For spatial coregistration, it is argued that a previously proposed coregistration metric is well suited, by characterizing the full shape and position differences between PSFs of different bands. A relatively simple method for imaging the PSF is demonstrated, based on tomographic reconstruction from line spread functions measured in different directions. The method is used to characterize two presumably comparable commercial hyperspectral cameras, a SpecIm PFD dating from 2012 and a HySpex VNIR-1600 dating from 2016. The measurements reveal significant differences, but it is pointed out that the present measurements by themselves do not constitute a proper comparison of the camera types. The main point made in the paper is that full characterization of the PSF in all bands is possible with moderate experimental effort, and provides clarifying measurements of actual camera performance.

Keywords: Hyperspectral imaging, Multispectral imaging, Remote sensing, Optical design, Coregistration, Spectroscopy, Modulation transfer function, Standardization

1. INTRODUCTION

The quality of spectra collected by a hyperspectral camera depends critically on the spatial coregistration between bands. When imaging an inhomogeneous scene, coregistration errors will lead to a crosstalk from spatial contrasts to spectral contrasts. It is well known that even small coregistration errors can lead to large errors in the measured pixel spectrum [1-3]. For a scene containing some distribution of black and white areas, an ideal camera should yield spectrally grey pixels at scene boundaries. In this paper, it is shown that a real camera can deviate significantly from this ideal behavior. We also discuss how image sharpness can be characterized in a way relevant to hyperspectral image analysis.

The spatial coregistration performance of commercial hyperspectral cameras is often specified in terms of the "keystone" distortion, essentially the pixel centroid shift resulting from residual wavelength-dependence of magnification. However, this measure of pixel shift does not account for wavelength dependencies in the width or shape of the PSF. Thus there is a very real possibility for hyperspectral cameras with the same coregistration specification to exhibit large differences in their actual coregistration performance.

There is at present no widely accepted standard for full characterization of coregistration errors. Several works have considered coregistration beyond the simple keystone measure. A recent example is [4], where a scanned line source is used to measure the line spread function (LSF) in the "along-track" and "across-track" directions of an imaging spectrometer. A Gaussian fit was used to create a map of the LSF position and width. In [5], a set of metrics was proposed for measuring sharpness and coregistration in the across-track direction of an imaging spectrograph. In [6], a coregistration error metric is proposed and shown to be a generalization of keystone error, while also capturing differences in width and shape of the point spread function (PSF). In [7], we showed that a tomographic imaging scheme is feasible for forming an image of the PSF. Further review and discussion of PSF measurement is given in [8].

*torbjorn.skauli@ffi.no
Figure 1. Illustration of the opposing concerns for image sharpness and coregistration. The squares illustrate the pixel grid for two spectral bands, red and blue, overlaid on a complex remote sensing scene. The figure assumes that the optics introduces a coregistration error in the form of a constant offset. Left: Larger pixels, with lower spatial resolution, but better coregistration. Right: Smaller pixels, with better spatial resolution, but large coregistration error relative to the pixel size.

2. METRICS FOR SPATIAL COREGISTRATION AND SHARPNESS

2.1 Coregistration

Coregistration error can arise if there are differences between bands in PSF position or PSF shape. In a given pixel, all bands should have identical PSFs. If the pixel is viewing a homogeneous scene, then the measured spectrum will be the same regardless of differences in PSF position or shape between bands. If the scene is inhomogeneous within the pixel, then the ideal spectrum is a weighted sum of the scene constituents, with the same weightings for each band. For this to hold regardless of the shape of the scene, the PSFs of all bands must be identical. Breaking it down to the simplest case, we can consider spectral imaging with just two bands. Any difference between the two PSFs has potential to change the weighting of scene components from one band to another. Figure 1 illustrates simple cases where the pixel grid of two bands has a relative offset, leading to different weighting of scene constituents. Such weighting errors will lead to pixel spectra which are not linear combinations of the spectra of the endmember components that are present in the pixel. In this way, coregistration error violates the assumption of linear mixing inherent in many image processing methods. Also, regardless of the processing method, coregistration errors will tend to increase the spectral dimensionality of the image data. This is particularly true for scenes containing spatial contrast on many scales, such as the important case of a sunlit outdoor scene, where shadows range in size from mountains to grains of sand.

It is important to consider coregistration and spatial resolution together. Observe that coregistration can be improved simply by defocusing the camera: Assuming that the optics behaves like a conventional camera, which is true for many spectral imaging architectures, defocusing will tend to make the PSFs approach a disk shape with the same diameter in all bands (since the plane of the defocused image sensor array intersects the focused light cone away from the focus). With increasing defocus, the growing diameter of the disk increases overlap between bands, and any residual differences between PSFs will tend to become negligible. In this way, it is in principle possible to achieve arbitrarily good coregistration. Another way to improve coregistration is to replace each pixel by an average over a neighborhood of pixels ("binning"), as illustrated in Figure 1. In both cases, improved coregistration comes at the expense of reduced image sharpness, of course. Therefore, design of a spectral camera will to some extent be a tradeoff between resolution and coregistration[9], but not in the sense that these characteristics are inverses of each other. A good spectral camera will be characterized by a combination of good image sharpness and low coregistration error.

In [6], it was proposed to characterize coregistration error between two bands as the volume between the normalized PSFs. For two bands with band indices $n$ and $m$, the spatial coregistration error metric $\varepsilon_{s,mn}$ is an integral of PSF difference over the image $(x,y)$ plane

$$
\varepsilon_{s,nm} = \frac{1}{2} \iint_{x,y} |f_m(x,y) - f_n(x,y)| \, dx \, dy
$$

(1)
where \( f_n(x,y) \) is the PSF for band \( n \), normalized to unity integral. In [6], it is shown that the metric value represents an upper bound on the weighting error of spectral components in a mixed pixel containing different endmember constituents. This metric characterizes coregistration independently of image sharpness, and will properly account for the improved coregistration that results from defocus or binning. An overall metric for coregistration performance of a camera can be formed from the mean or maximum value of \( \varepsilon_{s,\text{nm}} \) over all band pairs.

A somewhat different coregistration metric was proposed in [5], for characterization of coregistration errors in the across-track direction of an imaging spectrometer, a case of significant practical interest. This alternative metric is calculated from differences in pixel enslaved energy, which is easily measured by scanning a one-dimensional point source across the field of view (FOV) in the across-track direction. To be applicable in a wider range of cases, this metric should ideally be generalized to two spatial dimensions, in order to capture the full range of coregistration errors. Figure 2 illustrates cases where one-dimensional characterization of coregistration will fail to capture significant errors. Also, a coregistration metric of the kind proposed in [5], calculated from enslaved or ensquared energy, would not capture differences between PSFs within the nominal pixel. Such differences could lead to signal errors for a mixed pixel (containing different scene constituents), and should be captured by the coregistration metric. We therefore proceed using the metric (1), which accounts for the full shape of the PSF.

That said, [5] makes an important observation about the metric (1), namely that it will not correctly represent the largest possible error when imaging a point source. In [6], the metric (1) is shown to represent an upper bound on errors when the scene is composed of different extended sources, characterized by their radiance or excitance. Imaging of point sources, which are characterized by their intensity (in W/sr), is an important separate case, which can be relevant in applications such as astronomy or single molecule imaging. In such cases, it is quite clear that the largest coregistration error occurs when the point source is located at the point where the difference between PSFs is largest. The relevant metric for maximum error is this largest PSF difference, for point sources. The metric (1) is then essentially an expectation value for the error in the point source case, averaged over all point source positions.

![Figure 2. Illustration of the need for two-dimensional characterization of the PSF shape to determine coregistration. The contours illustrate PSFs of two bands in the same pixel, with a large coregistration error. Left: PSFs which appear to be equal when judged by a vertical LSF. Right: PSFs which have nearly equal LSFs in both horizontal and vertical directions. A characterization of coregistration from a single LSF may fail to detect the coregistration error in these cases.](image)

2.2 Sharpness

In conventional imaging, sharpness is essential, for preservation of spatial contrasts. Conventional characteristics such as MTF are then appropriate for the application. In spectral imaging, on the other hand, most forms of image exploitation start by processing the spectra of individual pixels. Sharpness is then better assessed by the degree of preservation of the pixel spectrum. Thus ideally, a pixel in a spectral image should contain an average spectrum over the nominal pixel footprint, with minimal influence from adjacent areas. When coregistration is characterized independently by the metric (1), sharpness can be measured based on the mean PSF over all bands. Ideally, this mean PSF should coincide with the boundaries of the nominal pixel FOV. A relevant measure of sharpness is then the ensquared energy of the mean PSF within a nominal pixel. This is essentially a generalization to two dimensions of the approach for characterizing sharpness proposed in [5].
For this measure of sharpness, it is necessary to define the nominal pixel area. It is not uncommon among commercial cameras to have a specified pixel instantaneous field of view (IFOV) larger than the pixel sampling interval. Therefore, a fair measure of sharpness is the enclosed energy within the specified pixel IFOV. Here we will report this measure of sharpness, as well as the ensquared energy within a pixel size given by the pixel sampling interval.

In the following, dimensions in the image plane are given in pixel units. For the case of line scanning cameras considered here, the pixel unit is taken to be the pixel sampling interval in the across-track direction. The same distance is then taken to be the pixel size in the along-track direction. This is reasonable here, given that the cameras under test employ detector arrays with square pixels.

A fair measure of sharpness is to evaluate the enclosed energy within a pixel IFOV of the specified size, positioned relative to the mean PSF so that the enclosed energy is maximized. Here we simplify the evaluation of enclosed energy by positioning the nominal IFOV at the centroid of the measured mean PSF. Since the measured PSFs generally tend to be symmetric about the across- and along-track directions for the cameras under test here, this simplified evaluation should produce values very close to the optimum. For cases where the IFOV exhibits strong asymmetry, it may be necessary to numerically optimize the position of the IFOV region over which the enclosed energy is evaluated.

It can be noted that in the case of imaging of point sources, it appears that a better measure of sharpness will be the average ensquared energy [10] over all possible source positions. We have not analyzed this situation in detail, since an extended source is by far the most common case for hyperspectral imaging.

3. MEASUREMENT

3.1 Experimental setup

In this work, we employ an unconventional technique for measuring the full shape of the point spread function (PSF) of each spectral band in a given pixel. The technique is based on scanning a line source across the pixel in different directions, and using tomographic reconstruction to form an image of the PSF. The technique was demonstrated in [7] and is based on earlier work on tomographic scanning imaging [11].

The line source is based on an optical slit with a width corresponding to a small fraction of the pixel width. We use a 22 $\mu$m $\times$ 25 mm slit formed in a metal film on glass. The slit is illuminated from the back with a broadband incoherent source. We use a 5 cm integrating sphere illuminated with 2$\times$20 W halogen lamps, supplemented in the blue with LEDs at 405 and 420 nm. The slit is mounted on a translation stage, which in turn is mounted on a rotation stage and arranged such that the slit scans across a 25mm optical port in the center of rotation. The slit needs to move with subpixel precision, particularly with respect to keeping the center of rotation constant. We have found that sufficient precision is obtained using regular-grade motorized stages, combined with normal experimental care to ensure stable mounting, avoid strain from cables etc. The light path has been enclosed in a tube to eliminate residual blur due to air turbulence. For the work presented here, the cameras are viewing the slit directly, focused (except where noted) at a distance of about 1 m. The focus was adjusted by optimizing the pixel ensquared energy, as discussed below. (In an alternative setup, the slit could be used as the light source in a collimator, to test cameras focused at infinity.)
The line source is scanned while reading out data from the camera. The scan speed is adapted so that there are about 30 scan steps across a pixel. Depending on the scan length, magnification, and pixel size of the camera, the scanned region is typically several pixels across. Here we report measurements recorded near the center of the field of view of the cameras under test. To characterize coregistration variation across the full FOV, it is necessary to repeat the measurement after rotating the camera so that the scanning slit is seen in different parts of the FOV.

Figure 4. Example of a measured PSF, recorded from the SpecIm camera at 699 nm wavelength. (a) Contour map of the PSF with contours at 50, 75 and 90% enclosed energy. (b) and (c) Measured line spread function and cross section in the along (b) and across (c) track directions. Also shown are line spread functions calculated back from the reconstructed PSF (blue), as well as cross sections through the peak (black). (d) 3D plot of the PSF.

3.2 PSF reconstruction and evaluation

The resulting time series recorded during a scan from a pixel in a particular band is the line spread function (LSF) for that band. The LSF is a projection of the PSF in a direction orthogonal to the scan direction. By making linear scans in different directions around the circle, a set of PSF projections is obtained, from which the two-dimensional PSF for each band is recovered by an inverse Radon transform. With \( N \) scan steps across a pixel, and \( M \) different scan directions, the transform produces an image of the PSF resolved in \( \approx N \times M \) pixels. Outside the main PSF peak, the reconstruction tends to produce a radial pattern of artifacts with the same angular periodicity as the scan pattern. To reduce these artifacts, we have employed spectral filtering in the tangential direction, after transformation to cylindrical coordinates. Further discussion of the tomographic imaging technique is given in papers by Hovland [11].

Figure 4 shows an example of a measured PSF. The directly measured LSF in the along (b) and across (c) track directions is shown together with an LSF calculated from the reconstructed PSF, showing that the reconstructed PSF reproduces the measured data after the relatively complex reconstruction process. Also shown is a cross section through the peak of the PSF in each direction, from which we measure the peak full width at half maximum (FWHM).

From the set of PSFs for all bands, and for a small neighborhood of pixels, we can straightforwardly calculate a variety of characteristics, and even simulate measured spectra for arbitrary scenes. The sharpness is evaluated from the mean PSF averaged over all bands, and even simulate measured spectra for arbitrary scenes. The sharpness is evaluated from the mean PSF averaged over all bands, by integrating over the specified pixel FOV, as discussed above. For calculation of coregistration error, consider that the 3D image (d) of the reconstructed PSF shows that there is some residual noise and artifacts in the region of near-zero response outside the main PSF peak. The noise and artifacts will tend to average out to zero. However, if the metric integral (1) includes these areas, the coregistration error will be overestimated, due to the absolute value operator. To minimize the contribution from noise in the estimation of coregistration error, a threshold is set for the PSF in each band so that 95% of the PSF energy is above the threshold. The metric integral (1) is then evaluated over the area where both PSFs are above their respective thresholds.
Table 1. Summary of results for coregistration and sharpness.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Coregistration error, mean</td>
<td>0.22</td>
<td>0.20</td>
<td>0.09</td>
</tr>
<tr>
<td>Coregistration error, 90 percentile</td>
<td>0.35</td>
<td>0.36</td>
<td>0.16</td>
</tr>
<tr>
<td>Coregistration error, max</td>
<td>0.78</td>
<td>0.60</td>
<td>0.27</td>
</tr>
<tr>
<td>Mean PSF FWHM across track (pixels)</td>
<td>2.96</td>
<td>1.11</td>
<td>2.99</td>
</tr>
<tr>
<td>Mean PSF FWHM along track (pixels)</td>
<td>4.27</td>
<td>1.64</td>
<td>4.05</td>
</tr>
<tr>
<td>Peak ensquared energy in 1×1 pixels (%)</td>
<td>7.1</td>
<td>33.6</td>
<td>80.8</td>
</tr>
<tr>
<td>Peak ensquared energy in specified pixel IFOV (%)</td>
<td>23.8</td>
<td>52.8</td>
<td>80.8</td>
</tr>
</tbody>
</table>

3.3 Cameras under test

Here we show measurement results for two hyperspectral transmission-grating spectrograph cameras. These cameras are normally used in a line-scanning configuration. Here, we keep the cameras stationary, so that there is no PSF broadening due to the scan motion. Both cameras cover the spectral range from 400 to 1000 nm, approximately.

The first camera is a SpecIm PFD camera with 1312 pixels across the field of view, dating from 2012. The camera has an OLE23 23 mm objective lens with adjustable focus. The slit width is 30 µm, imaged 1:1 through the spectrograph onto
detector pixels with 8 µm pitch. Thus, the nominal pixel size is 8x8 µm in the slit plane. This camera was set to record images with no pixel binning. The specification then implies an IFOV of 1 pixel across and 30/8=3.75 pixels along track.

The second camera is a HySpex VNIR-1800 camera with 1800 pixels, dating from 2016. This camera is used with a 1 m fixed focus lens accessory. This camera operates by default with 2x binning in the along-track direction, and is specified as having an IFOV of 1x2 pixels. For this camera, focus was adjusted by translating the camera relative to the slit.

4. RESULTS

The main results are summarized in Figure 5 and Table 1. The top part of Figure 1 shows a contour map of the spatial coregistration error $\varepsilon_{x,m,n}$ for all band pairs. Each point in this symmetric matrix represents the difference between the corresponding pair of PSFs. The bottom part of the figure shows selected pairs of PSFs corresponding to different points in the matrix. As apparent from the figure, the main difference between the two cameras is that the HySpex camera exhibits significantly better image sharpness. The SpecIm FWHM is about 2.6 times larger in both the along- and across-track directions. (It can be noted that several different focus settings were tested for the SpecIm camera, and the reported results are for the best overall setting, see discussion of focus below.) The coregistration matrix turns out to look qualitatively similar for the two cameras. This similarity reflects the fact that for both cameras, the PSF evolves with wavelength, and is sharpest in the middle part of the spectral range. In terms of the coregistration metric, the HySpex camera exhibits only about 10% better coregistration, on average. Thus, when viewing a scene with strong spatial contrast on all scales, such as a sunlit outdoor scene, the fraction of misregistration-induced spectral artifacts will be comparable for the two cameras. Of course, this most favorable comparison does not take into account the differences in spatial resolution. For a more fair comparison, consider a case where the HySpex image is binned down to a spatial resolution comparable to the SpecIm camera. We have approximated this situation by forming a new PSF in each band from the sum of several copies of the HySpex PSF, shifted by an integer number of pixels. The numbers in the last column of Table 1 are obtained for a 3×3 binning of image pixels (which corresponds to 3×6 detector pixels, since the HySpex camera by default uses 2× along-track binning). Then the binned PSF has an area comparable to the unbinned SpecIm PSF, so that the resolution of the resulting images would be comparable. The coregistration matrix for this binning pattern is shown in the upper half of the matrix on the right in Figure 5. The binning leads to a significant reduction in coregistration errors, particularly for the wavelength pairs that exhibit a large PSF difference in the unbinned case. The mean coregistration error is then reduced to 0.09.

![Figure 6](https://www.spiedigitallibrary.org/conference-proceedings-of-spie) variation of coregistration error with focus for the HySpex camera. This graph shows the relative variation of the mean pixel signal for different focus settings when the light source is centered on a pixel and oriented in the along- or across-track directions. The pixel reading is used as a measure of the degree of focus. The "total" curve is the sum for the two slit orientations. The corresponding measurements of coregistration error illustrate the opposing relation between coregistration and resolution.
Figure 7. Reflectance spectra which would be measured for a scene consisting of a transition from black to white, calculated from the measured PSF. The spectra are calculated for different slit positions, and for two orientations of the edge. Numbers indicate the offset of the edge from the pixel center, for the indicated groups of curves. Insets show the edge at position -2. Black: SpecIm camera. Blue: HySpex camera, unbinned. Red: HySpex camera binned 3x3.

Figure 6 shows a set of data illustrating the focusing method, and the effect of defocus. For the comparison of cameras in Figure 5 and Table 1, the cameras were focused by varying the focus setting and observing the peak value of the mean LSF. Figure 6 shows examples of how the LSF peak height varies and reaches a maximum at a certain focus setting. It turns out that the position of this maximum depends on the orientation of the slit. Slightly different focus settings are found for the along- and across-track orientations for both cameras. An overall focus measure can be obtained by averaging the values for the two slit orientations, as shown. The data in Figure 5 and Table 1 are recorded at the peak of this total energy curve for both cameras. As a check, we have also measured the PSF at the peak of the two other curves. We find that such small changes in focus do not affect the value of the coregistration metric significantly.

Figure 6 also shows the mean coregistration error for a subset of focus settings where a full PSF measurement has been performed. The results illustrate clearly that the sharpest focus tends to be associated with the largest coregistration error. By simply defocusing the camera, it is possible to trade off resolution against coregistration error, within constraints set by the basic imaging quality of the instrument.

The coregistration metric value does not carry any information about the shape of the spectral artifacts that would be expected. To illustrate the effect of coregistration on image data, we have used the measured PSFs to calculate the reflectance spectra that would be measured for a scene consisting of a transition from black to white, shown in Figure 7.
The curves show spectra that would be observed for different positions and orientations of the edge. The curves show, above all, that the measurements would deviate significantly from the expected shape of a flat, gray spectrum near the edge. Both cameras exhibit deviations by tens of percentage points in reflectance. The main difference between the cameras lies in the faster transition from white to black for the HySpex camera, which for this scene would lead to a smaller fraction of incorrect pixel spectra in the image. It can be observed that the degree of spectral artifact is less when the edge is centered on the pixel, due to the approximate mirror symmetry of the PSF in the two scan directions used here. For other edge orientations, the symmetry will be less helpful.

Figure 7 also includes predicted spectra for the HySpex camera after 3×3 binning. The binning gives a clear reduction of the error in the spectra, which are still quite different from the ideal gray, flat shape.

5. DISCUSSION AND CONCLUSIONS

This work has demonstrated the feasibility of detailed measurement of PSF shape, using a relatively simple experimental setup. The measured data provide full information about spatial coregistration and resolution. A variety of performance metrics can be extracted from these data. The results in Figure 5 illustrate that the tomographic reconstruction technique can be used to extract rich information about the detailed characteristics of hyperspectral cameras.

Note that although the results in the Figure 7 are simulated from our measurement of PSF shape, the same spectra could be measured directly, by translating a knife edge in front of a uniform light source, or by a single slit scan. Such measurements could be made with a simpler experimental setup, and could form the basis of a simplified coregistration testing scheme, but many forms of coregistration error would then be missed, as illustrated in Figure 2. A possible alternative to the technique used here would be a direct PSF measurement by 2D scanning of a point source based on broadband laser sources which have become available in recent years.

The cameras tested here are comparable in the sense that their main characteristics are similar, and they target the same applications and market segments. The considerable differences between the two commercial cameras, which are not reflected in their specifications, illustrate that there is a need for better standards and practices in characterization of spectral cameras[12]. The purpose of this work is not primarily to make a comparative review of the cameras, however. A full comparison would involve several other characteristics in addition to coregistration and resolution. Also, we have not taken any steps to ensure that the tested cameras are representative units. Furthermore, we point out that the camera which is found to exhibit better sharpness is also several years newer than the other.

We have argued that the metric (1) and the ensquared energy in the specified IFOV are good metrics for characterizing coregistration and image sharpness, respectively. These metrics are independent of each other, and they are not limited to measuring particular forms of imperfections (such as for keystone). These preferred metrics reflect the requirements of most hyperspectral image analysis methods, which start by processing the spectral dimension.

ACKNOWLEDGMENTS

The authors are grateful to Ingunn Burud and Espen Olsen of the Norwegian University of Life Sciences for lending us the SpecIm camera for testing, and to Harald Hovland and Thomas Opsahl for advice on processing PSF measurements. This work is based on results presented at SPIE DCS 2018 in paper 10644-51.

DISCLOSURES

FFI has ongoing collaborations with NEO, makers of the HySpex camera. The work presented here, however, has been carried out independently by FFI. The HySpex camera has been purchased by FFI on commercial terms following a public tender.
REFERENCES