

LA-UR- 97-2605

Approved for public release;
distribution is unlimited.

CONF-970706--

Title:

ALGORITHMS USING INTER-BAND
CROSS-CORRELATION FOR PIXEL REGISTRATION
AND JITTER RECONSTRUCTION IN
MULTI-CHANNEL PUSH-BROOM IMAGERS

Author(s):

James P. Theiler, NIS-2
Bradley G. Henderson, NIS-2
Barham W. Smith, NIS-2

Submitted to:

SPIE Annual Meeting, July 28-Aug. 1, San
Diego, CA

Proc. SPIE 3163, "Signal and Data
Processing of Small Targets 1997", O.E.
Drummond, editor

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED *ph*

MASTER

Los Alamos
NATIONAL LABORATORY

Los Alamos National Laboratory, an affirmative action/equal opportunity employer, is operated by the University of California for the U.S. Department of Energy under contract W-7405-ENG-36. By acceptance of this article, the publisher recognizes that the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or to allow others to do so, for U.S. Government purposes. Los Alamos National Laboratory requests that the publisher identify this article as work performed under the auspices of the U.S. Department of Energy. The Los Alamos National Laboratory strongly supports academic freedom and a researcher's right to publish; as an institution, however, the Laboratory does not endorse the viewpoint of a publication or guarantee its technical correctness.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, nor any of their employees, make any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

Algorithms Using Inter-band Cross-correlation for Pixel Registration and Jitter Reconstruction in Multi-channel Push-broom Imagers

James Theiler, Bradley G. Henderson, and Barham W. Smith

Astrophysics and Radiation Measurements Group, NIS-2
Nonproliferation and International Security Division, MS-D436
Los Alamos National Laboratory, Los Alamos, NM 87545 USA

RECEIVED
NOV 03 1997
OSTI

ABSTRACT

We present two algorithms for determining sensor motion of a multi-spectral push-broom imager for use in subsequent image registration. The first algorithm, termed the "pairwise" method, performs cross-correlations between individual pairs of channels. The offsets of maximum correlation are formulated into a system of linear equations whose solution gives an estimate of the jitter function. The second algorithm performs cross-correlations between channels and a reference image called the "baseline" which is constructed by averaging together all the channels in the image cube. An estimated jitter time series is computed for each channel, all of which are overlapped and averaged to obtain a best estimate of the jitter function. The pairwise method is more general in that it can handle a wider range of jitter scenarios. The baseline method, although more restricted, is very simple to implement, and its accuracy can be improved substantially through iteration. In this paper, we describe both methods in detail and present results of simulations performed on thermal-infrared data cubes.

Keywords: registration, jitter, multispectral, push-broom imager, algorithm

1. INTRODUCTION

Registration is the process by which multiple images of the same scene are "lined up" so that corresponding pixels represent the same geographic location on the ground. It is a fundamental process in the analysis of remotely sensed imagery and is essential for many applications.¹ For change detection studies, images of the same scene acquired at different times must be registered so that changes in the scene can be recognized by an analyst or a computer algorithm. In other applications, images of the same scene acquired by different sensors are co-registered before applying "fusion" algorithms for enhancing the spatial resolution of one data set or for improving the total information content relative to the individual images themselves.^{2,3}

Another important application of registration lies in the exploitation of multi-spectral data. Multi-spectral images must be co-registered before retrieval algorithms (for vegetation,⁴ columnar water vapor,⁵⁻⁷ surface temperature,^{8,9} *etc.*) can be applied to individual pixels. Although many multi- and hyper-spectral sensors acquire images already registered, those with physically separated detectors, such as push-broom systems, do not.

In the focal plane of a push-broom imager is a linear array (or "frame") of pixels. The array is pushed across a scene of interest, and the pixels in the frame are periodically read out into computer memory from which an image of the scene is reconstructed. In effect, an image is built up one row at a time as the array is pushed across the scene. The sampling interval in the cross-track direction is determined by the pixel spacing on the focal plane, and the sampling interval in the along-track direction is determined by the ground-track speed and the pixel readout rate. The reconstruction of the scene requires that the motion of the imager be known very accurately; in particular, any jitter in the scanning motion will degrade the image. If this jitter can be measured (for instance, with accurate on-board gyroscopes or accelerometers), then it can be accounted for in the image reconstruction.

For multichannel push-broom imagers, the linear arrays are physically separated on the focal plane so that each channel "sees" a particular part of the scene at a different time. Although this adds difficulty to the problem of

Work supported by the U.S. Department of Energy. Email: {jt,bghenderson,bwsmith}@lanl.gov.

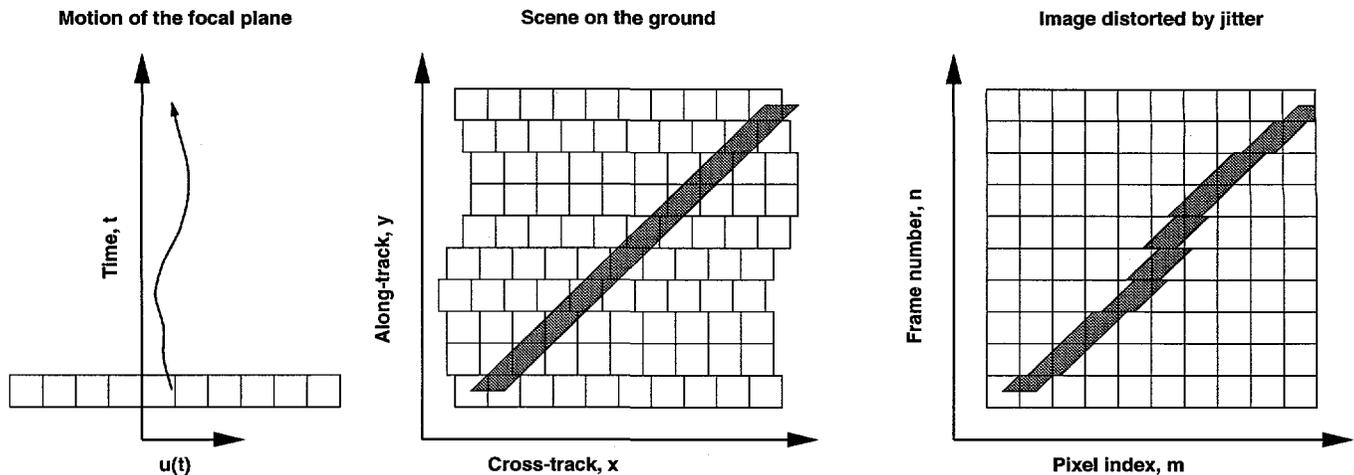


Figure 1. The effect of cross-track jitter on a push-broom imager is illustrated here. A scene on the ground (represented as a diagonal line) is imaged in pixels which are slightly offset from one row to the next. The lower left pixel, for instance, is from frame $n = 0$, pixel number $m = 0$, and is identified with a position on the ground given by $x = x_o + u(t_{io})$, and $y = y_{io}$.

registering the different channels to each other, it actually provides information that can be exploited in reconstructing the unknown jitter in the scanning motion, which can then be used for more accurate pixel registration and image reconstruction.

In this paper, we present two methods for determining sensor motion for use in registration of multi-spectral push-broom images. The underlying notion is that if the pointing direction of the sensor is known during imaging, it can be used to co-register the individual channels. Both algorithms use image correlation techniques. The first method performs cross-correlations between pairs of individual channels, the maxima of which are formulated into a system of linear equations which is solved to give an estimate of the jitter function. The second method performs cross-correlations between individual channels and a reference "baseline" which is an average of the multi-spectral images. It should be mentioned that the algorithms presented here are not intended to be used for general image matching. They are designed specifically for registration of multi-spectral images acquired on one pass of a push-broom system.

The next two sections will describe the "pairwise" and the "baseline" approaches, and these will be followed by a numerical simulation of a jittered push-broom imager along with a characterization of the accuracy with which these algorithms are able to reconstruct the simulated jitter.

2. PAIRWISE INTER-BAND CROSS-CORRELATION

The first algorithm to be presented, which we will refer to as the "pairwise" method, uses cross-correlations between pairs of individual channels to infer sensor motion. The method lends itself most readily to computing the cross-track jitter, the mechanics of which will be described first. The cross-track jitter estimation is done in two parts. The first part uses the images to obtain cross-correlation maxima; these maxima correspond to the difference in the cross-track jitter at two different times. The second step combines these individual estimates of jitter differences into a set of linear equations whose solution provides a single estimate of the cross-track jitter function. We will also briefly discuss prospects for inferring the along-track jitter which is also based on pairwise cross-correlations.

2.1. Notation

We will begin with a fairly detailed description of the notation. These details are of course necessary for an actual implementation of the pairwise inter-band cross-correlation algorithm, but they will also be useful for clarifying the similarities between the pairwise and the baseline methods.

- Let $r_i(m, n)$ denote the data value (radiance, typically) in the m 'th pixel on the n 'th frame of the i 'th channel.

- Let $x_i(m, n)$ and $y_i(m, n)$ denote the cross-track and along-track position on the ground scene that that pixel is pointed at, and
- Let $t_i(m, n)$ be the time that that pixel is exposed.
- Let $u(t)$ and $v(t)$ be the cross-track and along-track jitter (respectively) as a function of time t .

The scanner movement is dominated by uniform motion in the along-track direction, but is modified by jitter in both the along-track and cross-track direction. We can make this explicit with expressions for x , y , and t :

$$t_i(m, n) = t_{i0} + \gamma n \quad (1)$$

$$\begin{aligned} y_i(m, n) &= y_{i0} + \beta n + v(t_i(m, n)) \\ &= y_{i0} + \beta n + v(t_{i0} + \gamma n) \end{aligned} \quad (2)$$

$$\begin{aligned} x_i(m, n) &= x_o + \alpha m + u(t_i(m, n)) \\ &= x_o + \alpha m + u(t_{i0} + \gamma n) \end{aligned} \quad (3)$$

In particular, note that t and y depend on n but not m , and that x depends on n only through the jitter function $u(t)$. Here t_{i0} is the "turn-on" time for the i 'th channel, and y_{i0} is the along-track ground position at which the i 'th channel's frame is pointing at turn-on time. We will assume that t_{i0} and y_{i0} are known quantities (note that they don't need to be known absolutely, but only relative to the other channels). For simplicity, and without loss of generality, we will choose our units so that $\alpha = \beta = \gamma = 1$.

Note that m and n are integers, but we can still define interpolated data values $r_i(m', n')$ for m' , n' not integers, using a suitable interpolation scheme. Piotrowski *et al.*¹⁰ have noted that for their registration efforts, interpolation did not play a major role.

- Let $n_i(y)$ denote the frame number at which the i 'th channel is pointed at position y on the ground. Note that $n_i(y) = y - y_{i0}$ when the along-track jitter is negligible, but that in the more general case, $n_i(y)$ has to be defined with the nonlinear implicit formula with

$$y = y_{i0} + n_i(y) + v(t_{i0} + n_i(y)). \quad (4)$$

Keep in mind that $n_i(y)$ is not necessarily an integer.

- Let $R_i(m, y) = r_i(m, n_i(y))$ be the radiance at the m 'th pixel in the i 'th channel for the frame n in which the i 'th channel is pointing at position y in the along-track direction. When $n_i(y)$ is not an integer, then r_i can be computed by interpolation.
- Let $Y_i = t_{i0} - y_{i0}$. The difference $|Y_i - Y_j|$ is a fixed characteristic of the imager; it corresponds to the distance on the ground between where channels i and j are pointing at any given time. Since the absolute offsets for t_{i0} and y_{i0} are arbitrary, it is convenient to choose them so that $Y_i \geq 0$ for all i , and $\min_i Y_i = 0$.

Note that the i 'th channel $R_i(m, y)$ will be pointing at a position in the cross-track direction given by $x = x_o + m + u(t_{i0} + y - y_{i0}) = x_o + m + u(y - Y_i)$, while the j 'th channel $R_j(m', y)$ will be pointing at $x = x_o + m' + u(y - Y_j)$; They will both be pointing at the same location as long as

$$m - m' = u(y + Y_j) - u(y + Y_i). \quad (5)$$

Our goal is to estimate this difference $u(y + Y_j) - u(y + Y_i)$ by looking at the value of $m - m'$ that maximizes the cross-correlation between $R_i(m, y)$ and $R_j(m', y)$. Our working assumption (verified in simulation) is that the maximum cross-correlation between channels occurs when both channels are pointed to the same stripe of ground.

- Let $C_{ij}(z, y) = \langle R_i(m + z, y) R_j(m, y) \rangle_m$ be the cross-correlation as a function of this shift z .

Note that although m is properly speaking an integer, we can define $R_i(m', y)$ for non-integer m' by interpolation, and thereby obtain $C_{ij}(z, y)$ for non-integer z . Finally,

- Let $z_{ij}(y)$ be the value of z that maximizes $|C_{ij}(z, y)|$. We take the absolute value so that we can find the point of maximum *anti*-correlation for channel-pairs that are anti-correlated.

Since we expect this maximum to occur when R_i and R_j are pointing to the same part of the scene, we expect

$$z_{ij}(y) \approx u(y + Y_j) - u(y + Y_i). \quad (6)$$

If we know that the function $u(t)$ has some intrinsic smoothness, then we can often improve the reconstruction by smoothing these values of $z_{ij}(y)$. The utility of this step will obviously depend on how smooth the actual jitter function is. For the results that we will show later, no smoothing was done.

2.2. Counting equations and unknowns

If there are N frames per channel, and B channels, then there will be at most $B(B - 1)N/2$ distinct equations of the form in Eq. (6). This maximum is achieved if the the channels are each turned on at different times in just such a way that they all observe the same N "frames" of the ground scene. In general, however, there will be somewhat fewer than $B(B - 1)N/2$ equations.

We can be more precise about the number of equations, by noting that $z_{ij}(y)$ is computable only for values of y in which both $R_i(m, y)$ and $R_j(m, y)$ have been measured. If we

- Let N_i denote the number of frames of data taken in the i 'th channel, and
- Let $H_i(y)$ indicate values of y for which $R_i(m, y)$ has been measured; that is $H_i(y) = 1$ for $y_{i0} \leq y \leq y_{i0} + N_i - 1$, and $H_i(y) = 0$ otherwise; then we can
- Let $H_{ij}(y) = H_i(y)H_j(y)$ be an indicator function which is 1 if $z_{ij}(y)$ has been computed, and zero otherwise.

In terms of the indicator function, we can write the total number of equations as $\sum_{ijy} H_{ij}(y)$.

The unknowns are the values of $u(t)$ that are required for values of t ranging from the minimum to the maximum of $y + Y_i$. At N frames per channel, there will be at least N unknowns. If turn-on times are adjusted so that all channels are flown over the same part of the ground, then the number of unknowns is $N + \max_i Y_i - \min_i Y_i$.

As long as N is substantially larger than $|Y_i - Y_j|$, and as long as there are three or more channels, then there will be many more equations than unknowns. With only two channels, however, the problem is necessarily under-determined. Note that even with $B \geq 3$, the problem is not completely solvable, because adding a constant C to any solution $u(t)$ of the set of equations in Eq. (6) gives another valid solution. The interchannel cross-correlation algorithm does not resolve this ambiguity, but the ambiguity does not affect the quality of channel-to-channel registrations.

2.3. Least-squares formulation

One way to "solve" for the jitter is to express the equations in Eq. (6) in terms of a least-squares problem. Again using the indicator function $H_{ij}(y)$, we can write

$$\chi^2 = \sum_{ijy} w_{ij} H_{ij}(y) [z_{ij}(y) - u(y + Y_j) - u(y + Y_i)]^2, \quad (7)$$

where w_{ij} is a weight assigned to the ij pair of channels. Channel pairs that are more reliably cross-correlated will get higher weights.

To find the jitter function $u(t)$ which minimizes χ^2 , we differentiate it with respect to $u(t)$ and set the derivative to zero. That is,

$$0 = \frac{\partial \chi^2}{\partial u(t)} = 2 \sum_{ij} w_{ij} \{ -H_{ij}(t - Y_j) [z_{ij}(t - Y_j) - u(t) + u(t - Y_j + Y_i)] + H_{ij}(t - Y_i) [z_{ij}(t - Y_i) + u(t) - u(t - Y_i + Y_j)] \} \quad (8)$$

In particular, if we define a vector \mathbf{b} with components

$$b(y) = \sum_{ij} w_{ij} [H_{ij}(y - Y_j)z_{ij}(y - Y_j) - H_{ij}(y - Y_i)z_{ij}(y - Y_i)] \quad (9)$$

$$= \sum_{ij} w_{ij} [H_{ij}(y - Y_j)z_{ij}(y - Y_j) + H_{ji}(y - Y_i)z_{ji}(y - Y_i)] \quad (10)$$

where we made the equation more symmetrical by using $z_{ij} = -z_{ji}$; and a matrix \mathbf{A} with elements

$$A(y, t) = \sum_{ij} w_{ij} [H_{ij}(y - Y_j)\delta(y, t) + H_{ij}(y - Y_i)\delta(y, t) - H_{ij}(y - Y_i)\delta(y - Y_i + Y_j, t) - H_{ij}(y - Y_j)\delta(y - Y_j + Y_i, t)], \quad (11)$$

then we can write $\sum_t A(y, t)u(t) = b(y)$; or, as a matrix equation,

$$\mathbf{A}\mathbf{u} = \mathbf{b} \quad (12)$$

whose solution \mathbf{u} provides the jitter function $u(t)$. Note that the matrix \mathbf{A} is banded, sparse, and symmetric. Because of the endpoints, it is only approximately Toeplitz. The matrix is also nonnegative definite, but note that it does have a zero determinant. In particular, there cannot be a unique solution to Eq. (12), because $\mathbf{u} + C$ is also a solution for any constant C .

2.4. Iterative solution of the matrix equation

The least squares design matrix \mathbf{A} in Eq. (12) is quite large, and solving the equation is generally an $O(N^3)$ process where N is the size of the matrix (which in this case is the number of individual times t that we want to estimate $u(t)$, and is typically in the hundreds).

However, \mathbf{A} is dominated by the elements on its diagonal, and so one can find an approximate inverse fairly rapidly, and use that for an iterative solution to Eq. (12). In fact, the diagonal elements of the matrix \mathbf{A} are all nonzero, so our first approximation for an inverse is the diagonal matrix whose elements are the reciprocals of the elements of \mathbf{A} . That is, $\mathbf{B}^{(0)}$ has elements only along the diagonal: $b_{ii}^{(0)} = 1/a_{ii}$. One can successively improve this estimate of the inverse using

$$\mathbf{B}^{(2k+1)} = 2\mathbf{B}^{(k)} - \mathbf{B}^{(k)}\mathbf{A}\mathbf{B}^{(k)} \quad (13)$$

This is generally expensive because it involves two $O(N^3)$ matrix multiplications. However, since $\mathbf{B}^{(0)}$ is a diagonal matrix, the first iteration can be performed with only $O(N^2)$ operations. We use this first iteration to compute $\mathbf{B}^{(1)}$ and then solve Eq. (12) by repeated application of

$$\mathbf{u}^{(k+1)} = \mathbf{u}^{(k)} + \mathbf{B}^{(1)}(\mathbf{b} - \mathbf{A}\mathbf{u}^{(k)}), \quad (14)$$

each iteration of which requires $O(N^2)$ operations. See the discussion in Ref. 11 for more details.

2.5. Along-track jitter

The analysis presented so far assumes that there is no along track jitter; that is, $v(t) = 0$. If the along-track jitter function $v(t)$ is known, then it can be used in the equation for $n_i(y)$, which is used to compute $R_i(m, y)$. This leads to a more accurate estimate of the cross-track jitter.

In practice, of course, the along-track and cross-track jitter are both unknown. Pairwise cross-correlation can still be used to infer these jitter functions from the multispectral image data. The basic idea is to consider a cross-correlation function of two variables:

$$C_{ij}''(w, z, y) = \langle R_i(m + z, y + w)R_j(m, y) \rangle_m \quad (15)$$

and to find the functions $w_{ij}(y)$ and $z_{ij}(y)$ that *simultaneously* maximize the cross-correlation. Separate matrix equations can be used to estimate the cross-track and along-track functions: $u(t)$ and $v(t)$. In our implementation, we start with $v(t) = 0$ and we infer $u(t)$, using the cross-correlation approach discussed above. Then we take the

estimated $u(t)$ as "known," and estimate $v(t)$. We then use this $v(t)$ to go back and improve the estimate of $u(t)$, and so on (though, in practice, we see little improvement beyond the third step).

Unfortunately, the cross-correlation in the along-track direction is less accurate and more complicated. The inaccuracy stems from the fact that push-broom images are acquired one (cross-track) row at a time. In a given row, pixels are strictly contiguous and simultaneously acquired; neither of these conditions hold for the pixels in an (along-track) column. The bookkeeping is also more difficult in the along-track case, because we cannot use the along-track distance y as a uniformly varying index. Let

$$n_i^*(y) \equiv y - y_{io} \quad (16)$$

be an approximation to the frame number $n_i(y)$ defined in Eq. (4) that is correct when there is no along-track jitter. The idea is to compute a new cross-correlation function

$$C'_{ij}(y, w) = \langle R_i(m - u(y + w - Y_i), n_i^*(y) + w) R_j(m - u(y - Y_j), n_j^*(y)) \rangle_m \quad (17)$$

and note that it is expected to have a maximum at $w = w_{ij}(y)$ where

$$n_i^*(y) + w_{ij}(y) - n_j^*(y) = n_i(y) - n_j(y) \quad (18)$$

That is,

$$w_{ij}(y) = (n_i(y) - n_i^*(y)) - (n_j(y) - n_j^*(y)). \quad (19)$$

In the small amplitude limit (actually, what we require is that $dv/dt \ll 1$), we can approximate Eq. (4) with

$$v(y + Y_i) \approx n_i(y) - y + y_{io} = n_i(y) - n_i^*(y) \quad (20)$$

and therefore

$$w_{ij}(y) \approx v(y + Y_i) - v(y + Y_j) \quad (21)$$

provides the approximate equalities, which like those of Eq. (6) for the cross-track motion, can be combined with a large matrix equation into an expression for $v(t)$ in terms of the various functions $w_{ij}(t)$. In fact, the matrix \mathbf{A} used for the cross-track estimation is identical to the matrix used in the along-track case.

3. BASELINE CORRELATION

Instead of performing all possible pairwise cross-correlations, a simpler approach is to average the images obtained individually in each of the channels into a "baseline" image, and then to cross-correlate this baseline with each of the individual channels. The motivation here is that the baseline provides an approximate image of the ground that is less distorted by jitter than any of the individual bands. Each of these cross-correlations provides a separate but direct estimate of the jitter function (there are no matrix equations to be solved); the individual estimates can then be averaged to provide a single overall estimate of the jitter function. An earlier implementation of the baseline method has been described elsewhere⁹; we will describe our current version here, using the notation already developed for the pairwise inter-band cross-correlation method.

3.1. Method

First, a "baseline" image is generated by averaging all the images in all the bands of interest. In particular, the following function is defined:

$$R_*(m, y) = \langle R_j(m, y) \rangle_j = \sum_j w_j R_j(m, y) \quad (22)$$

where w_j is the weight on the j 'th channel. In practice, unequal weights are employed here as scaling factors so that the R 's have the same range (or the same variance). A cross-correlation is performed between the baseline and each of the individual channels. That is,

$$C_i(z, y) = \langle R_*(m + z, y) R_i(m, y) \rangle_m \quad (23)$$

and the maximum of this cross-correlation defines a function

$$z_i(y) = \operatorname{argmax}_z C_i(z, y). \quad (24)$$

Since we are assuming that the baseline image is an approximation to an unjittered image, we take $z_i(y) \approx u(y + Y_i)$ as approximation to the jitter function from the i 'th channel. Since a separate estimate of $u(t)$ is made for each channel, they can be averaged to obtain an overall estimate of the jitter history:

$$\hat{u}(t) = \langle z_i(t - Y_i) \rangle_i \quad (25)$$

Note that a separate set of weights could be used to make this a weighted average. In practice, these weights are generally equal for all channels.

Note that the cross-correlation in Eq. (23) is between channel i , and a baseline (computed in Eq. (22)) which includes channel i . We have also considered a version of the baseline method in which a separate baseline is computed for each channel by averaging all the other channels. That is

$$C'_i(z, y) = \langle [R_*(m + z, y) - w_j R_i(m, y)] R_i(m, y) \rangle_m. \quad (26)$$

but we found that this approach was less stable.

Now we can compare this result to the pairwise inter-band cross-correlation approach, by first writing Eq. (23) in terms of the pairwise cross-correlation functions:

$$C_i(z, y) = \langle \langle R_j(m + z, y) \rangle_j R_i(m, y) \rangle_m \quad (27)$$

$$= \langle R_j(m + z, y) R_i(m, y) \rangle_{m,j} \quad (28)$$

$$= \langle C_{ji}(z, y) \rangle_j. \quad (29)$$

Since $C_i(z, y)$ is the average of $C_{ji}(z, y)$, it is a fair approximation to write $z_i(y) = \langle z_{ji}(y) \rangle_j$. In this case, we can use Eq. (6) to write

$$z_i(y) = \langle u(y + Y_i) - u(y + Y_j) \rangle_j = u(y + Y_i) - \langle u(y + Y_j) \rangle_j. \quad (30)$$

Finally, averaging the z_i 's, we get

$$\hat{u}(t) = \langle z_i(t - Y_i) \rangle_i \quad (31)$$

$$= u(t) - \langle u(t - Y_i + Y_j) \rangle_{i,j}. \quad (32)$$

In particular, this provides an error estimate for the baseline method of $\langle u(t - Y_i + Y_j) \rangle_{i,j}$. We can see that for many bands and/or for jitter functions $u(t)$ with short decorrelation times, this error term will be small. On the other hand, the error scales linearly with the amplitude of the jitter amplitude, and if there is significant autocorrelation over timescales of order $|Y_i - Y_j|$, then the error will be more substantial.

The baseline correlation method depends on the assumption that the distortion present in the individual images is removed by the averaging process used to create the baseline. Thus it would seem that a large number of bands is necessary in order to obtain enough statistics to create a reasonable baseline. However the method can succeed with a moderate number of bands due to the double averaging inherent in the method. First, each pixel in the baseline is computed by averaging pixels from m different bands. Second, each point in the final time series is computed by averaging together m points from the individual band time series. Because of the offset of the linear arrays on the focal plane, a given point in time in the m individual time series represents m different points on the baseline. Thus, a single point in the final time series comes from m different points in the individual time series, each of which comes from m different points on the baseline.

3.2. Iterated Baseline Correlation

The accuracy of the baseline correlation method can be improved significantly through iteration. The steps involved in this process are mapped out in Fig. 2. The procedure starts with the image cube to which the baseline method is applied to obtain an initial estimate of the jitter motion. Next, the computed jitter time series is used to co-register all the bands. Because the initial jitter reconstruction is not exact, there remains some mis-registration in

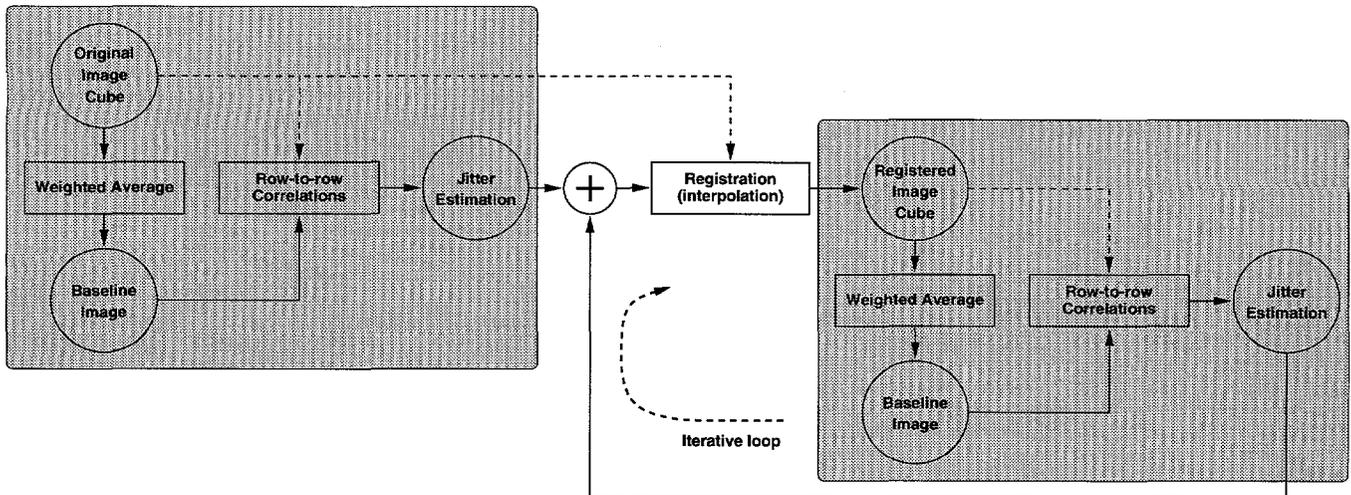


Figure 2. Flow chart showing the steps involved in the iterated baseline correlation method. The module in gray is the single baseline correlation step. Note that registration is always performed on the original image cube.

this intermediate (partially registered) image cube. The baseline method is now applied again to the intermediate image cube to obtain another jitter time series. This time series, because it was performed on the partially registered images, is actually a perturbation which is added to the original jitter estimate to obtain an improved estimate of the actual time series. To minimize the information loss via interpolation, the improved jitter estimate is applied to the original image cube, not the intermediate registered cube, during the registration step. This process can be iterated a number of times to improve the desired accuracy.

4. SIMULATION AND RESULTS

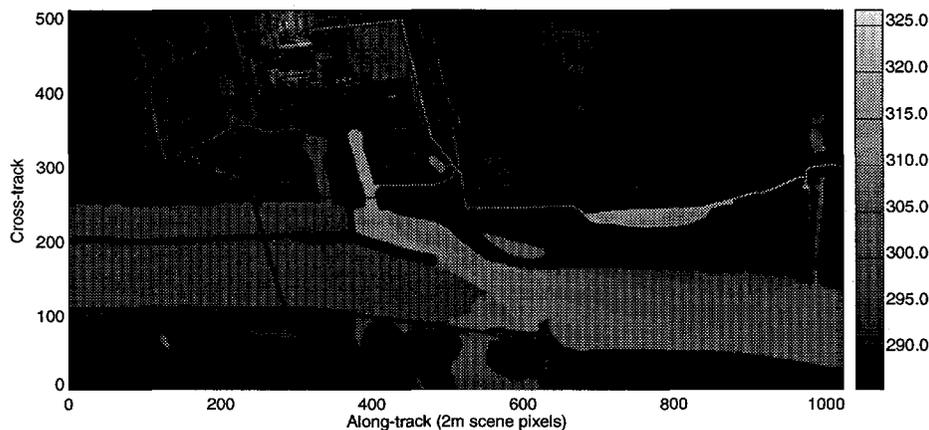


Figure 3. Input scene used for the simulation – a night-time temperature (in degrees Kelvin) map of the Bull Run coal-fired plant on the Clinch River near Oak Ridge, Tennessee, acquired by band 11 ($8.5\text{--}14\ \mu\text{m}$) of the Daedalus-1268 scanner. The cross-track and along-track dimensions are 512×1024 in two-meter scene pixels. We placed four of these temperature scenes end-to-end, to achieve a 512×4096 scene. The radiance images were then 51×409 in twenty-meter image pixels.

To test our registration algorithms, we used thermal imagery to simulate image acquisition as might be done for multi-spectral land surface temperature retrieval. As an input scene, we started with a calibrated brightness temperature

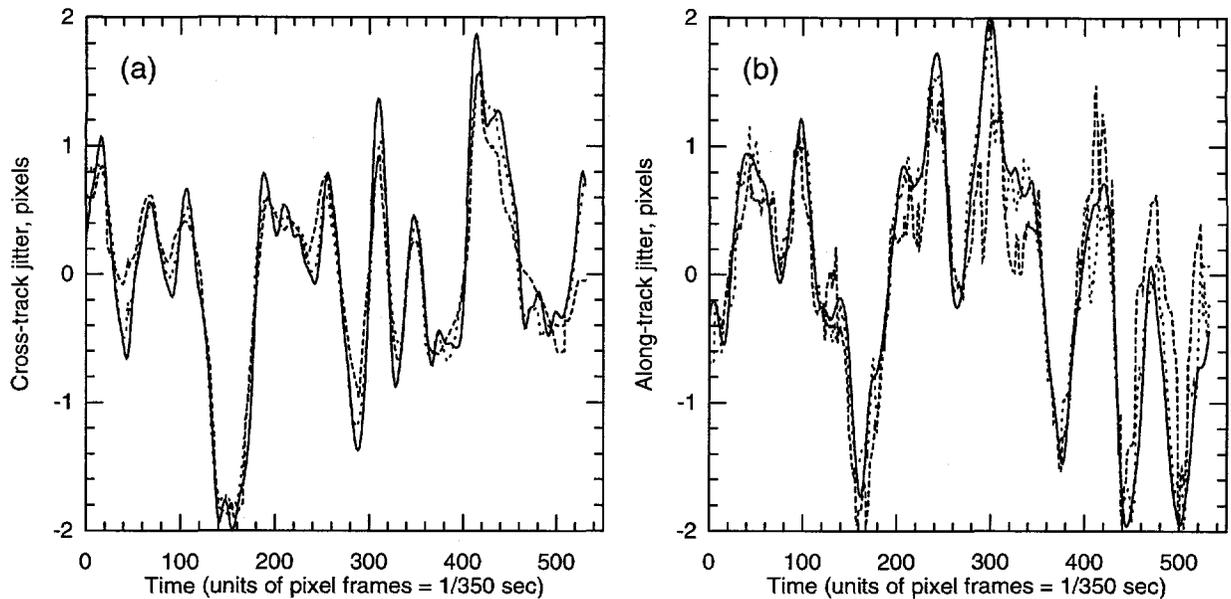


Figure 4. This figure shows (a) cross-track, and (b) along-track jitter reconstruction from the simulated five-channel multispectral image described in the text. The solid line indicates the “actual” simulated jitter motion; the dotted line is the pairwise reconstructed jitter, and the dashed line is the baseline reconstructed jitter. The pairwise method gave root-mean-square estimation errors of 0.162 pixels and 0.226 pixels, respectively, for the cross-track and along-track reconstructions. The baseline method achieved values of 0.276 pixels and 0.475 pixels for these two jitter reconstruction errors.

image of the Bull Run coal-fired plant on the Clinch River near Oak Ridge, Tennessee, acquired by band 11 (8.5–14 μm) of the Daedalus-1268 scanner (Fig. 3). We placed four of these scenes end-to-end to achieve a longer along-track dimension. The brightness temperature map was converted to top of atmosphere (TOA) radiances for 5 bands in the mid- and long-wave IR by assuming unit surface emissivity and using an algorithm based on MODTRAN.¹² Those radiances were then blurred by the system MTF, which, for this particular simulation, included the effects of diffraction-limited optics, motion of a push-broom imager, and the pixel size. Next, hypothetical jitter time series were computed to describe the sensor motion in the cross- and along-track directions during imaging. The jitter motion was taken to be a random realization of a specified Fourier power spectrum; in this case, we took the spectrum to be flat from zero to a cutoff frequency $f_o = 10\text{Hz}$, and to scale as $f^{-\alpha}$, with $\alpha = 8$ for higher frequencies. So the cutoff at f_o is quite sharp. The linear arrays were then “flown” over the input scene at a ground-track speed of 7 km/s to acquire images in each band a row at a time, shifting in x and y as dictated by the cross- and along-track jitter time series, respectively. The scene and image pixels were 2 m and 20 m respectively, resulting in 10×10 scene pixels per sensor pixel during imaging. The frames for the five bands were at relative positions 0.0, 22.54, 45.08, 100.08, and 123.08 in units of the 20m image pixels. (These correspond to the variables Y_i in the earlier exposition). The results presented here did not include noise in the calculations. However, previous trials showed that reasonable amounts of noise have a negligible effect on the jitter reconstruction accuracy. Yaw motion was not included in the simulation.

A comparison of the pairwise and baseline methods of jitter estimation is shown in Fig. 4. For a single iteration, the pairwise method achieves a considerably smaller error than the baseline method. However, as seen in Fig. 5, this difference is seen to evaporate when the iterative scheme is applied to the baseline method; the iterated baseline and the pairwise were found to attain similar accuracies. In all of these comparisons, we did not permit any smoothing of the final or intermediate results. If the jitter function has some smoothness (as our example did), then smoothing can substantially improve the final estimation. In this example, by adding knowledge of the jitter function power spectrum, we were able to see errors of 0.055 and 0.150 pixels, respectively, for the cross-track and along-track jitter reconstruction errors, using the iterated baseline method.

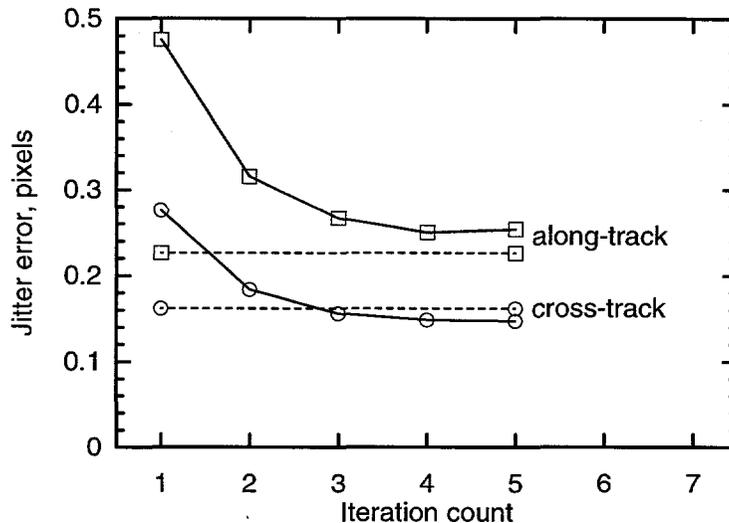


Figure 5. Iteration of the baseline method permits improved jitter estimation. In this figure, cross-track estimates are indicated with circles, and along-track estimates with squares. The dashed lines indicate the errors obtained from the uniterated pairwise method. This experiment suggests that the pairwise and the iterated baseline methods are roughly comparable.

5. DISCUSSION AND CONCLUSION

We have presented two methods for reconstructing jitter motion of a multi-spectral push-broom imager. The computed jitter time series is intended to be used for registration of the multi-spectral image cube. The first method, the pairwise method, performs cross-correlations between individual pairs of channels, whereas the baseline method performs cross-correlations between channels and a reference baseline computed by averaging together all bands in the image cube. The pairwise method is more general in that it is subject to fewer assumptions and can handle a wider range of jitter types in terms of amplitude and frequency content of the time series. However, it is more cumbersome to use and does not as readily lend itself to computing the along-track jitter, or to incorporating into an iterative scheme. The baseline method is more limited in that the characteristic de-correlation time of the jitter must be short relative to the spacing of the channels on the focal plane. The baseline method, however, is simple to implement and can be improved substantially through iteration.

Success of these jitter reconstruction methods relies on adequate correlation between the spectral bands in the image cubes. In our simulation, we used night-time thermal infrared imagery, which are well correlated spectrally. We have not tested these algorithms in other spectral regions. Presumably the performance will not be as good when correlating bands of widely different spectral regions. Within a given spectral region, solid materials on the earth's surface typically demonstrate a fair amount of correlation so that the methods presented here should still be applicable. Furthermore, one can always choose the weighting factors to favor those spectral regions which exhibit the most correlation.

ACKNOWLEDGEMENTS

We are grateful to our colleagues Christoph Borel, William Clodius, Pawel Smolarkiewicz, and Paul Weber, for useful discussions and comments during the course of this work.

REFERENCES

1. L. M. G. Fonesca and B. S. Manjunath, "Registration techniques for multisensor remotely sensed imagery," *Photogrammetric Engineering and Remote Sensing* **62**, pp. 1049-1056, 1996.
2. T. W. Carper, T. M. Lillesand, and R. W. Kiefer, "The use of intensity-hue-saturation transformation for merging panchromatic and multispectral image data," *Photogrammetric Engineering and Remote Sensing* **56**, pp. 459-467, 1990.

3. C. K. Munechika, J. S. Warnick, C. Salvaggio, and J. R. Schott, "Resolution enhancement of multispectral image data to improve classification accuracy," *Photogrammetric Engineering and Remote Sensing* **59**(1), pp. 67-72, 1993.
4. C. J. Tucker, "Red and photographic infrared linear combinations to monitor vegetation," *Rem. Sens. Environ.* **8**, pp. 127-150, 1979.
5. R. O. Green, V. Carrère, and J. E. Conel, "Measurement of atmospheric water vapor using the airborne visible/infrared imaging spectrometer," in *Am. Soc. Photogram. and Remote Sensing, Workshop on Image Processing, Sparkes, Nevada, 23-26 May, 1989*.
6. C. Borel and D. Schläpfer, "Atmospheric pre-corrected differential absorption techniques to retrieve columnar water vapor: theory and simulations," in *Proceedings of the 6th JPL Airborne Earth Science Workshop (AVIRIS '96)*, 1996.
7. C. Borel, W. B. Clodius, and J. Johnson, "Water vapor retrieval over many surface types," in *Algorithms for Multispectral and Hyperspectral Imagery II*, A. Iverson, ed., vol. 2758 of *Proc. SPIE*, pp. 218-228, 1996.
8. C. Tornow, C. C. Borel, and B. J. Powers, "Robust water temperature retrieval using multi-spectral and multi-angular ir measurements," *Proc. IGARRS '94*, pp. 441-443, 1994.
9. B. G. Henderson, C. C. Borel, J. Theiler, and B. W. Smith, "Image quality degradation and retrieval errors introduced by registration and interpolation of multispectral digital images," in *Signal and Data Processing of Small Targets 1996*, O. E. Drummond, ed., *Proc. SPIE* **2759**, pp. 149-160, 1996.
10. C. C. Piotrowski, M. S. Farber, S. J. Hemple, and W. S. Helliwell, "An analysis of advantages and disadvantages of scene registration as a part of space-time processing," in *Signal and Data Processing of Small Targets 1996*, O. E. Drummond, ed., *Proc. SPIE* **2759**, pp. 98-110, 1996.
11. W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C*, Cambridge University Press, Cambridge, p. 57, 1992.
12. A. Berk, L. S. Bernstein, and D. C. Robertson, "MODTRAN: A moderate resolution model for lowtran 7," Tech. Rep. GL-TR-89-0122, Geophysics Laboratory, Air Force Systems Command, Hanscom AFB, MA, 1989.
13. C. L. Chan *et al.*, "Performance of an adaptive algorithm for suppressing intraframe jitter in a 3-d signal processor," in *Signal and Data Processing of Small Targets 1996*, O. E. Drummond, ed., *Proc. SPIE* **2759**, pp. 42-61, 1996.
14. N. Nabaa and R. H. Bishop, "A recursive solution to the sensor registration problem in a multiple sensor tracking scenario," in *Signal and Data Processing of Small Targets 1996*, O. E. Drummond, ed., *Proc. SPIE* **2759**, pp. 385-396, 1996.