Journal of Applied Statistics
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713428038

Alignment and Sub-pixel Interpolation of Images using Fourier Methods

C. A. Glasbey[a]; G. W. A. M. Van Der Heijden[b]
[a] Biomathematics and Statistics Scotland, Edinburgh, UK [b] Biometris, Wageningen, The Netherlands

To cite this Article Glasbey, C. A. and Van Der Heijden, G. W. A. M.(2007) 'Alignment and Sub-pixel Interpolation of Images using Fourier Methods', Journal of Applied Statistics, 34: 2, 217 — 230
To link to this Article: DOI: 10.1080/02664760600995155
URL: http://dx.doi.org/10.1080/02664760600995155

Journal of Applied Statistics
Vol. 34, No. 2, 217–230, March 2007

Alignment and Sub-pixel Interpolation of Images using Fourier Methods

C. A. GLASBEY* & G. W. A. M. VAN DER HEIJDEN**

**Biomathematics and Statistics Scotland, Edinburgh, UK, ***Biometris, Wageningen, The Netherlands*

ABSTRACT *A method is proposed for both estimating and correcting a translational mis-alignment between digital images, taking account of aliasing of high-frequency information. A parametric model is proposed for the power- and cross-spectra of the multivariate stochastic process that is assumed to have generated a continuous-space version of the images. Parameters, including those that specify misalignment, are estimated by numerical maximum likelihood. The effectiveness of the interpolant is confirmed by simulation and illustrated using multi-band Landsat images.*

KEY WORDS: Aliasing, coherency, complex Gaussian distribution, cross-spectrum, landsat image, phase spectrum, power spectrum, sub-pixel

Introduction

A digital image consists of a set of pixels, which are typically sampled points on a rectangular spatial lattice. The sampling involves a spatial convolution or smoothing of a process in continuous space, plus the addition of noise. In image analysis, we often need values of the process at non-lattice points, mainly for image enlargement (zooming). These values are usually obtained by interpolation, which implicitly assumes that it is sufficient to recover the smoothed version of the process and that noise is negligible. Standard interpolation methods include the bilinear, bicubic and b-spline interpolators. The sinc interpolator is optimal if sampling is in accord with the Nyquist criterion, i.e. the sampling frequency is two times the highest frequency in the smoothed process. However, if the image contains aliasing, the sinc interpolator is not optimal. A general theory for optimal linear interpolation is provided by kriging (Parrott *et al.*, 1993). For a recent review of image interpolation, see Meijering (2002).

As well as for enlargement, another reason for interpolation is to align images that have sub-pixel translation shifts. For example, microscope images obtained using different imaging modalities can be translationally shifted (Glasbey & Mardia, 2001), as can the different bands in remotely sensed images (Berman *et al.*, 1994). To illustrate the second case, consider bands 1–5 and 7 of a Landsat TM image, shown in Figure 1 and previously analysed by Glasbey & Horgan (1995). Figure 2 shows phase differences between Fourier transforms of all pairs of Landsat bands, in which translations between bands manifest themselves as

Correspondence Address: C. A. Glasbey, Biomathematics and Statistics Scotland, James Clerk Maxwell Building, King's Buildings, Edinburgh, EH9 3JZ, Scotland, UK. Email: chris@bioss.a.c.uk

0266-4763 Print/1360-0532 Online/07/020217–230 © 2007 Taylor & Francis
DOI: 10.1080/02664760600995155

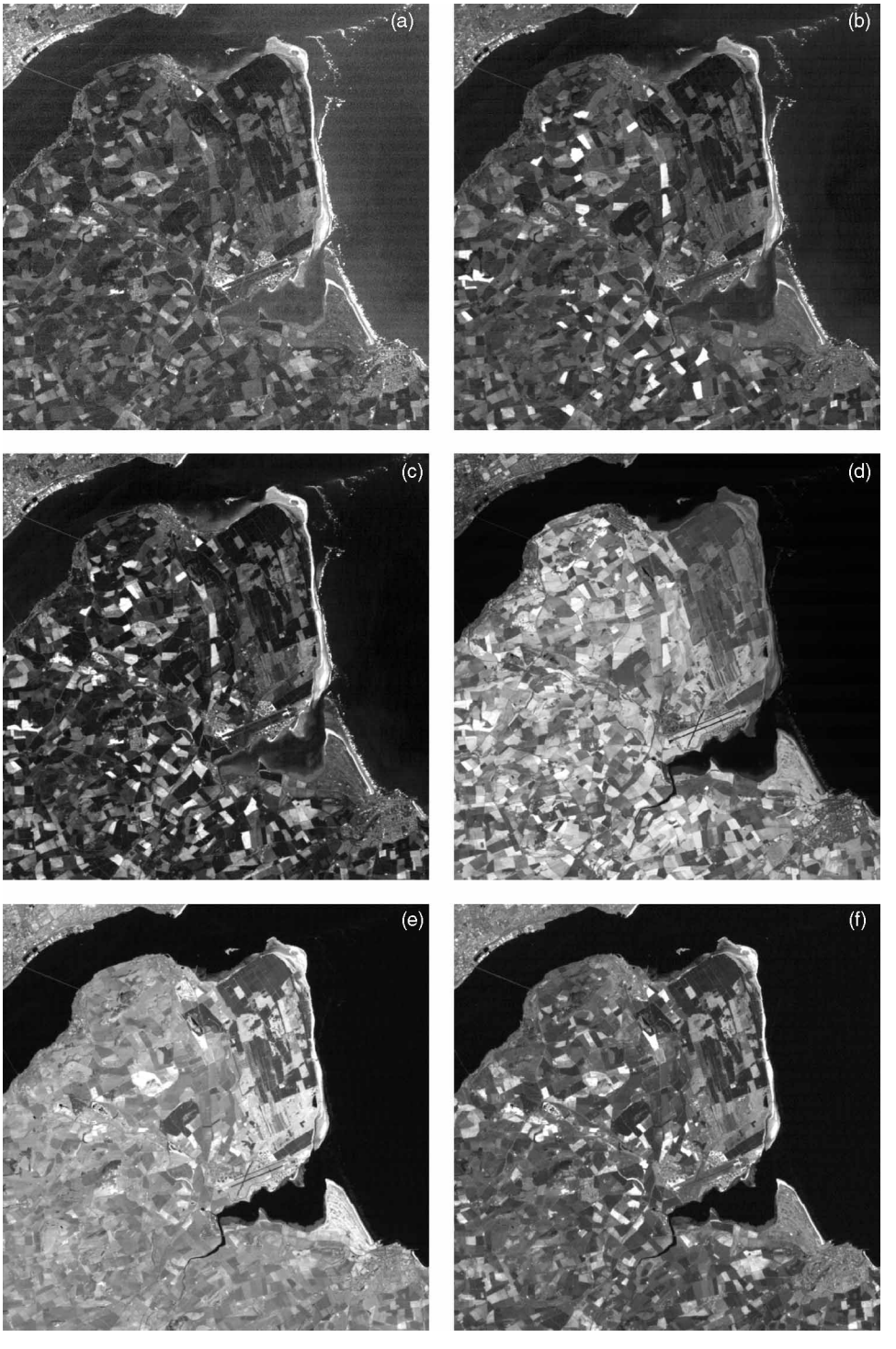

Figure 1. Landsat Thematic Mapper (TM) images: bands 1–5 and 7, for a region between the river Tay and the town of St Andrews on the east coast of Scotland, in May 1987. There are 512×512 pixels, each 30^2 m. (© National Remote Sensing Centre Ltd, Farnborough, Hampshire)

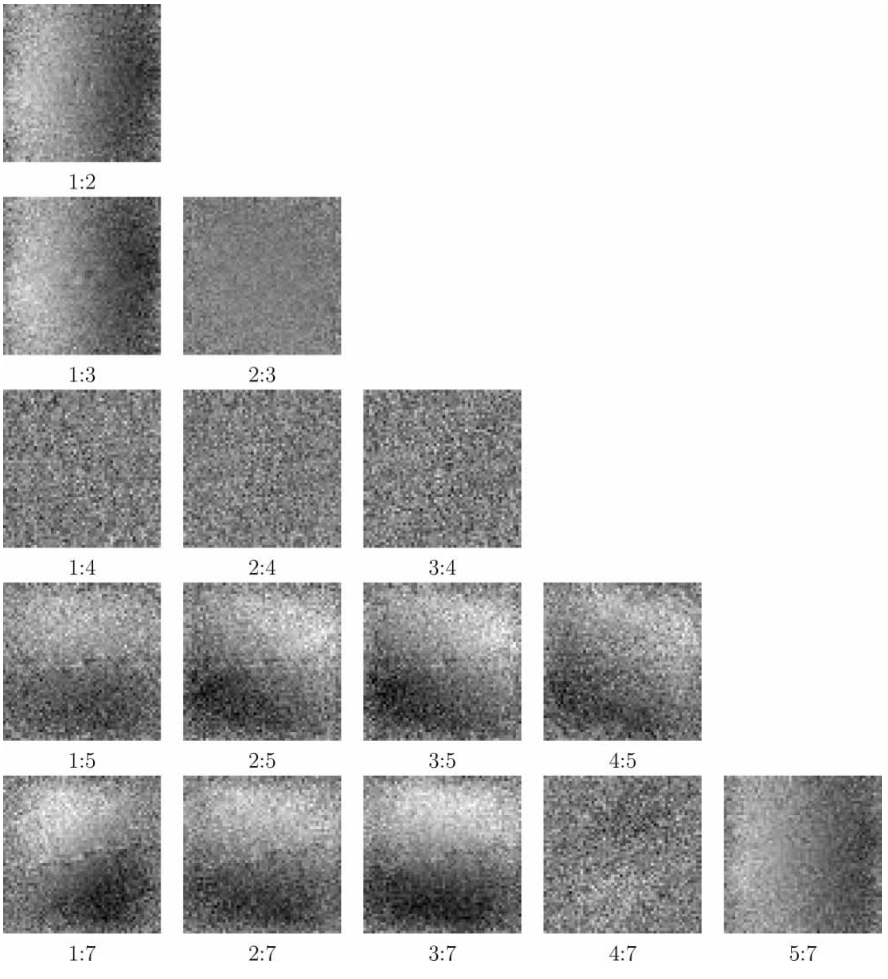

Figure 2. Phase differences between Fourier transforms of pairs of Landsat bands in Figure 1, averaged over 11×11 blocks of frequencies, with angles displayed on a monotonically increasing grey scale over the range $[-1, 1]$, as black for angles < -1 and as white for angles >1

linear trends (see the next section for further details). We see evidence for bands 2 and 3 being translated horizontally relative to band 1, band 5 being translated vertically and band 7 being translated diagonally. We note that these trends cover the full range of frequencies except for tapering to zero at the highest frequencies. This means that noise is negligible in these bands, and the tapering effect is due to aliasing. Band 7 also appears to be translated vertically relative to bands 2 and 3 and translated horizontally relative to band 5. In such cases, there is a need both to estimate the translation, and to correct for it by interpolation.

Estimation of a translation is most elegantly done in the Fourier domain. For example, using Fourier transforms the cross-correlation between images at all translations that are whole numbers of pixels can be computed simultaneously. Berman *et al.* (1994) took account of aliasing by empirically modelling the effect on phase differences. Kaltenbacher & Hardie (1996) and Luengo Hendriks & van Vliet (2000) estimated sub-pixel translations by approximating one image by a first-order Taylor series expansion of a second image.

After estimating the translation, interpolation is possible to sub-pixel accuracy, by making use of cross-correlation between images. This is often referred to as super-resolution (Hunt, 1995). One special case is where the shifted images are sampled from a common process, for example where movement of a camera provides sub-pixel shifted frames. Several methods have been proposed (Borman & Stevenson, 1998; Luengo Hendriks & van Vliet, 2000), which can be used to increase the resolution of camera images. Another special case is digital colour imaging, where a colour mosaic filter is used to obtain red, green and blue bands: the misalignment between bands is known, but there is still a need to correct for it by interpolation, and an opportunity to use the information from the other colour bands. The most common colour mosaic filter is the Bayer filter, where every second pixel in a chequerboard pattern is green and the others alternate between red and blue pixels. A common approach to interpolate these colour images is demosaicking, developed by Freeman (1988). Bands are interpolated using a standard interpolator, such as bilinear interpolation, and then a median filter is applied to the difference image of, for example, red-green and blue-green. A mathematical framework is proposed by Trussel & Hartwig (2002), and Ramanath *et al.* (2002) describe this and several other approaches.

In this paper we propose an approach that simultaneously estimates the amount of aliasing, the misalignment between the images, and the coherency between them. The estimated parameters are used to create a powerful interpolant and also to align the images. The method is formulated in the next section, and applied to the Landsat data in the third section. Finally, in the fourth section, the methodology is discussed.

Method

Let $F_{j,xy}$ denote the pixel value at spatial location (x, y) in the jth of J digital images, for integer values of $x = 0, \ldots, (K-1)$, $y = 0, \ldots, (L-1)$. We assume that sampling noise is negligible, so that these are sampled values on the integer lattice of an unobserved smoothed process (f) in continuous space, with $F_{j,xy} \equiv f_j(x, y)$. We further assume that $f(x, y)$ is a realisation of J stationary, cross-correlated, stochastic processes on a continuous 2D rectangular domain: $0 \leq x < K$ and $0 \leq y < L$.

The Fourier representation of f_j is

$$f_j(x, y) = \sum_{k=-\infty}^{\infty} \sum_{l=-\infty}^{\infty} f^*_{j,kl} \exp\left[2\pi\iota\left(\frac{xk}{K} + \frac{yl}{L}\right)\right] \tag{1}$$

where ι denotes $\sqrt{-1}$ and $f^*_{j,kl}$ is the complex Fourier transform of f_j at frequency $\omega_{kl} \equiv (\omega_{1,kl}, \omega_{2,kl})^T = 2\pi(k/K, l/L)^T$. As $K, L \to \infty$, the Fourier coefficients (f^*) converge in distribution to multivariate complex normal distributions, independently distributed at each frequency except for 180° rotational symmetry (i.e. $f^*_{j,-k,-l} \equiv \overline{f^*}_{j,kl}$, where $\overline{f^*}$ denotes the complex conjugate). For an introduction to complex distributions see, for example, Andersen *et al.* (1995: ch. 2). At frequency ω_{kl}, the J Fourier coefficients, $f^*_{1,kl}, \ldots, f^*_{J,kl}$, which we abbreviate to f^*_{kl}, have zero mean and $J \times J$ complex variance matrix, v_{kl}. So, for example,

$$\operatorname{cov}\left(\begin{bmatrix}\mathcal{R}(f^*_{i,kl})\\ \mathcal{I}(f^*_{i,kl})\end{bmatrix}, \begin{bmatrix}\mathcal{R}(f^*_{j,kl})\\ \mathcal{I}(f^*_{j,kl})\end{bmatrix}\right) = \frac{1}{2}\begin{bmatrix}\mathcal{R}(v_{ij,kl}) & -\mathcal{I}(v_{ij,kl})\\ \mathcal{I}(v_{ij,kl}) & \mathcal{R}(v_{ij,kl})\end{bmatrix} \tag{2}$$

where $\mathcal{R}(z)$ and $\mathcal{I}(z)$ denote the real and imaginary parts of complex variable, z. In particular, terms $v_{jj,kl}$ (for $k, l = -\infty, \ldots, \infty$), which are real, are the power spectrum of f_j, and $v_{ij,kl}$

$(i \neq j)$ is the cross-spectrum between f_i and f_j, which is complex, and can be represented as

$$v_{ij,kl} = |v_{ij,kl}| e^{\iota \phi_{ij,kl}}, \quad \text{where } \phi_{ij,kl} = \arg(v_{ij,kl}) \equiv \tan^{-1}\left(\frac{\mathcal{I}(v_{ij,kl})}{\mathcal{R}(v_{ij,kl})}\right) \tag{3}$$

where $|z|$ and $\arg(z)$ denote, respectively, the modulus and argument (or phase) of complex variable z. Terms $|v_{ij,kl}|$, for $k, l = -\infty, \ldots, \infty$, are the cross-amplitude spectrum, and terms $\phi_{ij,kl}$ are the phase spectrum, between f_i and f_j. The coherency,

$$\rho_{ij,kl} = \frac{|v_{ij,kl}|}{\sqrt{v_{ii,kl} v_{jj,kl}}} \tag{4}$$

lies in the interval [0, 1], and is a measure of correlation between f_i and f_j at frequency ω_{kl}. (For further on all the above, in the 1-D case, see, for example, Bloomfield, 2000:ch. 4.)

Translations in the spatial domain are equivalent to phase differences in the Fourier domain. Therefore, if images are misaligned due to translations, the phase spectra will be linear functions of ω. In particular, if image j is misaligned with respect to image 1 by $\mu_j \equiv (\mu_{1,j}, \mu_{2,j})^T$ pixels (i.e. $\mu_{1,j}$ pixels in the x-direction and $\mu_{2,j}$ pixels in the y-direction), then the misalignment between images i and j will be $(\mu_i - \mu_j)$ pixels, and

$$\phi_{ij,kl} = \omega_{kl}^T(\mu_i - \mu_j) \mod 2\pi \tag{5}$$

The Fourier representation of the observed data, F, is

$$F_{j,xy} = \sum_{k=-(K/2)}^{(K/2)-1} \sum_{l=-(L/2)}^{(L/2)-1} F_{j,kl}^* \exp\left[2\pi\iota\left(\frac{xk}{K} + \frac{yl}{L}\right)\right] \tag{6}$$

where the inverse Fourier transform is

$$F_{j,kl}^* = \frac{1}{KL} \sum_{x=0}^{K-1} \sum_{y=0}^{L-1} F_{j,xy} \exp\left[-2\pi\iota\left(\frac{xk}{K} + \frac{yl}{L}\right)\right] \tag{7}$$

Note that the k and l summations are over finite ranges, unlike in Equation (1), and there are only a finite number of Fourier coefficients, F^*, which are related to f^* by

$$F_{j,kl}^* = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} f_{j,k+Kp,l+Lq}^*, \quad k = -\frac{K}{2}, \ldots, \left(\frac{K}{2} - 1\right), l = -\frac{L}{2}, \ldots, \left(\frac{L}{2} - 1\right) \tag{8}$$

All coefficients at frequencies beyond half the sampling frequency (i.e. $|\omega_1|$ or $|\omega_2| > \pi$), are aliased with coefficients at lower frequencies. Since the Fourier coefficients F^* are linear combinations of asymptotically independently distributed complex random variables f^*, they are also approximately multivariate normally distributed, with zero mean, complex variance V, where

$$V_{kl} = \sum_{p=-\infty}^{\infty} \sum_{q=-\infty}^{\infty} v_{k+Kp,l+Lq} \tag{9}$$

and the log likelihood of F can be approximated by

$$\mathcal{L} = -\frac{1}{2} \sum_{k=-(K/2)}^{(K/2)-1} \sum_{l=-(L/2)}^{(L/2)-1} (\log|V_{kl}| + (F_{kl}^*)^T V_{kl}^{-1} F_{kl}^*) \tag{10}$$

The asymptotic approximation can be improved by tapering the borders of images using a cosine bell before obtaining Fourier coefficients (Bloomfield 2000: section 6.2).

We model the power spectra (v_{jj}) by

$$v_{jj,kl} = g(\omega_{kl}, \alpha_j) \tag{11}$$

and coherencies (ρ_{ij}) by

$$\rho_{ij,kl} = h(\omega_{kl}, \beta_{ij}) \tag{12}$$

where g and h are functions, which will vary with application, with unknown sets of parameters α and β. In combination, equations (5), (11) and (12) specify variance matrices v_{kl} and V_{kl} at all frequencies ω_{kl}. We estimate parameters μ, α and β, by numerically maximising the asymptotic log likelihood ($\mathcal{L}$) given by equation (10). However, note that in specific situations some parameters will be of known value. For example, if data are multiple copies of a single image, then coherencies $\rho \equiv 1$. For colour images produced using, for example, a Bayer colour mosaic filter, the misalignment (μ) between bands will be known.

Once parameters are estimated, we can infer the smoothed process in continuous space, f, and thereby correct for translation and interpolate. We first obtain estimates for the dealiased Fourier terms, conditional on the observed terms, for $k = -K/2, \ldots, K/2 - 1$; $l = -L/2, \ldots, L/2 - 1$:

$$\widehat{f^*}_{k+Kp,l+Lq} = \widehat{v}_{k+Kp,l+Lq} \widehat{V}_{kl}^{-1} F^*_{kl} \quad p = -P, \ldots, P, q = -Q, \ldots, Q \tag{13}$$

where P and Q are chosen to be sufficiently large that $\widehat{v}_{k+Kp,l+Lq}$ is negligible for $|p| > P$ or $|q| > Q$. This follows from Mardia *et al.* (1979: 63). Then we apply a phase adjustment,

$$\widehat{f^*}_{j,kl} \to \widehat{f^*}_{j,kl} \times \exp[-\iota \omega_{kl}^T \mu_j] \tag{14}$$

for all j, k, l, to reverse the translation. Finally, the inverse Fourier transform produces aligned images at sub-pixel resolution, which can be further interpolated using the sinc function.

Application

Model Identification

Before we can apply the methodology of the previous section to the Landsat bands shown in Figure 1, we need to use F^* to identify appropriate models for the power spectra and coherencies. We consider, in turn, displays and plots to identify the power spectra ($v_{jj,kl}$), confirm the model for phase spectra ($\phi_{ij,kl}$) and identify the model for coherencies ($\rho_{ij,kl}$). To do this, we assume at the identification stage that aliasing is not too severe, so

$$V_{kl} \approx v_{kl} \tag{15}$$

for $|\omega_{kl}|$ not too near π.

Figure 3(a) shows an image representation of $|F^*_{1,kl}|^2$, obtained from Landsat band 1 after tapering of the image boundaries using a cosine bell. Because

$$\mathrm{E}|F^*_{j,kl}|^2 = V_{jj,kl} \approx v_{jj,kl} \tag{16}$$

where E denotes expectation, this display helps us choose a model for the power spectrum. As Figure 3(a) and similar displays for the remaining bands (not shown), show approximate circular symmetry, we assume that v_{kl} is simply a function of $|\omega_{kl}|$. Figure 3(b) shows a

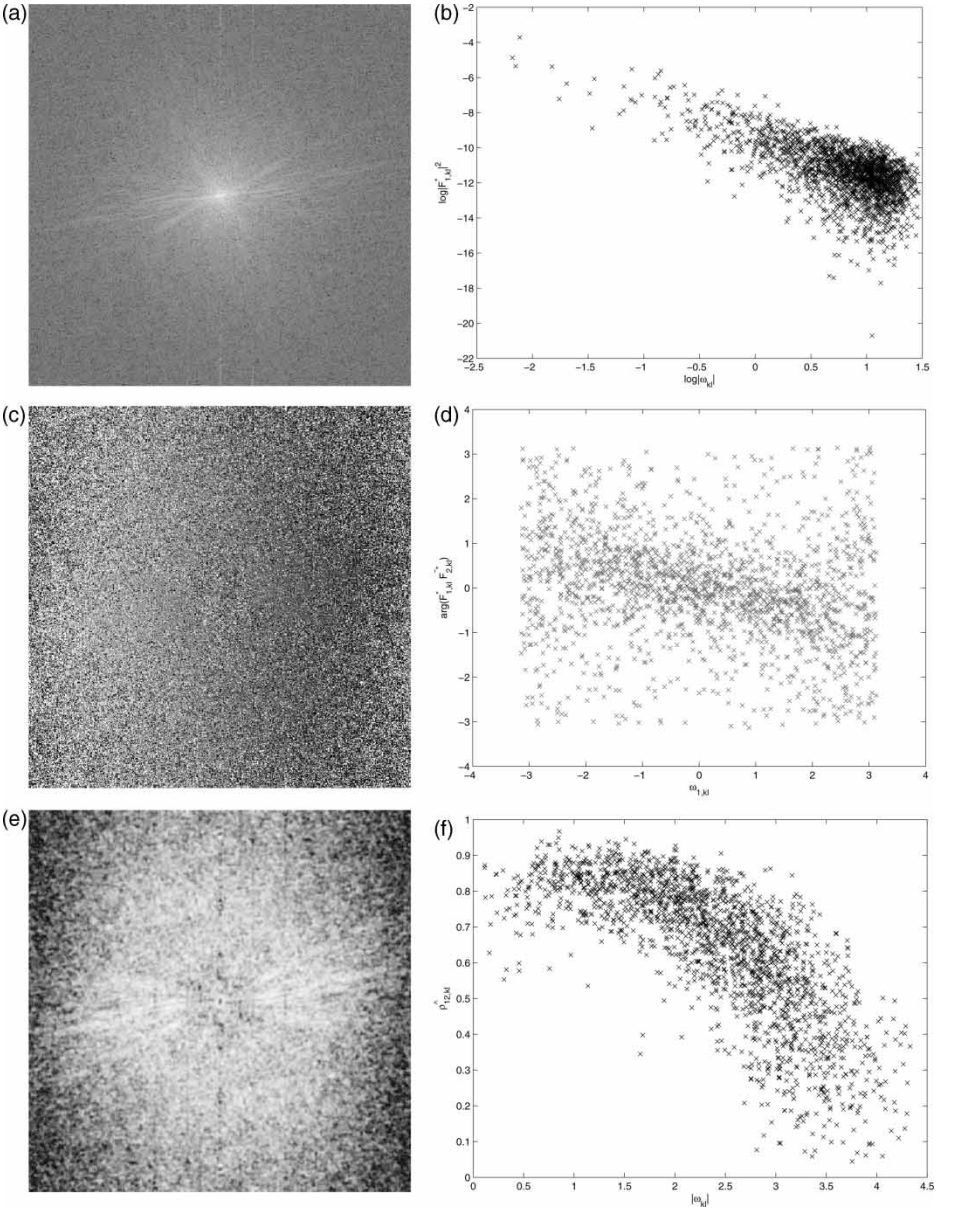

Figure 3. Illustrations, using Landsat in Figure 1, of displays and plots to identify model for v. (a) Power spectrum for band 1: $|F^*_{1,kl}|^2$ displayed as shades of grey (larger values shown as lighter shades), with k indexing columns and l indexing rows. (b) Plot on a log-log scale of 1% of the data in (a), versus $|\omega_{kl}|$. (c) Phase difference between bands 1 and 2: $\arg(F^*_{1,kl}\overline{F^*}_{2,kl})$, displayed as shades of grey (angles are displayed on a monotonically increasing grey scale over the range $[-(\pi/2), (\pi/2)]$, as black for angles $< -(\pi/2)$ and as white for angles $> (\pi/2)$). (d) Plot of 1% of the data in (c), versus $\omega_{1,kl}$. (e) Coherency between bands 1 and 2: $\hat{\rho}_{ij,kl}$, displayed as shades of grey (larger values shown as lighter shades, ranging from 0 as black to 1 as white), obtained from equation (21) by averaging over 5×5 squares. (f) Plot of 1% of the data in (e), versus $|\omega_{kl}|$

log-log plot of $|F^*_{1,kl}|^2$ against $|\omega_{kl}|$ for a random 1% of the values, the linearity of which is supportive of a model of the form

$$v_{jj,kl} = \alpha_{j,1}|\omega_{kl}|^{\alpha_{j,2}} \tag{17}$$

where α are unknown parameters to be estimated. Others have advocated the 2D Matern function (Handcock & Wallis, 1994; Stein, 1999),

$$v_{jj,kl} = \frac{\theta_{j,1}}{(\theta_{j,2} + |\omega_{kl}|^2)^{\theta_{j,3}+1}} \tag{18}$$

but our experience in fitting equation (18) to image data is that $\hat{\theta}_{j,2} \approx 0$ (Glasbey, 2001), which can lead to problems in its use.

Figure 3(c) shows an image representation of the phase differences between Landsat bands 1 and 2, i.e. $\arg(F^*_{1,kl}) - \arg(F^*_{2,kl}) \equiv \arg(F^*_{1,kl}\overline{F^*}_{2,kl})$, which are estimates of $\phi_{12,kl}$ because

$$\mathrm{E}(F^*_{i,kl}\overline{F^*}_{j,kl}) = V_{ij,kl} \approx v_{ij,kl} \equiv |v_{ij,kl}|\mathrm{e}^{\iota\phi_{ij,kl}} \tag{19}$$

A linear trend from left to right can be discerned in Figure 3(c), except towards the edges of the display where the approximation breaks down because of aliasing with higher frequencies. Figure 3(d) shows phase differences plotted against $\omega_{1,kl}$ for 1% of the values, and again a linear trend can be seen. These displays, and similar ones for other comparisons of pairs of Landsat bands, support the linear model for the phase spectra, given in equation (5).

The coherencies (ρ_{ij}) are symmetric in i and j, and $\rho_{jj} \equiv 1$, so we need only consider $i < j$. Our strategy is to model the coherencies only for $i = (j-1)$, and otherwise to model conditional coherencies (γ_{ij}) between non-adjacent bands. For $i \leq (j-2)$, let γ_{ij} denote the coherency between bands i and j conditional on bands $(i+1), (i+2), \ldots, (j-1)$. From the set of coherencies and conditional coherencies at frequency ω_{kl}, Appendix 1 shows how we obtain matrix $|v_{kl}|$, which gives terms in the cross-amplitude spectra. By constraining $\rho \in [0, 1]$ and $\gamma \in [-1, 1]$, we ensure that all $|v|$ are positive definite.

From equations (19), (4) and (16),

$$|\mathrm{E}(F^*_{i,kl}\overline{F^*}_{j,kl})| \approx |v_{ij,kl}| \equiv \rho_{ij,kl}\sqrt{v_{ii,kl}v_{jj,kl}} \approx \rho_{ij,kl}\sqrt{\mathrm{E}|F^*_{i,kl}|^2\mathrm{E}|F^*_{j,kl}|^2} \tag{20}$$

However, unlike for $v_{jj,kl}$ and $\phi_{ij,kl}$, we cannot obtain an estimate for $\rho_{ij,kl}$ using only $F^*_{i,kl}$ and $F^*_{j,kl}$, because $|F^*_{i,kl}\overline{F^*}_{j,kl}| \equiv |F^*_{i,kl}||F^*_{j,kl}|$. Therefore, we instead combine information over a small range of frequencies, to obtain

$$\hat{\rho}_{ij,kl} = \frac{|\sum_{pq} F^*_{i,pq}\overline{F^*}_{j,pq}|}{\sqrt{\left(\sum_{pq}|F^*_{i,pq}|^2\right)\left(\sum_{pq}|F^*_{j,pq}|^2\right)}} \tag{21}$$

where the summations are over the set $\{(p, q)\colon (k-n) \leq p \leq (k+n), (l-n) \leq q \leq (l+n)\}$, a $(2n+1)$-square centred on (k, l) for a small value such as $n = 2$ or 3. Figure 3(e) shows an image representation of the estimated coherency between Landsat bands 1 and 2. As Figure 3(e) and similar displays for other pairs of bands (not shown), exhibit circular symmetry, we assume that the coherency is simply a function of $|\omega_{kl}|$. Figure 3(f) shows a plot of estimated coherency against $|\omega_{kl}|$ for 1% of the values, which suggests a curvilinear relationship, that we can model using a second-order polynomial.

To ensure that coherencies lie in the interval [0, 1] we use a logistic link function. So, for coherencies between adjacent bands:

$$\rho_{(j-1),j,kl} = \frac{1}{1 + \exp\left[\sum_{m=0}^{2} \beta_{(j-1),j,m} |\omega_{kl}|^m\right]} \tag{22}$$

where $\beta_{(j-1),j}$ are unknown parameters to be estimated. For simplicity and to limit the number of parameters, we assume that the conditional coherencies (γ) are each constant over all frequencies, so, to ensure that they lie in the interval $[-1, 1]$, we specify

$$\gamma_{ij,kl} = \frac{1 - \mathrm{e}^{\beta_{ij}}}{1 + \mathrm{e}^{\beta_{ij}}} \tag{23}$$

where β_{ij} are also unknown parameters to be estimated.

Results

We fit the model specified by equations (17), (22) and (23) to the six Landsat bands shown in Figure 1 by numerically maximising $\mathcal{L}$, after having initialised the 47 parameters as follows. The ten parameters (μ) in the relative shift of each band relative to band 1 were all initialised at 0. The 12 parameters (α) in the power spectrum model were initialised by linear regression, ignoring aliasing. The coherency between adjacent bands required 15 parameters and the conditional coherencies a further ten parameters (β). These were all initialised at 0, which implies a coherency between bands i and j of $0.5^{|i-j|}$.

We used a quasi-Newton constrained function minimisation routine with Sequential Quadratic Programming, where an estimate of the Hessian of the Lagrangian function is updated at each iteration using the Broyden-Fletcher-Goldfarb-Shanno (BFGS) formula (Matlab, 2001). The calculation of the likelihood was written in a compiled C-routine, to speed up the estimation process. To simplify computation of $\mathcal{L}$, we assumed f was band-limited at the sampling frequency, i.e. $f^*_{j,kl} \equiv 0$ for $|\omega_1|$ or $|\omega_2| > 2\pi$, and therefore we have twofold undersampling with respect to the Nyquist criterion. We omitted Fourier coefficients at the lowest frequencies ($|\omega| < 0.10\pi$) from $\mathcal{L}$, since some of these terms are very large and so can dominate the estimation of α, whereas frequencies in the range 0.25π to 1.75π are more relevant for interpolation. In addition, to speed up the computations, we used only a random 10% of the Fourier coefficients. These proved to be sufficient to obtain good estimates of all parameters. A single likelihood calculation using 20,000 frequencies took about 4 seconds on a Pentium 4 1500 MHz PC, and the algorithm converged after 100 iterations, and required 5200 evaluations of $\mathcal{L}$

Figure 4 illustrates the model fit by showing plots matching those in Figure 3. The estimated shifts between the bands are given in Table 1. These are consistent with what we saw in Figure 2, with bands 2 and 3 misaligned horizontally by 1/3rd of a pixel relative to band 1, band 5 misaligned vertically by almost half a pixel and band 7 misaligned diagonally. In addition, we see that band 4 is misaligned horizontally by 1/3rd of a pixel.

To illustrate the results of interpolation, Figure 5 shows details of two small regions in the band 7 image. For comparison, we also show the results of bicubic interpolation. The improvement with our method is evident, with the bridge shown as a straighter line and other details more pronounced. Further, in the second enlargement, the blocking of the area is perpendicular to the field, whereas this is not the case with bicubic interpolation.

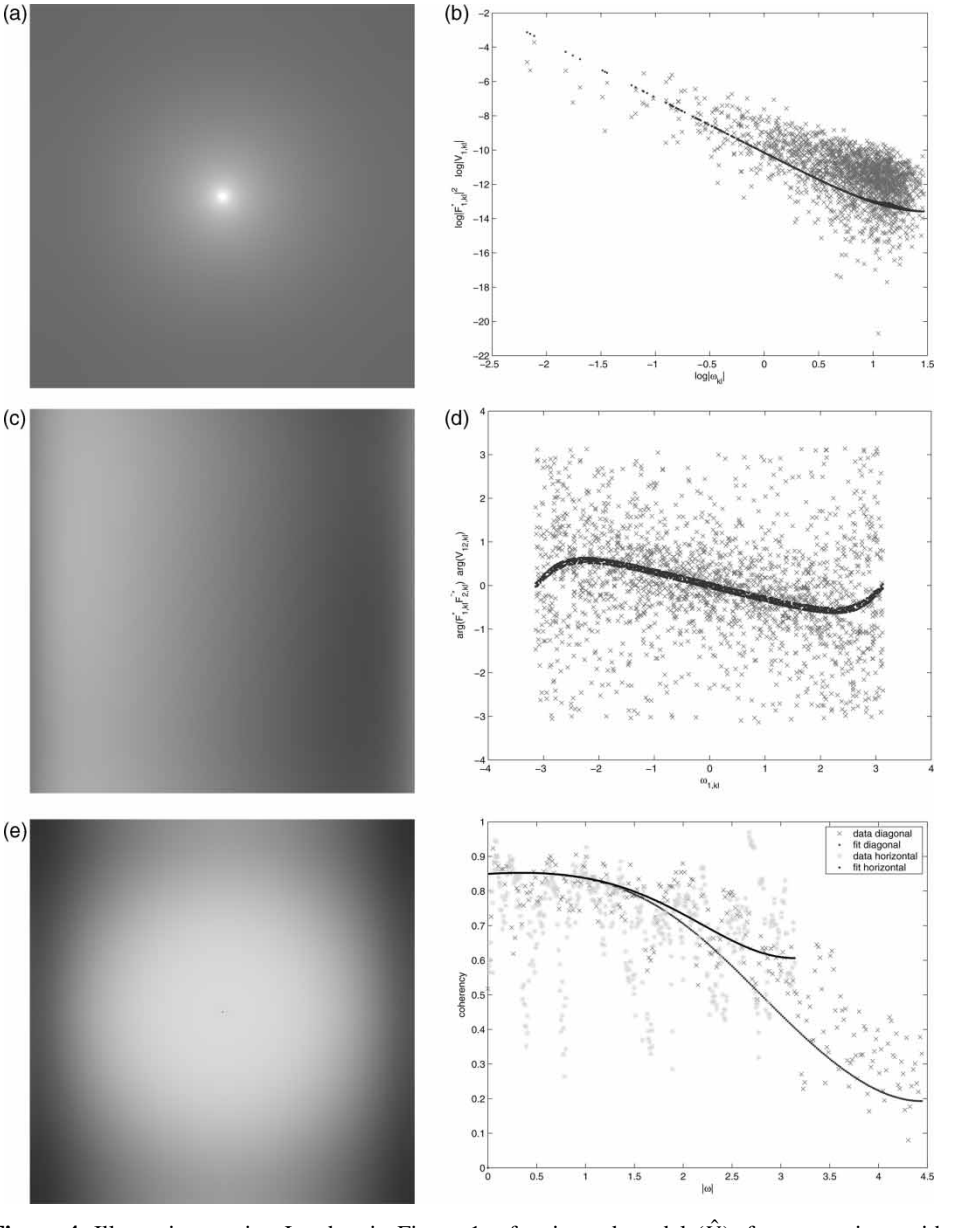

Figure 4. Illustrations, using Landsat in Figure 1, of estimated model ($\hat{V}$), for comparison with Figure 3. (a) Power spectrum for band 1: $\hat{V}_{11,kl}$. (b) Figure 3(b), with fit superimposed. (c) Phase difference between bands 1 and 2: $\arg(\hat{V}_{12,kl})$. (d) Figure 3(d), with fit superimposed. (e) Coherency between bands 1 and 2: $(|\hat{V}_{12,kl}|/\sqrt{\hat{V}_{11,kl}\hat{V}_{22,kl}})$. (f) Subset of Figure 3(f), with fit superimposed

Simulations

To test the method further, a small simulation study was conducted. A four-band image, 128×128 in size, was simulated, with all coherencies set to 0.8, and other parameters set to typical values to match Figure 3. From this, four images of size 64×64 were created by subsampling with a factor of 2 at offsets (0,0), (0,1), (1,0) and (1,1) respectively for the four bands. Parameters were then estimated as above. Our estimated values of $\hat{\mu}$ were all within

Table 1. Estimated misalignment between Landsat band 1 and other bands

$\hat{\mu}_2$	$\hat{\mu}_3$	$\hat{\mu}_4$	$\hat{\mu}_5$	$\hat{\mu}_7$
0.31	0.33	0.36	0.05	0.29
0.03	−0.02	−0.02	0.45	0.47

0.004 of their true values. We interpolated the four bands back to images of size 128 × 128. Table 2 shows the mean square difference between the original and interpolated image for band 1. For comparison, the mean squared difference of other common interpolants are also shown. We see that the new method is superior.

As a final test of the method, we simulated a Bayer colour mosaic filter, by subsampling a colour photograph 670 × 560 pixels in size. We started the green band at location (1,1), (2,1) for red, (1,2) for blue and (2,2) for second green band, with increments of 2 in x and y. We treated the GRBG quadruplet as a four-band image, and fitted the same model as for the Landsat data, except that the misalignment (μ) between bands is known, the coherency between the two green bands is set at 1 and they share a common set of power spectrum parameters, α. Table 3 shows the mean square difference between the original and interpolated images, for each colour band. For comparison the mean squared difference of other common interpolants are also shown, including that proposed by Freeman (1988). Again we see that the new method is superior.

Discussion

We have proposed a method for simultaneously estimating the misalignment between images and interpolating them, while taking account of aliasing. The images need not be identical, since the coherency between them is also modelled and estimated. In the application, we assumed that the images were approximately two-times undersampled, although the method in principle can cope with further undersampling. Further, the method works best if coherency between images is high and misalignment is by a non-integer number of pixels. Simulations have shown that the method outperforms the sinc interpolant if aliasing is present, and it also outperforms standard local interpolants such as b-splines. Furthermore, the simulation showed that the method provides estimates of sub-pixel shifts that are very accurate. Again, for a Bayer colour mosaic image, the method outperforms standard interpolants.

We have explicitly assumed that the smoothed process in continuous space, f, is stationary. This is an approximation, at best, in most applications. Where necessary, the method can be generalised by separately modelling each homogeneous region in a set of images. Implicitly we have assumed that f is approximately Gaussian, in order for the applied linear methods to be optimal. In some cases, nonlinear methods can outperform linear ones in imaging applications. Many nonlinear interpolants exist for single images, but it is less clear how these can be extended to the general multivariate situation considered in this paper.

The estimation method is slow, since it requires an iterative optimisation step, with various matrix operations per frequency. The estimation of the parameters can take many hours, especially when large images with many bands are used. Fortunately, it is not necessary to use all frequencies: about 10,000 randomly selected frequencies gave reasonable estimates in our application. Furthermore, it may not be necessary to estimate the parameters for each image. The interpolation can be used with parameters estimated from another image, taken

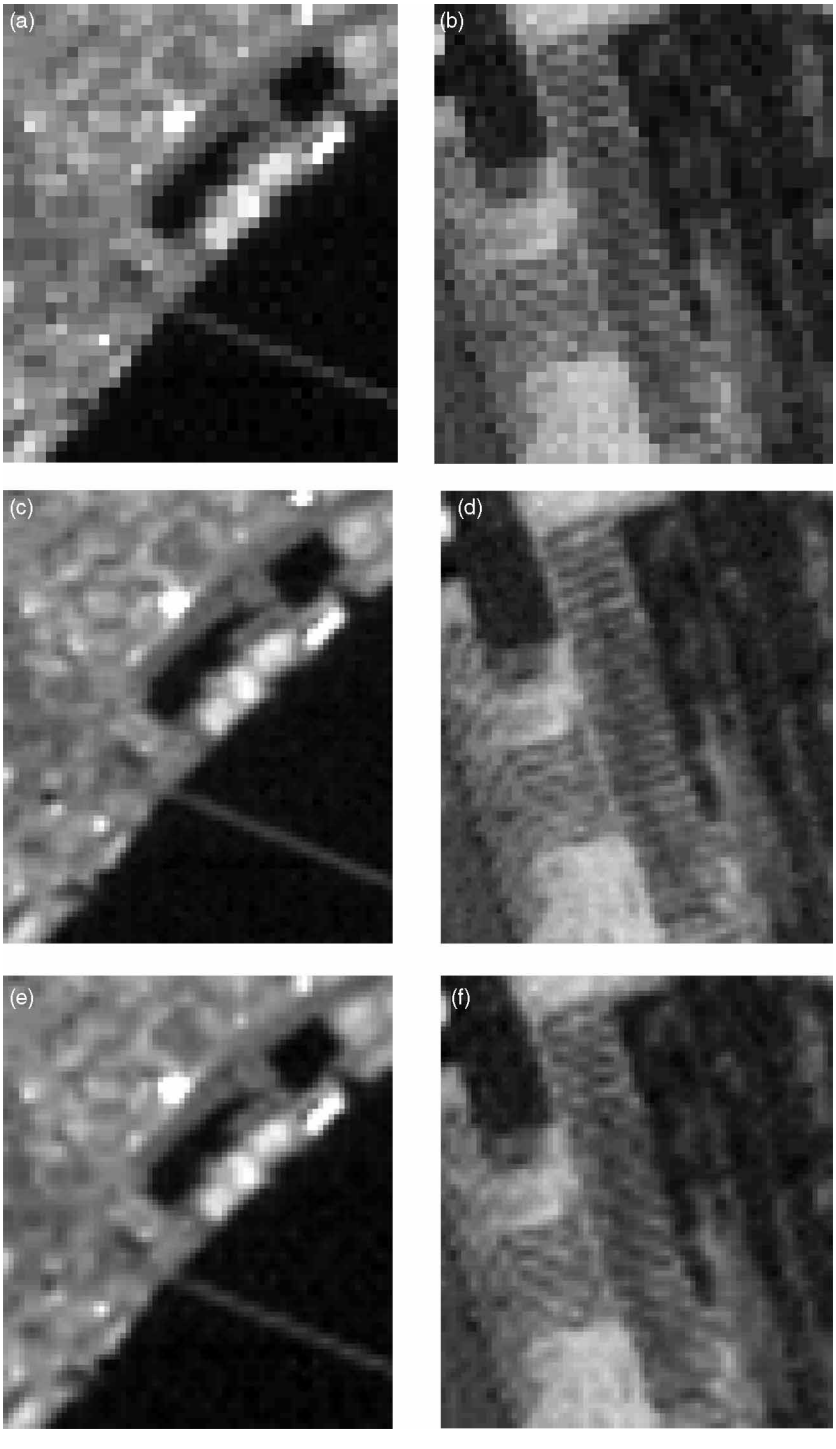

Figure 5. Two examples of interpolation for Landsat band 7: (a) and (b) original images of Tay bridge and pattern of fields; (c) and (d) results of new interpolant; (e) and (f)results of cubic interpolant

Table 2. Mean square difference between band 1 of simulated 128 × 128 image and result of interpolating from a four-band 64 × 64 image

Nearest-neighbour	Sinc	Bilinear	b-spline	Bicubic	New
36	20	17	18	17	13

Table 3. Mean square difference between original colour image and result of interpolating from an artifically generated Bayer colour mosaic filter

	Bilinear	Bicubic	Freeman	New
Red	47	42	33	27
Green	28	22	28	20
Blue	73	70	55	50

with the same device (same shift and amount of aliasing). In that case, the interpolation can be rather fast, although it still requires a forward and reverse Fast Fourier Transform. Clearly, the complexity is too high for this method to be implemented in real time situations. However, in special cases where real time aspects are less important, and a higher resolution is needed (e.g. in satellite images), the proposed method provides a very powerful interpolant, which may be difficult to beat.

Acknowledgements

The work of CAG was supported by funds from the Scottish Executive Environment and Rural Affairs Department, and GWAMH was supported by a travel grant from the Netherlands Organization for Scientific Research (NWO).

References

Andersen, H. H., Højbjerre, M., Sørensen, D. & Eriksen, P. S. (1995) *Linear and Graphical Models: For the Multivariate Complex Normal Distribution* (New York: Springer-Verlag).

Berman, M., Bischof, L. M., Davies, S. J., Green, A. A. & Craig, M. (1994) Estimating band-to-band misregistrations in aliased imagery, *CVGIP: Graphical Models and Image Processing*, 56, pp. 479–493.

Bloomfield, P. (2000) *Fourier Analysis of Time Series: An Introduction*, 2nd edn (New York: Wiley).

Borman, S. & Stevenson, R. L. (1998) Super-resolution from image sequences: a review, in: *Proceedings of the 1998 Midwest symposium on Circuits and Systems*, Notre Dame, IN.

Freeman, W. T. (1988) Median filter for reconstructing missing color samples (US Patent 4,724,395).

Glasbey, C. A. (2001) Optimal linear interpolation of images with known point spread function, in: *Scandinavian Image Analysis Conference – SCIA-2001*, pp. 161–168.

Glasbey, C. A. & Horgan, G. W. (1995) *Image Analysis for the Biological Sciences* (Chichester: Wiley).

Glasbey, C. A. & Mardia, K. V. (2001) A penalised likelihood approach to image warping (with discussion), *Journal of the Royal Statistical Society, Series B*, 63, pp. 465–514.

Handcock, M. S. & Wallis, J. R. (1994) An approach to statistical spatial-temporal modeling of meteorological fields (with discussion), *Journal of the American Statistical Association*, 89, pp. 368–390.

Hunt, B. R. (1995) Super-resolution of images: algorithms, principles, performance, *International Journal of Imaging Systems and Technology*, 6, pp. 297–304.

Kaltenbacher, E. & Hardie, R. C. (1996) High resolution infrared image reconstruction using multiple, low resolution, aliased frames, in: *Proceedings of SPIE*, 2751, pp. 142–152.

Luengo Hendriks, C. L. & van Vliet, L. J. (2000) Improving resolution to reduce aliasing in an undersampled image sequence, in: M. M. Blouke, G. M. Williams, N. Sampat & T. Yeh (eds) *Sensors and Camera Systems for Scientific, Industrial, and Digital Photography Applications. Proceedings SPIE*, 3965, pp. 214–222.

Mardia, K. V., Kent, J. T. & Bibby, J. M. (1979) *Multivariate Analysis* (London: Academic Press).

Matlab (2001) *12.1 Reference manual* (The Mathworks, Inc).

Meijering, E. (2002) A chronology of interpolation: from ancient astronomy to modern signal and image processing, *Proceedings of the IEEE*, 90, pp. 319–342.

Parrott, R. W., Stytz, M. R., Amburn, P. & Robinson, D. (1993) Towards statistically optimal interpolation for 3-D medical Imaging, *IEEE Engineering in Medicine and Biology*, 12(3), pp. 49–59.

Ramanath, R., Snyder, W. E., Bilbro, G. L. & Sander, W. A. (2002) Demosaicking methods for Bayer color matrixs, *Journal of Electronic Imaging*.

Stein, M. L. (1999) *Interpolation of Spatial Data: Some Theory for Kriging* (New York: Springer).

Trussel, H. J. & Hartwig, R. E. (2002) Mathematics for demosaicking, *IEEE Transactions on Image Processing*, 11, pp. 485–492.

Appendix. Computation of Cross-amplitude Spectra from Coherencies

From the set of coherencies and conditional coherencies at a particular frequency, we obtain the matrices of all cross-amplitude spectra, $|v|$, as follows.

We make use of a standard result that, if z_1 and z_2 are vectors, and $(z_1, z_2)^T$ is multivariate normally distributed with mean $(\nu_1, \nu_2)^T$, variance Λ (partitioned into submatrices Λ_{11}, Λ_{12}, Λ_{21} and Λ_{22}), then z_1 conditional on z_2 is multivariate normally distributed with mean $\{\nu_1 + \Lambda_{12}\Lambda_{22}^{-1}(z_2 - \nu_2)\}$, variance $\{\Lambda_{11} - \Lambda_{12}\Lambda_{22}^{-1}\Lambda_{21}\}$. See, for example, Mardia *et al.* (1979: 63).

Recursively for $j = 2, 3, \ldots, J$ we consider in descending order $i = (j-1)$, $(j-2), \ldots, 1$, and apply

$$|v_{ij,kl}| = |v_{ji,kl}| = A_{12} + \gamma_{ij,kl}\sqrt{(v_{ii,kl} - A_{11})(v_{jj,kl} - A_{22})} \tag{A1}$$

where A is a 2×2 matrix. If $i = (j-1)$, then all elements in A are set to zero, and

$$|v_{(j-1),j,kl}| = \rho_{(j-1),j,kl}\sqrt{v_{(j-1),(j-1),kl}\, v_{jj,kl}} \tag{A2}$$

Otherwise, $i < (j-1)$ and elements in A are functions of terms in v_{kl} already computed, expressed by matrix algebra as $A = BC^{-1}B^T$, where B is a $2 \times (j-i-1)$ matrix and C is a $(j-i-1)$-square matrix with

$$B_{1p} = |v_{i,(i+p),kl}|,\ B_{2p} = |v_{j,(i+p),kl}|,\ C_{pq} = |v_{(i+p),(i+q),kl}|,$$
$$\text{for } p, q = 1, 2, \ldots (j-i-1) \tag{A3}$$