Nonrigid registration of hyperspectral and color images with vastly different spatial and spectral resolutions for spectral unmixing and pansharpening

Yuan Zhou, Anand Rangarajan, Paul D. Gader
Department of Computer and Information Science and Engineering, University of Florida
Gainesville, FL 32611, USA
{yuan,anand,pgader}@cise.ufl.edu

Abstract

In this paper, we propose a framework to register images with very large scale differences by utilizing the point spread function (PSF), and apply it to register hyperspectral and hi-resolution color images. The algorithm minimizes a least-squares (LSQ) objective function with an incorporated spectral response function (SRF), a nonrigid freeform deformation applied on the hyperspectral image and a rigid transformation on the color image. The optimization problem is solved by updating the two transformations and the two physical functions in an alternating fashion. We executed the framework on a simulated Pavia University dataset and a real Salton Sea dataset, by comparing the proposed algorithm with its rigid variation, and two mutual information-based algorithms. The results indicate that the LSQ freeform version has the best performance for the nonrigid simulation and real datasets, with less than 0.15 pixel error given 1 pixel nonrigid distortion in the hyperspectral domain.

1. Introduction

Hyperspectral images have important applications in agriculture, forestry, geosciences, and astronomy, as the sensor can capture the reflectance at hundreds of wavelengths ranging from visible to shortwave infrared. However, due to limited amount of incident energy, it also suffers from a low spatial resolution. For example, the Hyperion system onboard the Earth Observing 1 (EO-1) satellite launched in 2000 acquires data covering wavelength 0.4μm to 2.5μm with a 30m spatial resolution [18] while the AVIRIS sensor covers the same spectral range in a flight with a 18m spatial resolution [19]. On the other hand, multispectral (e.g. color) or panchromatic (black and white) images are recorded with very few bands, but they can have a much better spatial resolution. The commercial satellite QuickBird (launched in 2001) can collect panchromatic imagery at 61 centimeter spatial resolution and multispectral imagery at 2.5m.

Combining these two types of images can be used to produce a both spatially and spectrally high-resolution image, a.k.a. pansharpening or image fusion [28, 7, 4, 25, 23, 10]. Also, a high resolution color image can guide finding the subpixel objects that contribute to a hyperspectral pixel, a.k.a. spectral unmixing [28, 25, 27, 26]. Before these applications, we require an efficient registration process that registers these images at subpixel accuracy such that we know which region in the (spatially) high resolution image corresponds to the hyperspectral pixel.

Despite a plethora of works in the pansharpening community, very few works have delved into this registration process [28, 7, 4, 25, 18, 23, 10, 22]. According to the review in [10], various techniques including component substitution, Bayesian methods, and matrix factorization, are proposed for image fusion. However, most of them are validated on simulated or semi-realistic datasets, where pixel correspondence is known in advance (for example, they validated their model on images obtained from spatially and spectrally degradation of an existing hyperspectral dataset). As for those works that mention the registration, they are usually aimed at small scale differences, e.g. two (a pixel corresponds to 4 pixels) in [2], or lack fine scale accuracy [24]. For example, in [24] the authors found control points with correspondence by correlation coefficients from pixels with prominent features (e.g. lakes, rivers), and applied interpolation for the remaining pixels.

If we move to the image registration community, few works are proposed to handle significant scale difference, though remote sensing is a major application area of image registration [29]. Among the many registration methods, they can be mainly categorized as intensity-based and feature-based. The intensity-based method calculates a metric based on the intensities of the images, e.g. least-squares [12] or mutual information (MI) [17, 20, 16]. Also, the intensities can be transformed to the frequency domain, in which case a direct solution can be obtained by phase cor-

relation [15, 6], though it only applies to simple transformations. The feature-based approach is more common in remote sensing as the images are usually very large, registering a number of feature points is more efficient than comparing the intensities of all the pixels. These methods usually select some control points through a scale-invariant Fourier transform [11] or Harris corner detector [8], and approximate the nonrigid transform through the thin-plate spline [3] or Gaussian radial basis function [14]. For example, in [13], a constraint from local linear embedding on the feature points is used in the objective function for registering various airborne images. In [1], the authors selected control points in the reference image by edge detection, found the corresponding points in the test image by template matching, and used the thin-plate spline to approximate the warping for nonrigid registration of satellite images. In [5], the Harris corner detector selected feature points are used in the MI objective for registering airborne infrared images. The problems they are trying to solve have one point in common: the remote sensing images are of similar spatial scales.

In the context of spectral unmixing or pansharpening, the following physical assumptions are usually used while ignored in the previous registration research:

1. The scale difference between two images can be so large that a point spread function (PSF) should be specified in the registration problem. For example, a pixel in the hyperspectral image may correspond to hundreds of pixels in the color image.

2. The spectral difference is also large such that we may need to consider multiple bands of both images and the spectral response function (SRF) in the registration. For example, a color band of a multispectral sensor could cover 100nm [9] while a band of a hyperspectral sensor could only cover 10nm [19].

In addition, there are other challenges specific to this problem. First, the raw data of these airborne hyperspectral images are distorted due to airstream turbulence, earth rotation etc. These are calibrated in post-processing, which brings spatial errors to the spectra collected. Second, the registration accuracy should be high enough for the subsequent applications. It has been shown that the registration accuracy should be within 0.1 - 0.2 of the low-resolution hyperspectral pixel size to ensure correct fusion or unmixing [28]. Fig. 1 uses an example to show the challenges, where the hyperspectral image is the Salton Sea dataset obtained from AVIRIS portal, the color image is obtained from Google Earth. Fig. 1(a) shows the huge scale difference such that a PSF may be needed in the registration; (b) shows neighboring pixels having the same spectra, which is probably caused by spatial calibration and therefore necessitates a nonrigid registration.

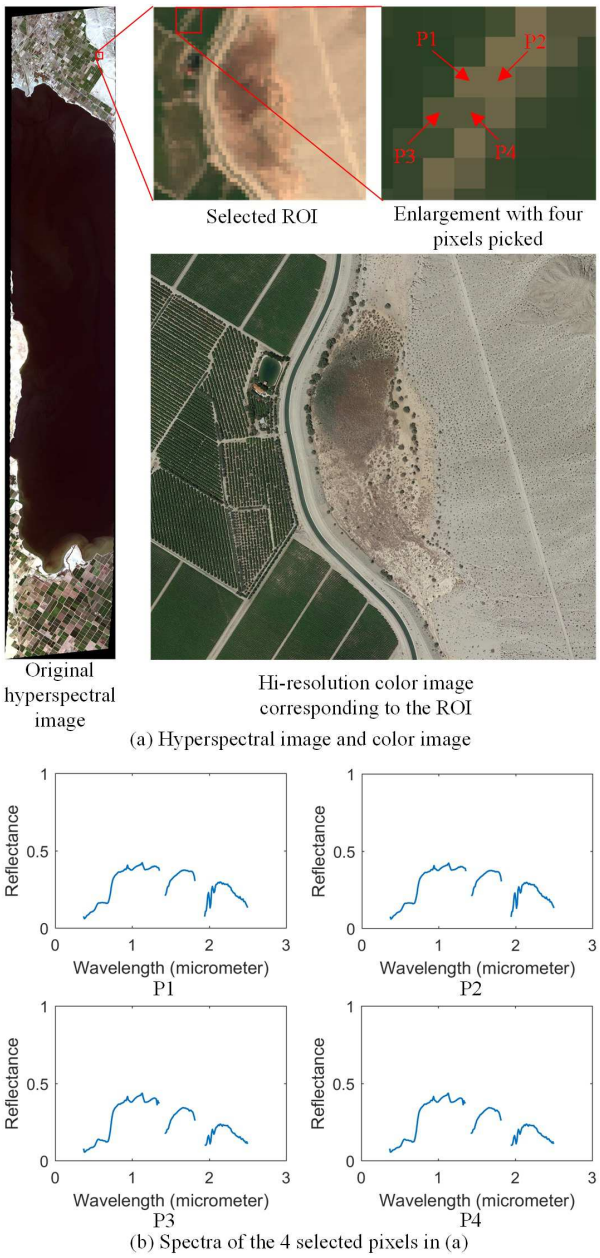


Figure 1. Challenges for registering a hyperspectral image and a color image. In (a) the ROI of the Salton Sea dataset has 56 by 51 pixels with 224 bands (shown as color image by extracting the bands close to 650nm, 540nm and 470nm) while the color image has 738 by 674 pixels with 3 bands. The scale difference is about 10. In (b) the manually picked pixels P1 and P2 have the same spectra, so do P3 and P4. Consider the thinness of the road, it is likely that they are spectra corresponding to the same location.

In this paper, we consider these assumptions and challenges in the registration, propose a least-squares (LSQ) objective function while considering multiple bands of the hyperspectral image with the SRF and the PSF. We apply a

rigid transformation to the color image and a nonrigid transformation to the hyperspectral image, and estimate them simultaneously. The proposed method can handle significant scale difference and spatial distortion with high accuracy. Hence, it can be used for pansharpening, spectral unmixing, and spatial calibration of hyperspectral images.

2. Problem Formulation

We will make the following physical assumptions. Let $\mathcal{D} \subset \mathbb{R}^2$, $\mathbf{x} = (x, y)$,

$$I : \mathcal{D} \to \mathbb{R}_+^B : I(\mathbf{x}) = [I_1(\mathbf{x}), I_2(\mathbf{x}), \ldots, I_B(\mathbf{x})]^T$$

be the hyperspectral image with B bands and $I_k : \mathcal{D} \to \mathbb{R}_+$ be the image at the kth band such that

$$I_k(\mathbf{x}) = \int_{\mathbb{R}^2} g(\mathbf{y} - \mathbf{S}\mathbf{x}) r_k(\mathbf{y}) d\mathbf{y} + n_k(\mathbf{x}), \; k = 1, \ldots, B \tag{1}$$

where $\mathbf{S} = \text{diag}(s_1, s_2) \in \mathbb{R}_+^{2\times 2}$ is the scaling matrix, $g : \mathbb{R}^2 \to \mathbb{R}$ is the PSF that is assumed to be positive and normalized, i.e.

$$g(\mathbf{x}) \geq 0 \, \forall \mathbf{x}, \int_{\mathbb{R}^2} g(\mathbf{x}) d\mathbf{x} = 1. \tag{2}$$

The PSF usually takes the form of a Gaussian function or a constant function over a circular region around the origin; $r_k : \mathbb{R}^2 \to \mathbb{R}_+$ denotes the fine scale reflectance at the kth wavelength; $n_k(\mathbf{x})$ is the noise.

Suppose the color or panchromatic image has b bands,

$$I' : \mathcal{D}' \to \mathbb{R}_+^b : I'(\mathbf{x}) = [I'_1(\mathbf{x}), \ldots, I'_b(\mathbf{x})]^T.$$

When the bandwidths of the hyperspectral sensor are narrow, the color image can be assumed to be a linear combination of the reflectances at selected wavelengths of the hyperspectral sensor,

$$I'_l(\mathbf{x}) = h_{0l} + \sum_{i=1}^{d} h_{il} r_{k_i}(\mathbf{x}) + n'_l(\mathbf{x}), \; l = 1, \ldots, b \tag{3}$$

where $k_1, k_2, \ldots, k_d$ are the selected indices of the B wavelengths, $h_{0l}, h_{1l}, \ldots, h_{dl}$ are the coefficients of the SRF; $n'_l(\mathbf{x})$ is again the noise function. Most research does not include h_{0l} in the SRF, but we use it for more general applications. An example SRF for the IKONOS satellite sensor (1m resolution for panchromatic and 4m for multispectral) can be seen in [9].

Combining (1) and (3) while ignoring the two noise functions, we have

$$\begin{aligned} h_{0l} + \sum_{i=1}^{d} h_{il} I_{k_i}(\mathbf{x}) &= h_{0l} + \sum_{i=1}^{d} h_{il} \int_{\mathbb{R}^2} g(\mathbf{y} - \mathbf{S}\mathbf{x}) r_{k_i}(\mathbf{y}) d\mathbf{y} \\ &= \int_{\mathbb{R}^2} g(\mathbf{y} - \mathbf{S}\mathbf{x}) I'_l(\mathbf{y}) d\mathbf{y}, \end{aligned} \tag{4}$$

where the change of summation and integration uses the property in (2). In the context of image registration, we usually have a reference (fixed) image and a test (moving) image whose coordinates are transformed. For our case, first note that to introduce the PSF, a rigid transformation should be applied to the coordinates of the color image, since we want to combine a correct set of hi-resolution pixels into a low-resolution pixel. For the nonrigid transformation, however, we can not add it to the rigid transformation. This is because the hi-resolution color image has a better spatial accuracy, distorting it will impede applications like image fusion. Hence, we separate the transformation into two parts, and apply the nonrigid transformation to the hyperspectral image. Let $\mathcal{T} : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation on the hyperspectral image, $\mathcal{T}' : \mathbb{R}^2 \to \mathbb{R}^2$ be the transformation on the color image. They are defined by

$$\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathbf{v}(\mathbf{x}), \; \mathcal{T}'(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{t} \tag{5}$$

where $\mathbf{v}(\mathbf{x}) = (u(\mathbf{x}), v(\mathbf{x}))$ is a nonrigid translation field on the coordinates, $\mathbf{A}$ contains only rotation since scaling is incorporated in $\mathbf{S}$ (we ignore shearing for the color image) and $\mathbf{t} = (t_1, t_2) \in \mathbb{R}^2$ is the translation vector. Then, the relation (4) becomes

$$h_{0l} + \sum_{i=1}^{d} h_{il} I_{k_i}(\mathcal{T}(\mathbf{x})) = \int_{\mathbb{R}^2} g(\mathbf{y} - \mathbf{S}\mathbf{x}) I'_l(\mathcal{T}'(\mathbf{y})) d\mathbf{y} \tag{6}$$

for $l = 1, \ldots, b$ and our problem is to find $\mathcal{T}$ and $\mathcal{T}'$ given I and I'.

The major difference of this formulation compared to the traditional registration is the separation of the transformation to the two images. Another difference is the introduction of the SRF and the PSF by separating the scaling from the rotation/translation. Note that the right hand side is actually a convolution if we ignore the scale $\mathbf{S}$. Hence it actually moves the color image and then performs low-pass filtering and downsampling.

3. Nonrigid Registration at fine-scale

We may scale the color image down to a similar scale as the hyperspectral image and perform a nonrigid registration. However, in the context of pansharpening or unmixing, we require the registration to be accurate at a subpixel level, while an initial scaling may not combine the correct hi-resolution pixels. Hence we will solve the fine-scale problem (6) with the PSF and SRF.

3.1. Rigid registration at fine-scale

We first consider the rigid registration and then extend it to the nonrigid case. Let $I'' : \mathcal{D} \to \mathbb{R}^b : I''(\mathbf{x}) = [I''_1(\mathbf{x}), \ldots, I''_b(\mathbf{x})]^T$ be the transformed color image, i.e.

$$I''_l(\mathbf{x}) = \int_{\mathbb{R}^2} g(\mathbf{y} - \mathbf{S}\mathbf{x}) I'_l(\mathcal{T}'(\mathbf{y})) d\mathbf{y}, \; l = 1, \ldots, b \tag{7}$$

where $\mathcal{T}'$ is defined in (5). Assume the PSF has the form of a Gaussian,

$$g(x,y) \propto H\left(\rho - \sqrt{x^2+y^2}\right) e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ determines the shape and ρ is the radius controlling the range of influence with the Heaviside function H [and ρ can be obtained from the instantaneous field of view (IFOV) and the flight height]. We can minimize the squared L_2 norm of the difference function,

$$\mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\sigma) = \int_{\mathcal{D}} \sum_l \left| h_{0l} + \sum_{i=1}^{d} h_{il} I_{k_i}(\mathbf{x}) - I''_l(\mathbf{x}) \right|^2 d\mathbf{x}, \tag{8}$$

w.r.t. $\mathbf{S}$, $\mathbf{A}$, $\mathbf{t}$, σ, where the selected indices $\{k_i,\, i=1,\ldots,d\}$ are determined by the visible range 0.4μm - 0.8μm. We can rewrite the continuous objective function in a discrete form. Let $\mathbf{Y} \in \mathbb{R}^{N\times B}$ be the discretized version of I, $\mathbf{X} \in \mathbb{R}^{N\times b}$ be the discretized version of I''. The selection of d bands can be encoded in a matrix $\mathbf{E} \in \mathbb{R}^{B\times d}$ where for the ith column, only the k_ith row is one while the others are zero. Let $\mathbf{H} := [h_{il}] \in \mathbb{R}^{d\times b}$, $\mathbf{h}_0 := [h_{01},\ldots,h_{0b}]^T \in \mathbb{R}^b$, $\widetilde{\mathbf{H}} := \left[\mathbf{h}_0, \mathbf{H}^T\right]^T \in \mathbb{R}^{(d+1)\times b}$, $\widetilde{\mathbf{Y}} := [\mathbf{1}_N, \mathbf{YE}] \in \mathbb{R}^{N\times(d+1)}$, we can rewrite (8) in the following discrete form

$$\mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\sigma) = \|\mathbf{X} - \widetilde{\mathbf{Y}}\widetilde{\mathbf{H}}\|_F^2. \tag{9}$$

Eq. (9) not only has the unknown registration parameters, but also the SRF $\widetilde{\mathbf{H}}$. We can remove this dependence by solving for it. It is an overdetermined problem to get $\widetilde{\mathbf{H}}$ from (9). A direct solution will introduce non-smooth SRFs. So we add a regularization term to enforce neighboring values to be similar,

$$\begin{aligned}\mathcal{E}\left(\widetilde{\mathbf{H}}\right) &= \|\mathbf{X} - \widetilde{\mathbf{Y}}\widetilde{\mathbf{H}}\|_F^2 + \frac{\lambda}{2}\sum_{i=1}^{d}\sum_{j=1}^{d} w_{ij}\|\mathbf{h}_i - \mathbf{h}_j\|^2 \\ &= \|\mathbf{X} - \widetilde{\mathbf{Y}}\widetilde{\mathbf{H}}\|_F^2 + \lambda \mathrm{Tr}\left(\mathbf{H}^T\mathbf{L}\mathbf{H}\right), \end{aligned} \tag{10}$$

where $\mathbf{h}_i := [h_{i1},\ldots,h_{ib}]^T \in \mathbb{R}^b$ (hence $\widetilde{\mathbf{H}} = [\mathbf{h}_0,\mathbf{h}_1,\ldots,\mathbf{h}_d]^T$). $w_{ij} = 1$ when $|i-j| = 1$ and $w_{ij} = 0$ otherwise. $\mathbf{L} \in \mathbb{R}^{d\times d}$ is the graph Laplacian matrix constructed from $\{w_{ij}\}$ [21]. λ is a parameter controlling the smoothness of the SRF. Taking the derivative of (10) w.r.t. $\widetilde{\mathbf{H}}$ and setting it to zero, we have

$$\widetilde{\mathbf{H}} = \left(\widetilde{\mathbf{Y}}^T\widetilde{\mathbf{Y}} + \lambda\mathbf{L}'\right)^{-1}\widetilde{\mathbf{Y}}^T\mathbf{X}, \tag{11}$$

where $\mathbf{L}' := \mathrm{diag}(0,\mathbf{L}) \in \mathbb{R}^{(d+1)\times(d+1)}$. Plug $\widetilde{\mathbf{H}}$ in (11) back into (9), the objective function becomes

$$\mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\sigma) = \|\mathbf{X} - \widetilde{\mathbf{Y}}\left(\widetilde{\mathbf{Y}}^T\widetilde{\mathbf{Y}} + \lambda\mathbf{L}'\right)^{-1}\widetilde{\mathbf{Y}}^T\mathbf{X}\|_F^2. \tag{12}$$

This is the final objective function used in the optimization.

The objective function (12) involves the parameters for the transform and for the PSF. We can split them into two sets and use block coordinate descent to solve for them. For each iteration, we alternately set

$$\mathbf{S},\mathbf{A},\mathbf{t} \leftarrow \arg\min_{\mathbf{S},\mathbf{A},\mathbf{t}} \mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\sigma),$$

$$\sigma \leftarrow \arg\min_{\sigma} \mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\sigma).$$

For the minimization problem w.r.t. $\mathbf{S},\mathbf{A},\mathbf{t}$, we can also use block coordinate descent, where each individual minimization can be achieved by brute force search. Given an initial coarse-scale registration (e.g. phase correlation), the search can be constrained to a small neighborhood and we can use multiple levels of search with diminishing step size to achieve high accuracy.

3.2. Nonrigid registration using calculus of variations

We can extend the rigid registration in Section 3.1 to the nonrigid case by adding the optimization w.r.t. $\mathbf{v}(\mathbf{x})$. Let $I''_l(\mathbf{x})$ be defined in (7). With the nonrigid transformation, Eq. (8) is written as

$$\begin{aligned}&\mathcal{E}(\mathbf{S},\mathbf{A},\mathbf{t},\mathbf{v},\sigma) \\ &= \int_{\mathcal{D}} \sum_l \left| h_{0l} + \sum_{i=1}^{d} h_{il} I_{k_i}(\mathcal{T}(\mathbf{x})) - I''_l(\mathbf{x}) \right|^2 d\mathbf{x}, \end{aligned} \tag{13}$$

where $\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathbf{v}(\mathbf{x})$. When $\widetilde{\mathbf{H}}$ is given as in (11), this is a functional w.r.t. the translation field $\mathbf{v}(\mathbf{x}) = (u(\mathbf{x}), v(\mathbf{x}))$. For optimization w.r.t. $\mathbf{v}(\mathbf{x})$, we also want the nonrigid transformation to be a smooth function, which leads to an additional constraint $\alpha\int \|\nabla u(\mathbf{x})\|^2 + \|\nabla v(\mathbf{x})\|^2 d\mathbf{x}$ added to (13). Using calculus of variations, we can obtain the necessary condition for its minimization, by setting the following Euler-Lagrange equation to zero,

$$\begin{aligned}\frac{\delta\mathcal{E}}{\delta\mathbf{v}} =& 2\sum_l \left\{ \left(h_{0l} + \sum_{i=1}^{d} h_{il} I_{k_i}(\mathcal{T}(\mathbf{x})) - I''_l(\mathbf{x}) \right) \right. \\ & \left. \left(\sum_{i=1}^{d} h_{il} \nabla I_{k_i}(\mathcal{T}(\mathbf{x})) \right) \right\} - 2\alpha\nabla^2\mathbf{v}(\mathbf{x}), \end{aligned}$$

where ∇ is the gradient operator and ∇^2 is the Laplacian operator. Directly solving $\frac{\delta\mathcal{E}}{\delta\mathbf{v}} = \mathbf{0}$ for $\mathbf{v}(\mathbf{x})$ is usually a difficult problem. We use gradient descent on $\mathbf{v}(\mathbf{x})$, giving us a partial differential equation (PDE) $\frac{\partial\mathbf{v}}{\partial t} = -\frac{\delta\mathcal{E}}{\delta\mathbf{v}}$ to solve. Using forward difference on $\frac{\partial\mathbf{v}}{\partial t}$, we have an update rule

$$\mathbf{v}(\mathbf{x}) \leftarrow \mathbf{v}(\mathbf{x}) - \Delta t \frac{\delta\mathcal{E}}{\delta\mathbf{v}} \tag{14}$$

for $\mathbf{v}(\mathbf{x})$, where Δt should be small enough to ensure a stable solution. Given an initial condition (e.g. $\mathbf{v}(\mathbf{x}) = \mathbf{0}$), we can update $\mathbf{v}(\mathbf{x})$ according to (14). Once it converges, we have a solution to $\frac{\delta \mathcal{E}}{\delta \mathbf{v}} = \mathbf{0}$.

The implementation involves the discretization of the gradient and the Laplacian operator. We use centered difference for the first order derivative and second centered difference for the Laplacian operator. Since the diffusion term $\nabla^2 \mathbf{v}(\mathbf{x})$ originates from the heat equation and we want the smoothing to be isolated within the domain, a homogeneous Neumann boundary condition is used for $\mathbf{v}(\mathbf{x})$ in the PDE.

Combining this optimization with the rigid version, we have our final update rules for the problem (13):

$$\mathbf{S}, \mathbf{A}, \mathbf{t} \leftarrow \arg\min_{\mathbf{S},\mathbf{A},\mathbf{t}} \mathcal{E}(\mathbf{S}, \mathbf{A}, \mathbf{t}, \mathbf{v}, \sigma),$$

$$\mathbf{v}(\mathbf{x}) \leftarrow \arg\min_{\mathbf{v}(\mathbf{x})} \mathcal{E}(\mathbf{S}, \mathbf{A}, \mathbf{t}, \mathbf{v}, \sigma),$$

$$\sigma \leftarrow \arg\min_{\sigma} \mathcal{E}(\mathbf{S}, \mathbf{A}, \mathbf{t}, \mathbf{v}, \sigma),$$

where the second minimization problem is solved by the update rule (14) while the remaining two follow Section 3.1. In the implementation, we only update the rigid parameters for the first few iterations for the sake of efficiency.

4. Results

We compared 4 algorithms in the experiments. Other than the proposed LSQ rigid registration in Section 3.1 and LSQ nonrigid registration with freeform deformation in Section 3.2, we also tried MI as a metric in our registration framework and compared a rigid and a nonrigid version. The parameters α, λ, Δt in our algorithms are fixed as $\alpha = 0.05$, $\lambda = 10^{-3}N$, $\Delta t = 1$, where α has the largest effect on the final result as the larger it is, the smoother nonrigid distortion we get. The two MI-based algorithms are embedded and implemented in our own framework to handle the significant scale difference. The rigid MI algorithm calculates the metric based on the red band of the color image and the closest band to 650nm of the hyperspectral image. The entropy was calculated by histogramming with 64 bins for the marginal distribution. The nonrigid version uses B-splines to model the deformation, with control points spaced at 8 pixels, then refined to 4 pixels, and finally 2 pixels for the first iteration. The remaining iterations use 2 pixel spaced control points from the previous iteration as initial conditions to further fine-tune the parameters.

The registration error can be calculated for the simulated dataset, where the rigid and nonrigid transformations are known. Though various parameters are known, we are especially interested in the pixel match error in the hyperspectral domain, since they affect the subsequent application. According to [28], the pixel match error should be within 0.1 - 0.2 pixels in the hyperspectral domain for meaningful spectral unmixing and pansharpening. To calculate this error, we transform the hyperspectral pixel coordinates to the corresponding points in the hi-resolution image according to the ground truth rigid and nonrigid transformations in (6), then apply the estimated transformations to transform them back to the hyperspectral domain, compare them to the expected coordinates, and calculate the Euclidean distances.

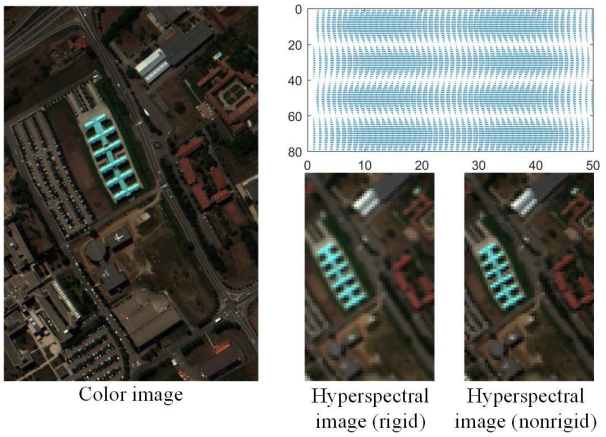


Figure 2. Simulated Pavia University dataset with color image and hyperspectral images (rotated by 5 degrees). The top right corner shows the translation field that was applied for the nonrigidly distorted version. The translation field moves 4 parts of the image by opposite directions, with maximal magnitude of 1 pixel. The distortion is obvious on the blue roof.

4.1. Simulated dataset - Pavia University

The Pavia University dataset was recorded by the Reflective Optics System Imaging Spectrometer (ROSIS) during a flight over Pavia, Italy, in July 2002. The scene is around the Engineering School at the University of Pavia. The image has spatial size 340 by 610 pixels with a resolution of 1.3 meters/pixel. In the spectral domain, it covers wavelengths from 430nm to 860nm by 103 bands. We use it to generate the experimental color and hyperspectral images.

The color image takes the visible bands and uses a Gaussian like SRF covering 120nm for each color band, centered at 650nm, 540nm and 470nm. We use the top part of the image with pixel size 340 by 500 as the final input color image (shown in Fig. 2). The hyperspectral image is generated by rotating the original image by 0 - 10 degrees and scaling it by $\mathbf{s} = (4.4, 4.5)$ with a PSF $\sigma = 10$, $\rho = 3$. The generated hyperspectral image has size 50 by 80 pixels with 103 bands. We considered two cases. For the rigid case, only the rotation and scaling generates the hyperspectral image. For the nonrigid case, a further nonrigid transformation with $\mathcal{T}(\mathbf{x}) = \mathbf{x} + \mathbf{v}(\mathbf{x})$, where $\mathbf{v}(\mathbf{x}) = \sum_k \mathbf{c}_k \mathcal{N}(\mathbf{x}|\boldsymbol{\mu}_k, \sigma^2 \mathbf{I}_2)$, is applied to the aforementioned image. We use 8 Gaussian components to simulate the aircraft instability. The translation field $\mathbf{v}(\mathbf{x})$ and the nonrigidly distorted hyperspectral image are shown in Fig. 2. Both the hyperspectral and color

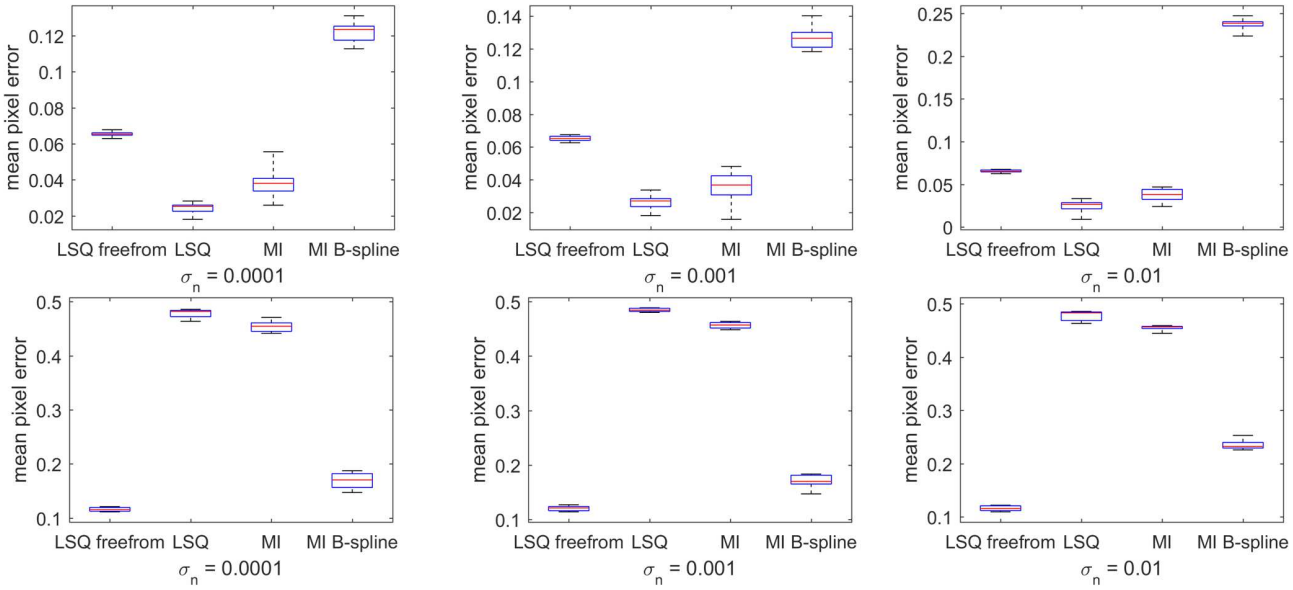


Figure 3. Quantitative comparison of error distribution for the simulated rigid (row one) and nonrigid (row two) Pavia dataset.

images are further contaminated by a zero-mean additive Gaussian noise with standard deviation σ_n on all the bands.

The quantitative errors for all the methods and noise levels are shown in Fig. 3. Among the 11 results for each noise level, we remove the largest 4 errors to better visualize the error statistics. The distribution shows that for the rigid dataset, the proposed rigid version gives the least error, while all except MI B-spline have errors below 0.1. For the nonrigid dataset, the nonrigid version performs best with errors below 0.15, followed by the errors around 0.2 from the MI B-spline approach. The average time costs for the 4 methods are 290, 95, 72, 792 seconds respectively.

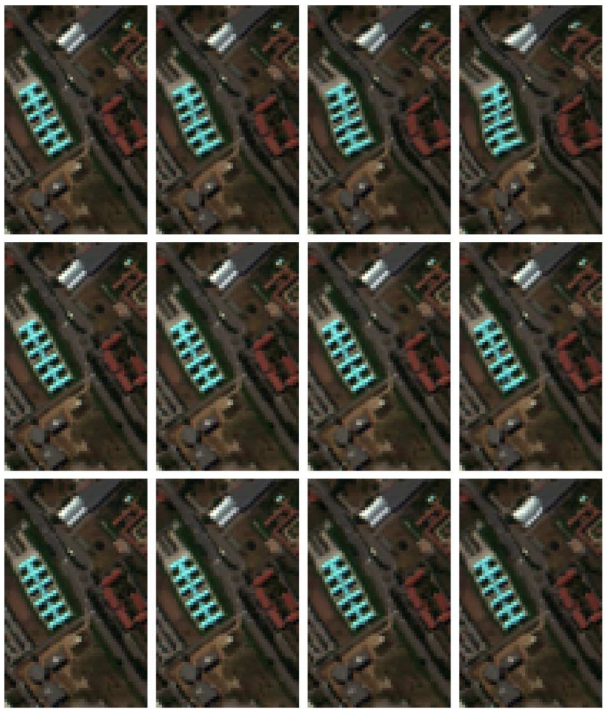
To test the nonrigid estimation part of the two nonrigid methods, we performed experiments on images generated with increasing distortion magnitude following the same pattern in Fig. 2, where the rigid parameters are fixed (assumed to be known) and only the nonrigid parameters are updated. The qualitative results on 4 examples and the quantitative errors for all are shown in Fig 4. We can see that although visually the MI B-spline result for the 4th example is not as good as LSQ freeform, the quantitative error is actually slightly lower. Compared to the result in Fig. 3, it implies that MI B-spline does not perform well in the whole framework despite slightly better nonrigid estimation capability.

4.2. The Salton Sea dataset

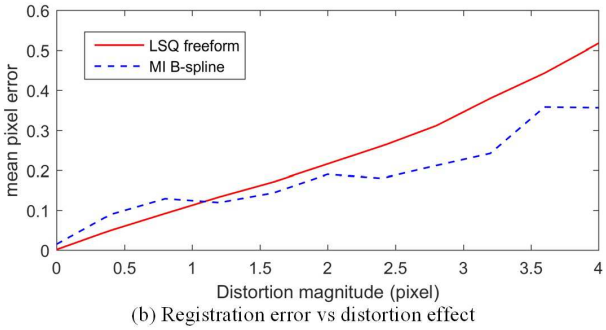
The Salton Sea dataset was collected by the AVIRIS onboard the ER-2 aircraft (20km above the ground) on March 31, 2014. The IFOV for one sample is about 1 milliradian, so each pixel covers approximately a 20-meter diameter region. Given its 16.9m spatial resolution, there is some overlap between the IFOVs of neighboring pixels. We selected a small ROI (56 by 51 pixel with 224 bands) containing vegetation, a river, rooftops and part of a hill. The color image (738 by 674 pixels with 3 bands) is obtained from Google Earth, but the original image was collected in March, 2015. Though there is a one year interval between the two images, they are visually similar and in the same month. We estimated an initial scale $s = 10.4$ between the two images, hence the PSF has $\rho = \lceil (20/16.9) \times 10.4/2 \rceil = 7$. The dataset is shown in Fig. 1.

Fig. 5 shows the registered hyperspectral and color images. We compared the original hyperspectral image and the transformed one from LSQ freeform visually, and marked three differences in red circles. For the top circle, the original one has a thick road segment, while the transformed one has a thinner road. Consider the narrowness of the road in the registered color image. It's clear that the transformed one gives a better spectra distribution. For the other two circles, the boundary between the vegetation and the road is not smooth compared to the boundary in the color image. In the transformed image, they correspond to the real scene better. If we shift to the transformed image from MI B-spline, the result appears to be fuzzier.

Fig. 6 shows the region-spectra correspondences for the arrow marked locations in Fig. 5. For the first column, the green region consists of all vegetation, the blue region consists of mostly road, however, the two regions have the same spectra for LSQ and MI. For LSQ freeform, the green region has spectra between the pure vegetation spectra (red) and the mostly road spectra (blue), which is a better interpretation. The same phenomena repeat for the second and third columns. For the first and fourth columns, the MI B-



(a) Distorted hyperspectral images (row one) and registration results for LSQ freeform (row two) and MI B-spline (row three)



(b) Registration error vs distortion effect

Figure 4. Qualitative and quantitative results on simulated images with increasing distortion effect based on only the nonrigid estimation part of the two nonrigid methods. The four columns in (a) correspond to distortion magnitude 0.4, 1.6, 2.8, 4 pixels.

spline method has very similar spectra for regions with visually different materials, implying that it tends to blur the features.

Since there is no ground truth for this dataset, we need to resort to a new metric for quantitative comparison. Considering that most pansharpening works assume the validity of the linear relationship through the PSF and the SRF, i.e. Eq. (4), we estimate the SRF, reconstruct the low-resolution color image, and calculate the root mean squared error for each pixel and the correlation coefficient for the whole region. Fig. 7 shows the error maps, from which it is clear that the LSQ freeform method achieves the least average error. When it comes to correlation coefficient, it also has the highest value 0.981.

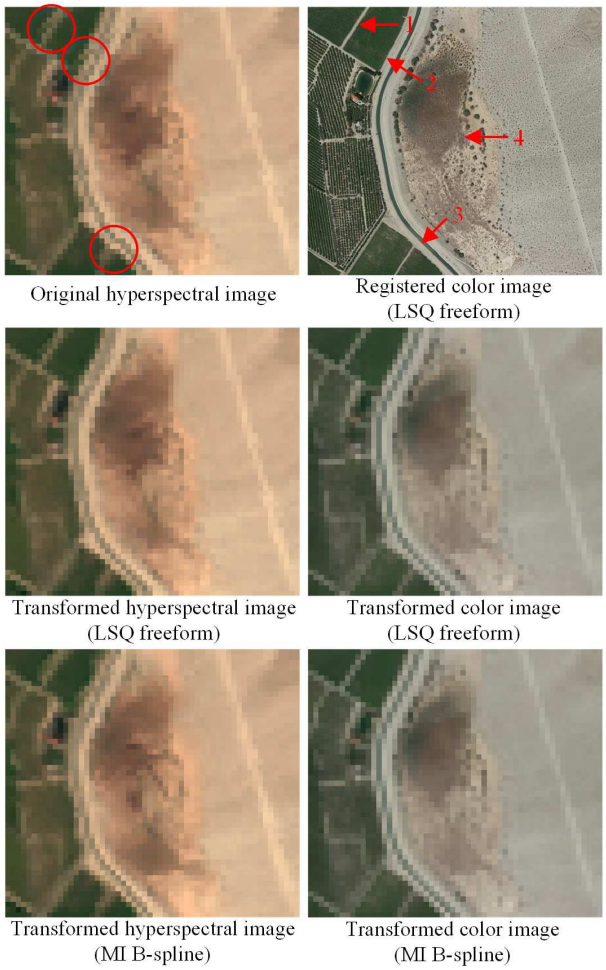


Figure 5. Qualitative results of the two nonrigid methods for the Salton Sea dataset. The 3 circle marked areas in the original hyperspectral image are improved in the transformed image of the proposed nonrigid algorithm. The detailed region-spectra correspondences for the 4 locations specified in the registered color image are shown in Fig. 6.

The running times for this dataset are 206, 121, 102, 905 seconds respectively. Similar to the simulated dataset, though LSQ freeform is slower than the two rigid methods, it is significantly faster than the other nonrigid method.

5. Conclusion

In this paper, we proposed a framework using the PSF to register images with different spatial scales, and applied it to registration of hyperspectral and hi-resolution color images. The algorithm is based on minimizing a LSQ objective function involving the PSF and the SRF, while a freeform deformation is applied to the hyperspectral image and a rigid transformation is applied to the color image. We compared the nonrigid method with its rigid variation and

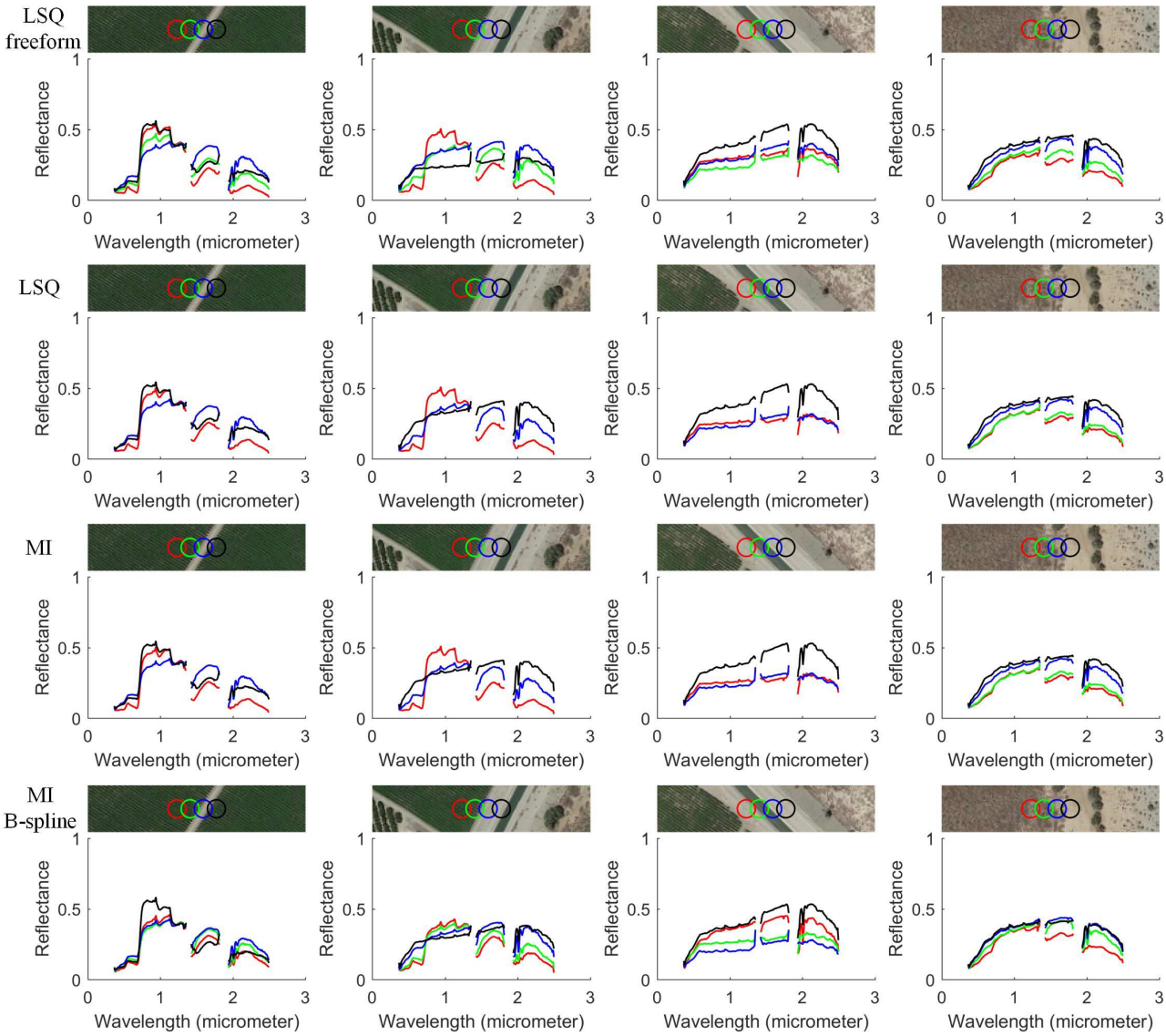


Figure 6. Qualitative region-spectra correspondences for the 4 locations specified in Fig. 5. The four rows are the results from LSQ freeform, LSQ, MI and MI B-spline respectively. For each location magnification, we select 4 contiguous pixels (in the hyperspectral image) and denote the PSF covered regions (in the color image) by 4 different colors. The plot below shows the spectra corresponding to the 4 circled regions above (if the green spectra is missing, it coincides with the blue spectra). We expect to see continuous spectra transition according to the transition of materials.

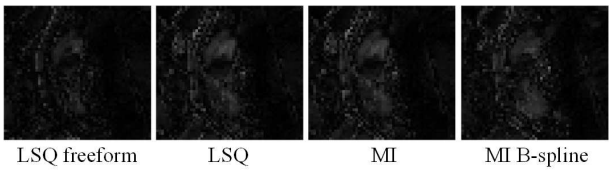


Figure 7. Residual images from reconstruction. The average residual errors are 5.62, 7.52, 7.62, 7.48 for the four methods respectively. The correlation coefficients for the transformed images are 0.981, 0.965, 0.964, 0.967.

two variations using the MI metric, on a simulated dataset generated by the Pavia University dataset, and a real Salton Sea dataset. The results indicate that the proposed nonrigid algorithm has the best accuracy overall (though it does not outperform the rigid version in the rigid simulation, the error is still acceptable). Note that the method is validated on airborne and space images, its application to airborne and ground images is to be further explored. Also, the impact of sensor parameters or terrain surface to accurate registration is still an open question. Future work will verify its application on spectral unmixing and pansharpening.

References

[1] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin. An automatic image registration for applications in remote sensing. *IEEE transactions on geoscience and remote sensing*, 43(9):2127–2137, 2005. 1

[2] P. Blanc, L. Wald, and T. Ranchin. Importance and effect of co-registration quality in an example of pixel to pixel fusion process. In *2nd International Conference Fusion of Earth Data: merging point measurements, raster maps and remotely sensed images*, pages 67–74. SEE/URISCA, Nice, France, 1998. 1

[3] H. Chui and A. Rangarajan. A new point matching algorithm for non-rigid registration. *Computer Vision and Image Understanding*, 89(2):114–141, 2003. 1

[4] M. T. Eismann and R. C. Hardie. Hyperspectral resolution enhancement using high-resolution multispectral imagery with arbitrary response functions. *IEEE Transactions on Geoscience and Remote Sensing*, 43(3):455–465, 2005. 1

[5] X. Fan, H. Rhody, and E. Saber. A spatial-feature-enhanced mmi algorithm for multimodal airborne image registration. *IEEE Transactions on Geoscience and Remote Sensing*, 48(6):2580–2589, 2010. 1

[6] H. Foroosh, J. B. Zerubia, and M. Berthod. Extension of phase correlation to subpixel registration. *IEEE transactions on image processing*, 11(3):188–200, 2002. 1

[7] R. C. Hardie, M. T. Eismann, and G. L. Wilson. Map estimation for hyperspectral image resolution enhancement using an auxiliary sensor. *IEEE Transactions on Image Processing*, 13(9):1174–1184, 2004. 1

[8] C. Harris and M. Stephens. A combined corner and edge detector. In *Alvey vision conference*, volume 15, page 50. Citeseer, 1988. 1

[9] K. A. Kalpoma and J.-i. Kudoh. Image fusion processing for IKONOS 1-m color imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3075–3086, 2007. 2, 2

[10] L. Loncan, L. B. de Almeida, J. M. Bioucas-Dias, X. Briottet, J. Chanussot, N. Dobigeon, S. Fabre, W. Liao, G. A. Licciardi, M. Simoes, et al. Hyperspectral pansharpening: a review. *IEEE Geoscience and remote sensing magazine*, 3(3):27–46, 2015. 1

[11] D. G. Lowe. Distinctive image features from scale-invariant keypoints. *International journal of computer vision*, 60(2):91–110, 2004. 1

[12] W. Lu, M.-L. Chen, G. H. Olivera, K. J. Ruchala, and T. R. Mackie. Fast free-form deformable registration via calculus of variations. *Physics in Medicine and Biology*, 49(14):3067, 2004. 1

[13] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian. Robust feature matching for remote sensing image registration via locally linear transforming. *IEEE Transactions on Geoscience and Remote Sensing*, 53(12):6469–6481, 2015. 1

[14] A. Myronenko and X. Song. Point set registration: Coherent point drift. *IEEE transactions on pattern analysis and machine intelligence*, 32(12):2262–2275, 2010. 1

[15] B. S. Reddy and B. N. Chatterji. An FFT-based technique for translation, rotation, and scale-invariant image registration. *IEEE transactions on image processing*, 5(8):1266–1271, 1996. 1

[16] D. Rueckert, L. I. Sonoda, C. Hayes, D. L. Hill, M. O. Leach, and D. J. Hawkes. Nonrigid registration using free-form deformations: application to breast mr images. *IEEE transactions on medical imaging*, 18(8):712–721, 1999. 1

[17] P. Thévenaz and M. Unser. A pyramid approach to sub-pixel image fusion based on mutual information. In *Image Processing, 1996. Proceedings., International Conference on*, volume 1, pages 265–268. IEEE, 1996. 1

[18] F. D. Van der Meer, H. M. Van der Werff, F. J. van Ruitenbeek, C. A. Hecker, W. H. Bakker, M. F. Noomen, M. van der Meijde, E. J. M. Carranza, J. B. de Smeth, and T. Woldai. Multi-and hyperspectral geologic remote sensing: A review. *International Journal of Applied Earth Observation and Geoinformation*, 14(1):112–128, 2012. 1

[19] G. Vane, R. O. Green, T. G. Chrien, H. T. Enmark, E. G. Hansen, and W. M. Porter. The airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment*, 44(2-3):127–143, 1993. 1, 2

[20] P. Viola and W. M. Wells III. Alignment by maximization of mutual information. *International journal of computer vision*, 24(2):137–154, 1997. 1

[21] U. von Luxburg. A tutorial on spectral clustering. *Statistics and Computing*, 17(4):395–416, 2007. 3.1

[22] Q. Wei, J. Bioucas-Dias, N. Dobigeon, J.-Y. Tourneret, M. Chen, and S. Godsill. Multiband image fusion based on spectral unmixing. *IEEE Trans. on Geoscience and Remote Sensing*, 54(12):7236–7249, December 2016. 1

[23] Q. Wei, N. Dobigeon, and J.-Y. Tourneret. Fast fusion of multi-band images based on solving a sylvester equation. *IEEE Transactions on Image Processing*, 24(11):4109–4121, 2015. 1

[24] N. Yokoya, N. Mayumi, and A. Iwasaki. Cross-calibration for data fusion of EO-1/Hyperion and Terra/ASTER. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(2):419–426, 2013. 1

[25] N. Yokoya, T. Yairi, and A. Iwasaki. Coupled nonnegative matrix factorization unmixing for hyperspectral and multispectral data fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 50(2):528–537, 2012. 1

[26] Y. Zhou, A. Rangarajan, and P. D. Gader. A Gaussian mixture model representation of endmember variability for spectral unmixing. In *IEEE Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS)*, 2016. 1

[27] Y. Zhou, A. Rangarajan, and P. D. Gader. A spatial compositional model for linear unmixing and endmember uncertainty estimation. *IEEE Trans. on Image Processing*, 25(12):5987–6002, 2016. 1

[28] B. Zhukov, D. Oertel, F. Lanzl, and G. Reinhackel. Unmixing-based multisensor multiresolution image fusion. *IEEE Transactions on Geoscience and Remote Sensing*, 37(3):1212–1226, 1999. 1, 1, 4

[29] B. Zitova and J. Flusser. Image registration methods: a survey. *Image and vision computing*, 21(11):977–1000, 2003. 1