Co-registration of LISS-4 multispectral band data using mutual information-based stochastic gradient descent optimization

S. Manthira Moorthi[1,2,*], D. Dhar[1] and R. Sivakumar[3]

[1]Signal and Image Processing Area, Space Applications Centre, ISRO, Ahmedabad 380 015, India
[2]Department of Physics and Nanotechnology, SRM University, Kattankulathur 603 203, India
[3]Department of Civil Engineering, SRM University, Kattankulathur 603 203, India

We propose a solution for automatic co-registration of LISS-4 MX radiometrically conditioned multispectral images issue by considering an optimization problem in which mutual information-based approach is used. Co-registration of multispectral images from the same sensor may also be a tough problem to tackle, when the payload imaging geometry is complex. The multispectral images acquired by ISRO Resourcesat-1/2 LISS-4 MX class of sensors pose such problems and demand an automatic registration solution for system-corrected product generation to cater to user needs. Optical remote sensing image registration is assisted by image geo-referencing or navigation information along with components such as feature detection, matching, correspondence, and resampling the input image to the reference geometry. Intensity-based methods employ an iterative registration framework, where similarity metric based image matching and correspondence is refined to find out optimum transform parameters. We could successfully employ mutual information-based adaptive stochastic gradient descent optimization algorithm to do sub-pixel level satellite image registration tasks by a careful choice of parameters and models related to metric, transform, optimizer, and interpolator in a robust image registration framework which is automatic for different terrain data. The performance is also compared to a recent scale invariant feature transform (SIFT)-based registration method.

Keywords: Image registration, LISS-4, mutual information, optimization.

LISS-4 MX, one of the sensors onboard Indian remote sensing satellite systems Resourcesat-1 & 2 launched in 2003 and 2011 respectively, provides ~5 m data with three spectral bands. The earlier platform had LISS-4 MX swath restricted to 23 km, whereas the Resourcesat-2 offers 70 km (ref. 1). This payload contains three detector configurations corresponding to three spectral bands placed in focal plane at different angles across and along the pitch axis of the spacecraft. LISS-4 camera also has off-nadir viewing capability by steering across-track direction within ± 26 deg. Individual band images are acquired with small, yet finite time gaps while the satellite is driven by a pre-determined yaw profile. In addition, the odd-even pixels are placed in a staggered fashion by a small fixed offset in line direction. Satellite images are corrected for radiometric degradations mainly introduced by detector response non-uniformities usually characterized before launch of the sensor, and the geometric infidelities caused by knowledge uncertainties in the satellite position, attitude, terrain reliefs, etc. Multispectral imaging sensor systems such as LISS-3 in the same platform assure onboard band-to-band registration (BBR) figures within ± 0.25 by design features. On the other hand LISS-4 MX design features (see Figure 1) do not ensure BBR accuracies within ± 0.25. Automated co-registration is one of the primary steps in ground data processing system for LISS-4 MX, irrespective of cloud or terrain conditions under which the images are acquired.

Background and literature survey

Resourcesat-1 LISS-4 MX co-registration complexity was initially reported by Moorthi *et al.*[2], wherein the co-registration was tackled using an in-flight calibration procedure along with attitude refinement for Level-2 corrections that generates geo-referenced data sets. The authors also discussed non-rigid image registration approach of using a thin plate spline model with dense matched points to achieve co-registration. Radhadevi *et al.*[3] reported a method for automatic co-registration of multiple bands of the LISS-4 MX camera using image navigation model and complemented it with an image-based matching technique to remove unaccounted mis-registration residuals. They also indicated that in-flight geometric calibration was a pre-requisite for defining the direction and orientation of imaging ray originating from a pixel and projected on to the correct ground location. Various factors affecting the band-to-band registration

*For correspondence. (e-mail: smmoorthi@sac.isro.gov.in)

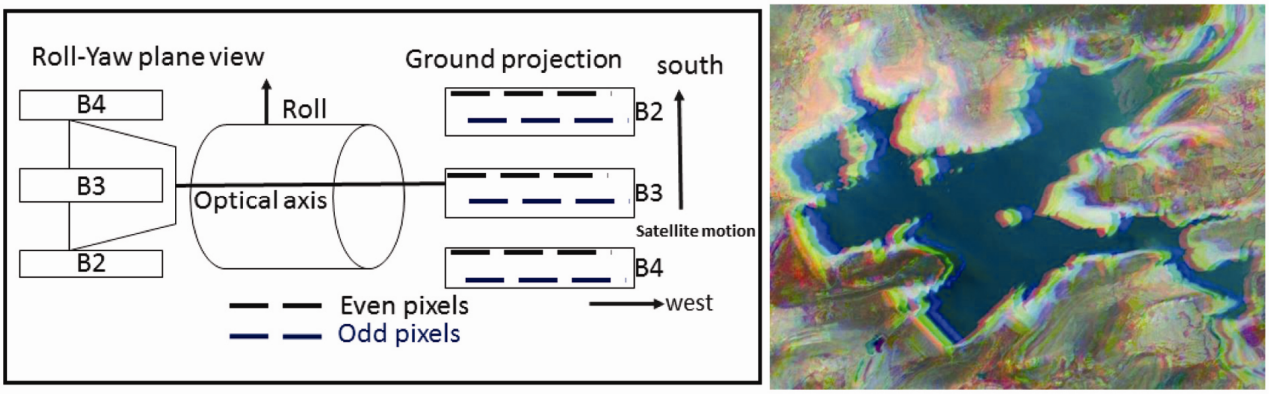

Figure 1. LISS-4 MX sensor schema and image colour composite before registration.

such as detector placement, platform tilt angle, spacecraft velocity effect and feature elevation effects were listed. There was another effort by Pillala *et al.*[4] developing a feature-based registration scheme compensating for residual band-to-band mis-registration after a geometric model update using ground control points. They used mutual information as a similarity measure in place of cross-correlation and a least square procedure to build final transform. Each one of the earlier methods had resolved the varying levels of complexities in the LISS-4 MX band-to-band registration issue at geo-referencing stage alone. They also did not attempt the band-to-band registration at radiometrically conditioned data stage that included a correction step accounting for staggered placement of detectors in the focal plane and photo-response non-uniformity correction across detectors. Similar co-registration problems existed in other remote sensing data systems developed for multispectral thermal imager (MTI) and multi-angle imaging spectroradiometer (MISR)[3].

The current work focuses on automatically co-registering LISS-4 MX three-band data corrected only for radiometric errors, so that we have the advantage of handling single geometry across bands right through geo-referencing. It will reduce complexities further down in the processing chain, if co-registration is achieved before geo-referencing step in which systematic errors are removed and resampled for regular ground sample distance. To be co-registered, remotely sensed images need to be warped and resampled onto some common reference system in a geometric sense. One particular way is to fix one of the band images as the reference or fixed or master image and to register the other image as moving or input image. Reference image in our task is the band-3 (B3) data of LISS-4 MX. This corresponds to visible red channel, which is also the middle of the triad of the channels, with the other two channels such as band-2 (B2) being green and band-4 (B4) being near infra-red, all of them imaged at times of order 1 to 2 second differences. Co-registration of radiometrically conditioned band is somewhat complex due to uncorrected geometric distortions, varying pixel to pixel within the image. Satellite images pose unique challenges for registration with issues like cloud pixels, noise in the images, systematic errors, non-uniform responses, terrain-induced distortions, etc. Few advantages of having an automatic band-to-band registration model at radiometric correction level are: (a) fully automated georeferencing workflow irrespective of input scene considerations; and (b) independent of the orbit, attitude parameters and terrain effects which cause errors pixel to pixel level in image registration. In addition, the automatic procedure should be able to handle some amount of cloud presence say up to 50%.

Two major categories of image registration methods are known as feature based or intensity-based methods and they are used in a variety of problems such as band-to-band registration, multisensor, multiview and multi-temporal data sets. Feature-based methods have a pre-processing step of collecting a number of features uniformly throughout the image regions (point or line or corner points or sub-images) for which the correspondence is established. Intensity-based methods directly use images (their joint histograms) without any preprocessing steps and hence the images can be engaged in registration immediately after acquisition.

Methods based on different combinations of the components used for image registration namely, a feature space, a search space, a search strategy and a similarity metric are employed for practical purposes[5]. A survey of scientific literature on image registration specifically for remote sensing is provided by Fonseca and Manjunath[6]. Previous surveys of the general image registration literature include Brown[5] and Zitova and Flusser[7]. These studies focused mainly on feature-based image registration methods, different methods for collecting features and employing cross-correlation for matching features. Surveys limited to medical image registration, with specific focus on mutual information, include Maintz and Viergever[8] and Pluim *et al.*[9]. In contrast to feature-based approach, intensity-based approaches rely on information

content in the joint histogram of images and the robustness is brought about by employing smarter optimization methods, metrics and transformations. Mutual information (MI) and intensity-based registration was first introduced in the medical imaging domain by Maes *et al.*[10] and Viola and Wells[11]. Readers are referred to studies by Thévenaz and Unser[12] and the review of MI-based methods in medical imaging by Pluim *et al.*[9] for understanding the evolution of this approach. Despite its wide use in medical imaging and good potential for multimodal fusion, MI-based methods are yet to be used extensively in remote sensing applications. Nevertheless, research-oriented articles have demonstrated results on limited image sets[13]. One very important difference to be noticed is that there are many optimization methods used in medical image registration tasks compared to remote sensing image registration methods. Though gradient descent methods are used in both disciplines, adaptive gradient descent methods where the step size is automatically estimated is not found to be reported much in literature for remote sensing images. We present here a particular adaptive stochastic gradient descent optimization method used in LISS-4 MX co-registration task based on mutual information and its gradient.

Intensity-based image registration method

In the registration of pairs of images, one of the images is called the moving or floating or input image f_B which is transformed to confirm geometrically with the other image, the fixed or reference image f_A. The quality of alignment is defined by a cost function, subjected to a transform T. Intensity-based image registration is usually treated as a nonlinear optimization problem. Define the fixed image $f_A(x): \Omega_A \subset R^D \to R$, the moving image, $f_B(x): \Omega_B \subset R^D \to R$, and a parameterized coordinate transformation $T(x, \mu): \Omega_A \times R^P \to \Omega_B$, where $\mu \in R^P$ represents the vector of transformation parameters. The following minimization problem is considered

$$\hat{\mu} = \arg\min C(f_A, f_B \circ T). \tag{1}$$

In eq. (1), $C(f_A, f_B \circ T)$ is the cost function that measures the similarity of the fixed image and the deformed moving image chosen from a multi resolution pyramid. The solution $\hat{\mu}$ is the parameter vector that minimizes that cost function. An optimization technique employed to find optimal transform parameters by maximizing the chosen similarity measure criteria, provides a robust image registration framework[14,15].

In this paper we describe a particular choice of intensity-based image registration that makes use of the following components in an iterative framework. The major components described here are: (a) MI as the similarity metric; (b) combining transformations (affine and Bspline) to take care of global as well as local deformations; and (c) adaptive stochastic gradient descent (ASGD) optimization using MI as cost function.

There are additional components such as multiresolution pyramids, sample selection and interpolation needed to complete the image registration task. It is also essential to remember that many optimization methods such as gradient descent, and simultaneous perturbation use first-order derivatives or their approximation in comparison to Newton methods where second-order derivatives or their approximations of the cost function are used. Subsequent sections bring out details on the individual components mentioned and the complete scheme is depicted later in the article.

Mutual information metric

MI has been known for some time to be effective for the registration of mono modal, as well as multimodal images in medical applications. In general, the correlation-based similarity metric provides reliable registration when the relationship between intensities of the two images to be matched is linear, but mutual information is theoretically more robust to complex variations between the intensities of image pairs, such as those that can occur between pairs of multimodal images[10].

Under rotation or translation, mutual information proved to produce sharper peaks than the most popular image matching metric cross correlation, and MI produces consistently sharper peaks at the correct registration values than correlation because it is less sensitive to noise, which is important for obtaining sub-pixel registration accuracy[13]. Moreover, sharper peaks are also produced at the lowest resolution of the images in the resolution pyramids. This indicates that MI can produce more accurate results than correlation in a multi-resolution registration scheme.

The cost function chosen here is the negative of mutual information $I(A, B)$ between two random variables A and B and is defined as

$$C(f_A, f_B \cdot T) = -I(A, B)$$

$$= -\sum_a \sum_b p_{A,B}\ (a,b) \log \frac{p_{A,B}(a,b)}{p_A(a) p_B(b)}. \tag{2}$$

It measures the distance between the joint probability density function (PDF) $p_{A,B}$ of the random variables A and B and the case associated with complete independence of A and B yielding $p_B(a)$ and $p_B(b)$ as $p_A = \sum_b p_{A,B}$ and $p_B = \sum_a p_{A,B}$, implying that mutual information is zero if and only if the two random variables are independent. It is stated that MI is a measure of the statistical dependence between two datasets indicating the amount of information that random variable A contains about random

variable B, and vice versa. Implementation issues which influence the accuracy of the registration results include estimation of the PDFs, interpolation methods for the PDFs and/or the images, and the choice of optimization methods. Also, the use of a multiresolution approach to increase speed and robustness needs to be considered[12].

The joint probability distribution, given a geometric transformation T, is estimated by first computing a two-dimensional frequency histogram $h(f, m, \mu)$ of the grey value combinations $(f_A(x), f_B(T(x)))$ for every grid point x of the floating image f_B that falls inside the region of overlap of the two images. Joint histogram is represented as a continuous function using Bspline based parzen windowing techniques[16].

The joint histogram is defined in terms of two Bspline window functions $\psi()$, which act as membership functions with f and m being the gray values from fixed and moving images

$$h(f,m,\mu)=\frac{1}{\varepsilon_A\varepsilon_B}\sum_{x_i\in V}\psi\left(\frac{f-f_A\ (x_i)}{\varepsilon_A}\right)\psi\left(\frac{m-f_B(x_i,\mu)}{\varepsilon_B}\right), \tag{3}$$

where ε_A, ε_B are the window scaling factors, V the overlap region between images and μ is the transform parameters. The joint probability is now expressed as

$$\hat{P}(f,m,\mu)=\alpha(\mu)h(f,m,\mu), \tag{4}$$

where

$$\alpha(\mu)=\frac{1}{\left(\sum_{f,m}h(f,m,\mu)\right)}. \tag{5}$$

Using a cubic Bspline Parzen window for the fixed image f_A and the transformation independence of $\hat{P}_A(f)$ and substituting α in eq. (4), the joint PDF can be expressed as

$$\hat{P}(f,m,\mu)=\frac{\alpha}{\varepsilon_A\varepsilon_B}$$

$$\times\sum_{x_i\in V}\beta^3\left(\frac{(m-f_B(T(x_i;\mu)}{\varepsilon_B}\right)\beta^0\left(\frac{f-f_A(x_i)}{\varepsilon_A}\right), \tag{6}$$

where β^n correspond to nth order Bspline function. Now the mutual information is restated as

$$I=\sum_{f=1}^{L_f}\sum_{m=1}^{L_m}\hat{P}(f,m,\mu)\log\left(\frac{\hat{P}(f,m,\mu)}{\left(\hat{P}_A(f)\hat{P}_B(m)\right)}\right). \tag{7}$$

In eq. (7) L_f and L_m are discrete sets of intensities associated with fixed and moving images and we have used 128 bins for the present exercises. Because the marginal PDF of the fixed image A has no influence on the derivative, a simple zero-order Bspline (nearest neighbour interpolation) is used to estimate PDF (ref. 16). Variation of MI value as a function of known translation between two images used in these exercises is depicted in Figure 2 which has a minimum in (0, 0) in x and y axes. The smooth surface shown in Figure 2 represents the search space in and around the optimal transform, which happens to be null translation in this case as this plot was generated for a single image reproduced as fixed and moving images with deliberately shifting them in x and y directions.

Bspline functions are a family of functions with several useful properties, detailed description and their numerical computation are given by Unser[17]. The sum of a Bspline function for all integral distances from a real value is one in addition to the *portion of unity* constraint. This means that no renormalization is required when histogramming. The nth order Bsplines are the convolution of any set of Bsplines whose order sums to n. The derivative of nth order Bspline is a function of two $(n-1)$th order Bsplines. As mentioned earlier, the derivative of cost function required to set up the optimization mechanics for gradient descent type is given in eq. (8).

$$\frac{\partial I}{\partial\mu_i}=\sum_{f=1}^{L_f}\sum_{m=1}^{L_m}\frac{\partial\hat{P}(f,m,\mu)}{\partial\mu_i}\log\left(\frac{\hat{P}(f,m,\mu)}{\hat{P}_B(m)}\right). \tag{8}$$

The derivative of the cost function usually involves computation of the spatial derivative of the moving image: $\partial f_B/\partial x$ which can also be estimated through Bspline image model, and the derivative of the transformation to

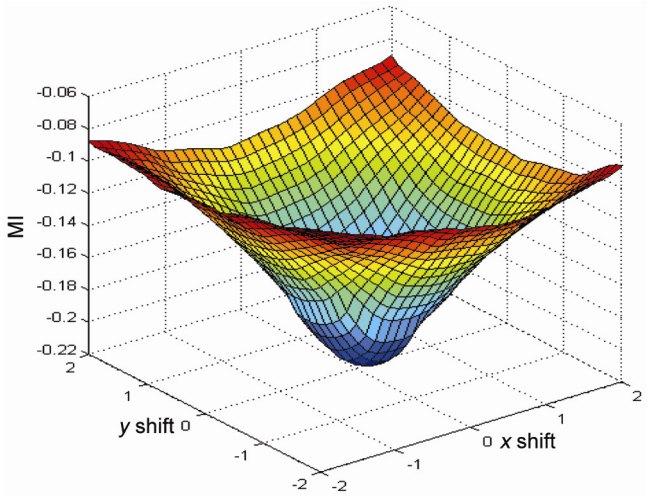

Figure 2. MI metric surface computed for a set of registered images with deliberate translations in x and y directions.

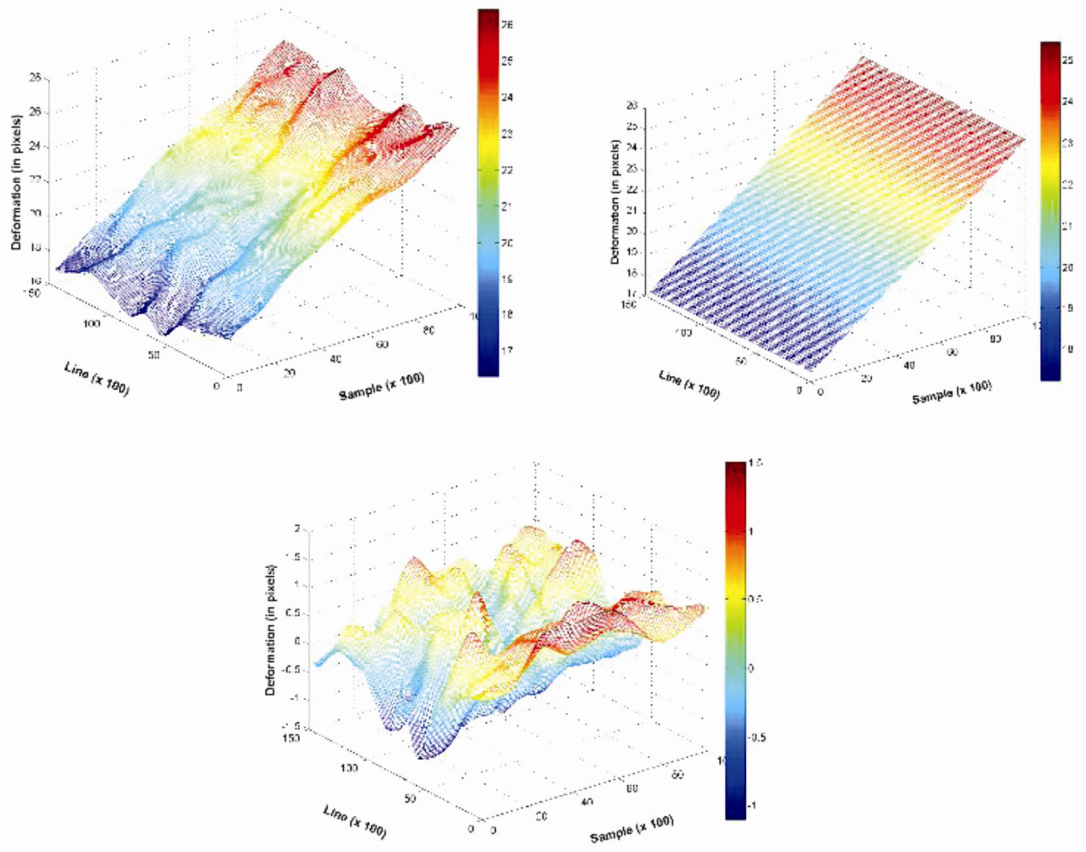

Figure 3. Deformation field portrayed are combined (left most), only affine (middle) and affine field subtracted from combine modelled by Bspline (right most) for the LISS-4 MX data acquired on 3 November 2011.

its parameters: $\partial T/\partial \mu$ called as Jacobian. Further any image can be modelled with Bspline as

$$f(x) = \sum_{x_i \in v} c(x_i)\beta^n(x - x_i), \tag{9}$$

where x_i are the control points, $c(x_i)$ the coefficient vector and v is the set of all control points within the compact support of the Bspline at position vector x. This image model is used to estimate image gradients, and facilitates interpolation of intensities[17].

Compose transform

It is known that translations, rotations, global scaling and local deformations are often necessary to align the images. Satellite images are acquired in a perspective projection, a slant view may be produced as input image, which varies in scale and other geometric errors pixel to pixel. These errors can be systematic or random. The systematic errors correspond to the known (to an order of accuracy) geometry of orbit, attitude, focal plane detector geometry and terrain effects. If the terrain information is not available as a reference data, the terrain-induced distortions cannot be modelled systematically. LISS-4 MX tilted data acquisitions can have terrain-related distortions introduced differently in each band data because of the time differences in imaging. Added to this error, other uncertainties also manifest as random errors which can be modelled only with the use of reference data. These errors can also be typed as global and local errors. The global errors can be modelled with single transformation but the local errors need location based model inputs at a scale corresponding to data itself.

Rigid and affine transformations are a good choice for modelling global errors and Bspline-based models are used to model local distortions considered as free-form deformations (FFD)[18]. A complete transformation model is to combine a global transformation with an additional local transformation to take care of distortions present in only radiometrically conditioned images. This kind of concatenated transforms can be employed to every image registration task, providing scope to model both global and local distortions[19].

Figure 3 depicts the modelled geometric deformation between LISS-4 MX B3 and B2 data, in which the extreme left is the total deformation field, next is just the affine field and the right is the Bspline filed characterized by the registration algorithm discussed here. The uniform

fields are modelled by affine transform and the rest (non-uniform varying and amorphous) is modelled by a Bspline field. We employed an affine transform first and then a Bspline transform as a complete model to take care of global and local distortions[18].

Stochastic gradient descent optimization

We used Klein's algorithm for ASGD optimizer[14], to co-register LISS-4 multispectral data. To solve eq. (1), an iterative optimization procedure was employed. In every iteration k, the current transformation parameters μ_k are updated by taking a step in the search direction d_k

$$\mu_{k+1} = \mu_k + p_k d_k,\ k = 0, 1, 2, \ldots, K, \tag{10}$$

with p_k a scalar that determines the step size. A wide range of optimization methods can be formulated in this way, each having different definitions of p_k and d_k. A common choice for the search direction is the derivative of the cost function $\partial C/\partial\mu$ evaluated at the current position μ_k. In this case, eq. (2) boils down to a gradient descent method. The stochastic gradient descent (SGD) method uses the following iterative scheme

$$\mu_{k+1} = \mu_k - \gamma_k \tilde{g}_k,\ \ k = 0, 1, 2, \ldots, K, \tag{11}$$

$$\tilde{g}_k = g(\mu_k) + \varepsilon_k, \tag{12}$$

where $\tilde{g}_k$ denotes an approximation of the true derivative $g = \partial C/\partial\mu$ at μ_k and ε_k is the approximation error. The distribution of g, $\tilde{g}_k$ and ε_k are independently determined from the selected set of samples. N instances of μ_n are generated around μ_0, and the exact cost function derivative g, the approximated derivative $\tilde{g}_k$, and the approximation error $\varepsilon_k = g(\mu_k) - \tilde{g}_k$ are computed for each of μ_n. The scalar gain factor $p_k = -\gamma_k$, the step size for ASGD, is determined by a predefined decaying function of the iteration number k. The Klein scheme for ASGD was originally proposed by Plakhov and Pedro Cruz[20] in the following way

$$\gamma_k \equiv \gamma(t_k) = a/(k + A)^{\alpha}, \tag{13}$$

$$t_{k+1} = t_k + \text{sigmoid}(-g_k^T g_{k-1}), \tag{14}$$

$$f(x) = f_{\text{MIN}} + \frac{f_{\text{MAX}} - f_{\text{MIN}}}{1 - (f_{\text{MAX}}/f_{\text{MIN}})\text{e}^{-x/\omega}}. \tag{15}$$

Setting values for a, A and α is an important exercise for the use of ASGD for image registration. With user specified constants ($a > 0$), ($A \geq 1$) and ($0 < \alpha \leq 1$), a choice of ($\alpha = 1$) gives a theoretically optimum rate of convergence when $k \to \infty$. The choice of value for a is much more critical, as the convergence of the process accelerates or slows down depending on whether it is large or small. The value for a should be independent of the choice of the cost function and it should be related to the resolution of the images in registration process, so that a meaningful choice of value is made. Another parameter δ is introduced as a user setting allowing a range of values (say a typical value $\delta = 1$ pixel), being the maximum allowed displacement per iteration of the deterministic gradient descent process to compute the stepsize parameter a. Estimates a, f_{MIN} and ω are done only once before the actual optimization starts, with defaultchoice of values $\delta = 1$, $A = 20$, $f_{\text{MAX}} = 1$ no of iterations, etc. Value for A can also be set as one tenth of the number of iterations – in our case it is 1000 iterations. The manual setting of step size can thus be avoided and satisfactory results obtained. If the gradients in two consecutive steps point in the same direction, the inner product will be made positive, which leads to a larger step size and therefore the time is reduced. The whole procedure is called an adaptive SGD optimizer, a detailed description of which can be found in Klein *et al.*[14]. Image pyramids can be used to create a sequence of reduced resolution images from the input image. In general, this coarse-to-fine hierarchical strategy applies to the usual registration methods, but it starts with the fixed and moving images on a coarse resolution[16]. The algorithmic steps for complete registration set-up using MI and ASGD can be summarized as below and depicted in Figure 4.

(1) Set-up registration choices, no. of resolution pyramids, no. of iterations, metric, transform choice, no. of samples for the estimation of metric.
(2) Initialize transform using known or null transform.
(3) Generate distributions through selection of samples for estimating MI and its gradient using the PDF and transform parameters.
(4) Compute step-size and execute the transform update rule according to ASGD.
(5) Check for the condition of exhausting the maximum number of resolution level, maximum number of iterations employed.
(5) Repeat the loop from step number 3.

Performance of ASGD

LISS-4 MX scenes are of 12,000 × 12,000 size covering 70 km × 70 km on the ground, and they may consist of heterogeneous features with different levels of details. Full size images are taken for multiresolution registration exercises with affine and Bspline transformations, and the performance of the procedure is discussed in the next section. However, to see the clear behaviour of ASGD mechanics, a sub-image of size 1024 × 1024 pixels was used

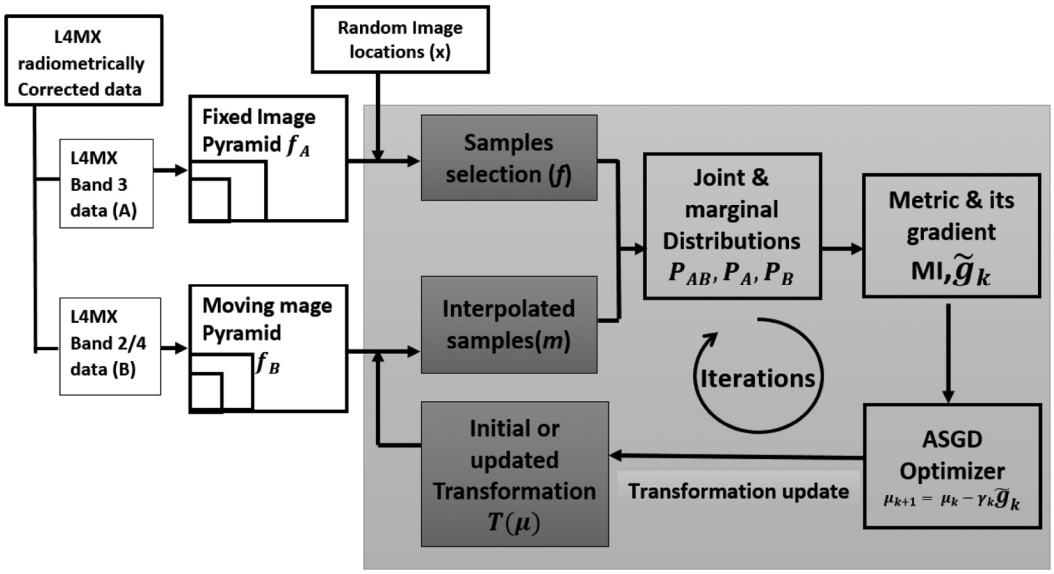

Figure 4. LISS-4 MX multiresolution co-registration scheme using MI and ASGD optimizer.

Figure 5. LISS-4 MX B3 sub image (left) and B2 and B3 checker display before (middle) and after (right) registration by only translation.

(Figure 5) with band 3 as reference, band 2 as image to be registered, with only translation as the transform model and with an initial translation of (15.0, 0.0). The final translation between these two images is estimated by the procedure as (21.29, 2.13). The initial translation was set by manual procedure and multi-resolution pyramid was not used for this sub-image registration case. Figure 5 shows the checker board displays to appreciate the registration achieved by investigating the continuity of linear features. Cost function (negative value of MI), its derivative and the step size values are computed in every iteration to drive the optimization. The step size (slowly decaying function) is automatically estimated using (a) the maximum allowed displacement per iteration as user input (default value is unity), and (b) first derivative of cost function as shown in Figure 6. The translation parameters estimation converges between 20 and 50 iterations for this specific case (Figure 7). Without multiresolution pyramids, the role of the initial transform is important to start with as the procedure works very well only in the neighbourhood of the exact solution. By employing multiresolution pyramids, the capture radius increases to some more extent, and the procedure will find the initial transform also to a reasonable level. With multiresolution set-up, the final translation estimated is (21.21 2.17), which is close to the single resolution with an initial transform setting manually (Table 1). Figure 8 shows the image pyramid and the *x*-shift estimated at every resolution refined at each level, and the same can also be seen from Table 1. Transform parameters estimated at coarse resolution can be multiplied by the resolution factor to arrive at the full resolution transform parameters as shown in the last column of Table 1. The multiresolution approach is more suitable for automatic registration as there will be no initial transform needed to be set-up. Multiresolution approach is more robust as the coarser resolution images are smoothed versions enabling steady convergence of the procedure without getting stuck up at local minima due to reduced noise levels.

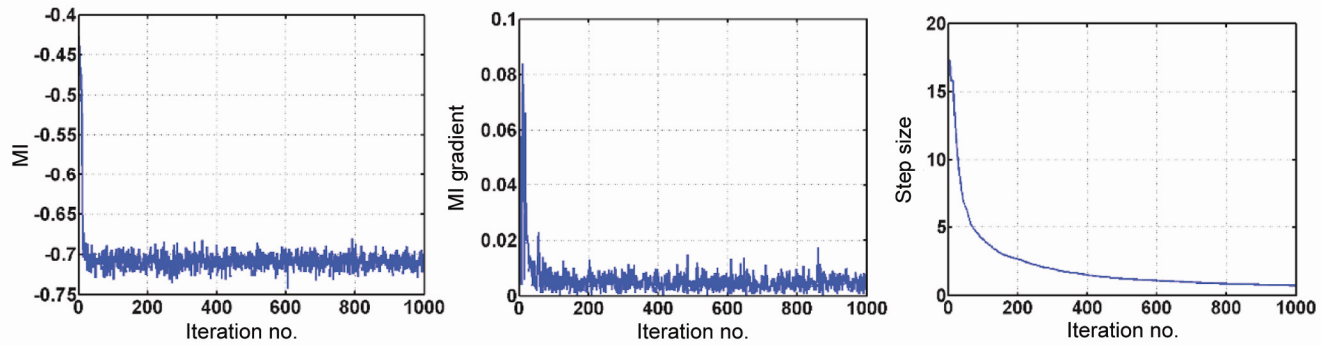

Figure 6. Mutual information, its derivative and step size during the registration process.

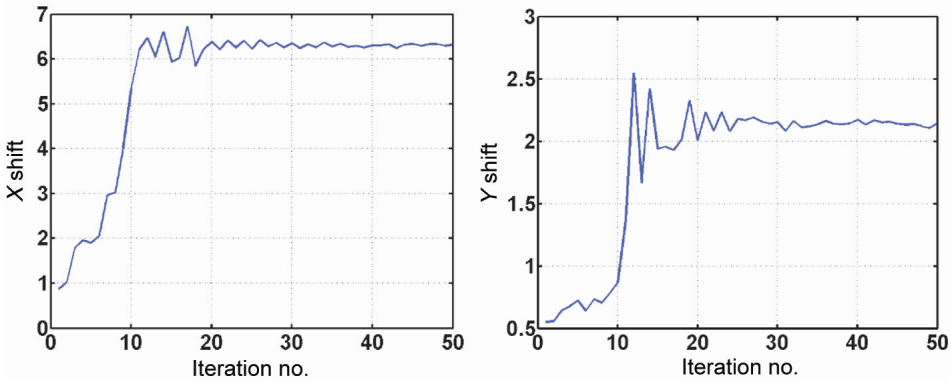

Figure 7. Translation in x and y found in ASGD process.

Table 1. Translation estimated through 4 level image pyramid

Resolution factor, pyramid level	Image size (scans, samples)	Initial transform (x shift, y shift)	Estimated transform over the initial transform	Final transform for full resolution
(8, 0)	(128, 128)	(0.0, 0.0)	(2.68, 0.25)	(21.40, 1.98)
(4, 1)	(256, 256)	(2.68, 0.25)	(0.01, 0.02)	(21.42, 2.00)
(2, 2)	(512, 512)	(5.3, 0.54)	(–0.06, 0.00)	(21.12, 2.71)
(1, 3)	(1024, 1024)	(10.62, 1.05)	(–0.02, 0.07)	(21.21, 2.17)

Performance of ASGD was satisfactory in these exercises with LISS-4 images shown in Figure 9 and listed in Table 2. Typical problems encountered in satellite images, such as cloud pixels, water bodies, etc. are tackled by random sampler choice and the number of spatial samples per iteration near to the 10% to 20% of the total population. In every iteration a fresh set of image samples were taken to estimate the cost function.

LISS-4 MX co-registration exercises and results

The performance of the described approach was tested on the listed LISS-4 MX full scenes and the results are shown in Table 2. The data sets are drawn from terrains typed by urban, undulating and vegetation features with varying degree of cloud pixel occurrence. A combined transform of affine and Bspline was employed to achieve registration to take care of global and local deformations. Many measures were taken (both automatic and manual) to evaluate the accuracy of the registration. As mentioned earlier, B3 was chosen as the reference data to be registered with the input bands being B2 or B4 from LISS-4 MX multispectral data set. Measuring the performance of the registration exercise is a critical step, wherein a global estimate (average position residual in both image axes) is desired generally. The performance can be measured automatically in terms of estimation of transformations that relate both the images, using only translations, or affine relation or complete deformation field. Such estimate can be done using the same procedure used for image registration itself. However, this step is subjected to manual scrutiny to make sure the claim made is true. Remote sensing users adopt a practice to identify control points in both the images after registration and compute

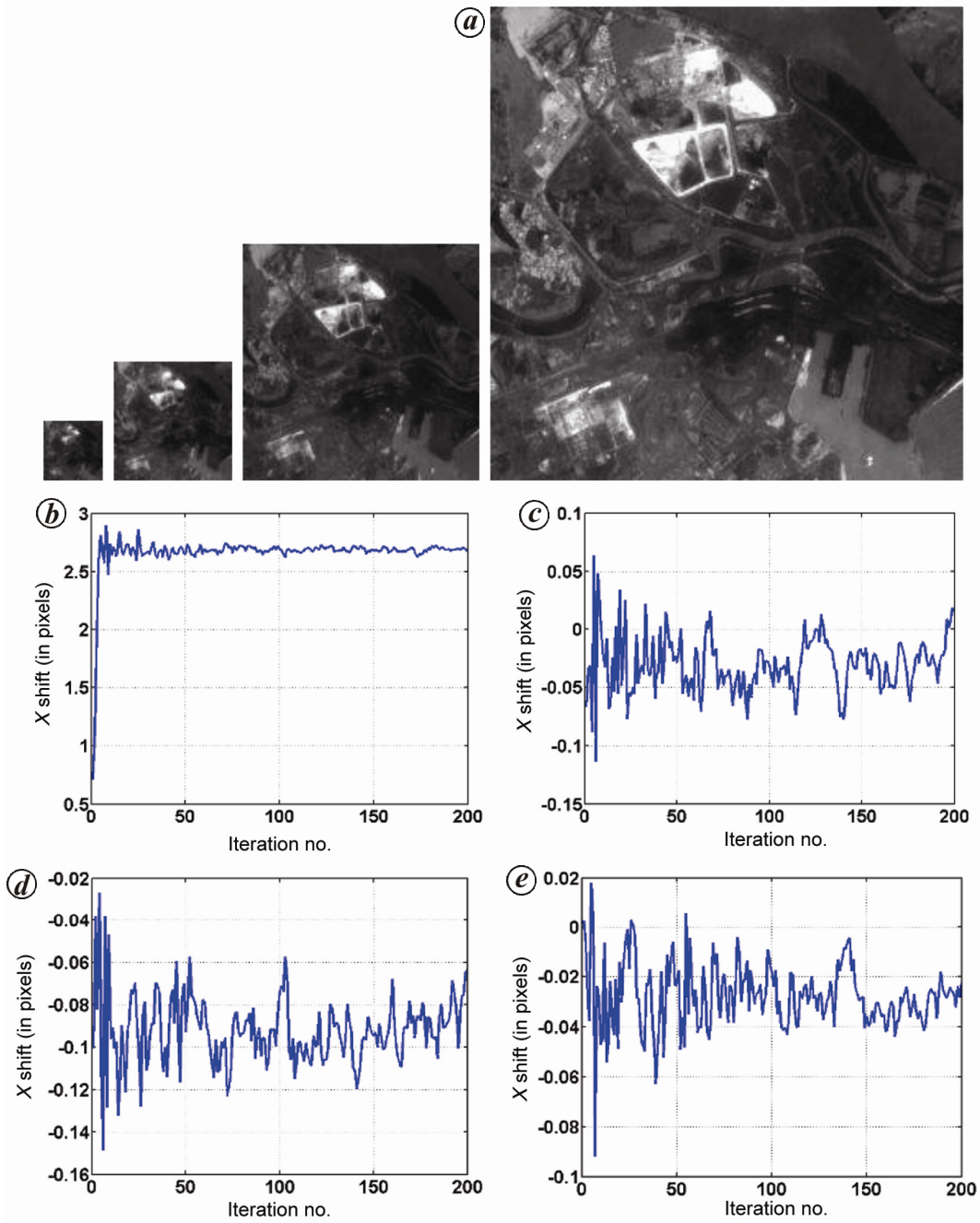

Figure 8. *x*-shift estimated every resolution image pyramid progressively in (***b***) to (***e***) for the image pyramid shown in (***a***).

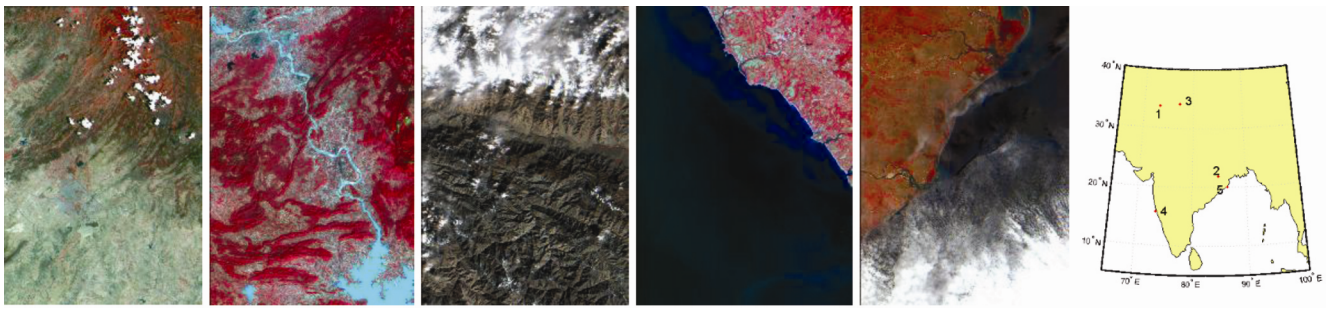

Figure 9. LISS-4 MX data sets (1–5) used in the co-registration exercises with map locations.

Table 2. LISS-4 MX co-registration performance evaluation

Date of imaging	Path/row sub-scene	Parameter	Performance measurement: Before registration B2-B3	Before registration B4-B3	Intensity-based B2-B3	Intensity-based B4-B3	SIFT-based B2-B3	SIFT-based B4-B3
15/06/12 878191	91/47A	*X* shift	20.184	–13.364	0.005	0.010	0.008	0.072
		Y shift	0.669	–1.528	–0.001	0.026	–0.019	0.035
		RMSE	20.239	–13.577	0.46	0.47	0.982	0.882
17/10/11 4101	105/57A	*X* shift	19.851	–13.095	–0.009	–0.005	0.022	0.008
		Y shift	2.097	2.264	0.000	0.030	–0.003	0.206
		RMSE	19.683	–14.200	0.27	0.27	0.763	0.703
03/11/11 1161	94/46D	*X* shift	19.609	–12.986	0.001	0.004	–0.142	0.043
		Y shift	8.346	–6.788	–0.005	–0.006	–0.248	0.260
		RMSE	19.503	–13.220	0.59	0.55	1.305	1.02
08/11/11 8191	95/62B	*X* shift	19.667	–10.337	–0.016	0.016	–0.075	0.226
		Y shift	0.235	–1.244	0.018	–1.239	0.062	0.389
		RMSE	21.453	–12.166	0.457	0.465	0.268	1.04
30/05/122671	107/58C	*X* shift	21.144	–13.980	–0.005	0.000	0.152	–0.275
		Y shift	7.808	–7.536	0.003	–0.018	0.110	–0.003
		RMSE	20.352	–14.483	0.46	0.53	0.859	1.022

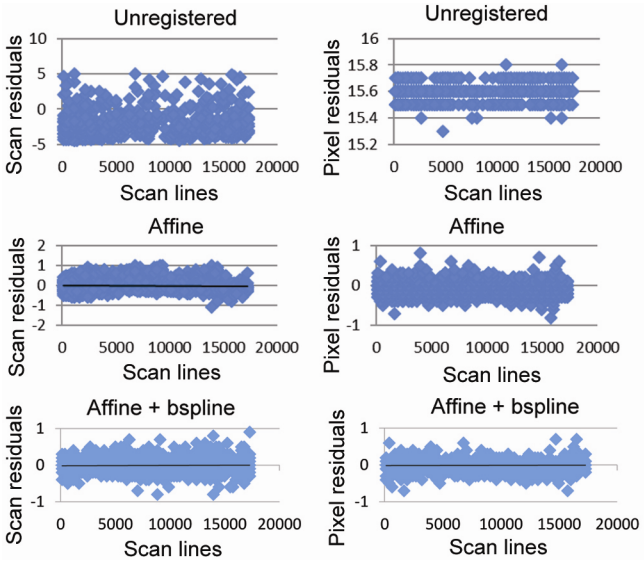

Figure 10. Residual plots for registered, only affine and combined affine and Bspline transform cases as titled for 105/57 A scene.

root mean square error (RMSE) of the residuals which amounts to 2D shifts at known feature points with the outliers removed before estimating the final accuracy. The proposed registration procedure was again used to estimate a translation between the registered images for automatic evaluation and these translation parameters (*X* shift, *Y* shift) are the global indicators of the band-to-band registration residuals. As we can see from Table 2, after registration, the significance in the translations estimated is in second or third decimal places, which means superb registration has been achieved between B2 to B3 and B4 to B3 bands data. It is sufficient to show translation errors up to two or three decimal places to appreciate the registration accuracy. RMSE measurements sensitivity is poor because, they were calculated using landmark or feature point identifications and refined with image correlations (accuracy up to 0.1 pixels) also show error less than 0.5 pixels with the proposed procedure. For comparison, we have also performed registration using a state-of-art scale invariant feature transform (SIFT) technique[21] to collect robust features along with affine transform for registration and its RMSE measures are tabulated in the last column of Table 2. In SIFT-based case, the overall RMSE errors are larger and in some cases even the residual translations are larger compared to the proposed intensity based registration results.

In further understanding the intensity-based registration results, the residual plots (Figure 10) also help us to visually see the results using control points used in the performance measure RMSE. The plots show residuals for unregistered pair, registered using only affine transformation and the combined case (affine plus the Bspline models) throughout the image length. The combined case shows the residuals are very well found within ±0.5 pixels with a trend line along zero axis.

Checker board displays created with reference and input image blocks are very useful to check the registration performance visually (Figure 11). In Figure 11, in each row, the first three columns of images are the checker board display generated without registration and the next three columns of images are with registration. The first and fourth columns of checker board displays depict checker board subimages of B2 to B3 before and after registration exercises, similarly second and fifth columns

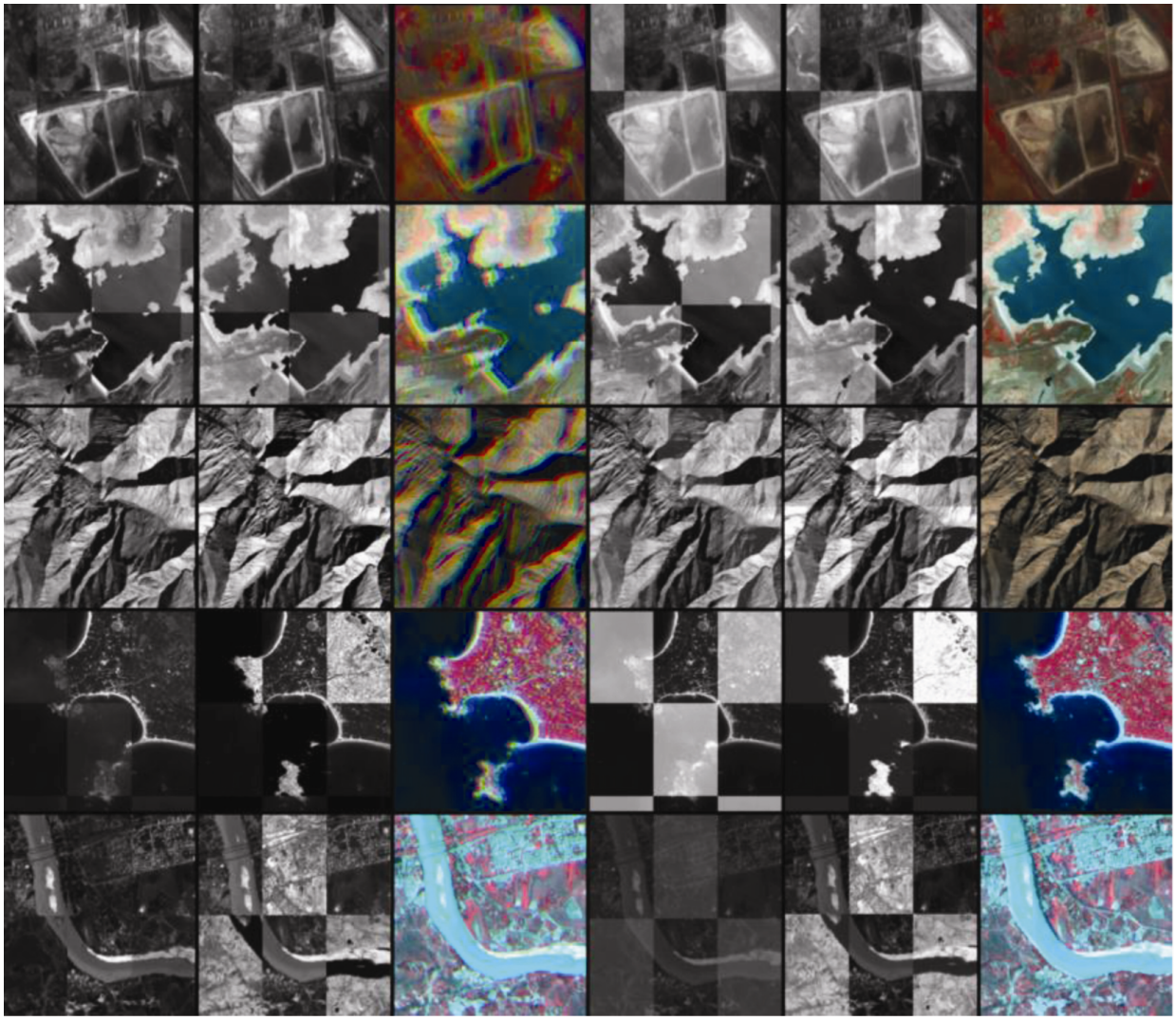

Figure 11. LISS-4 MX co-registration performance in checker board display (B2 and B4 w.r.t. B3) and color composites before (first 3 columns) and after registration (next 3 columns) for a selected portion of data sets in Table 2 row-wise.

of checker board display depict B4 to B3 data and the third and sixth columns display the colour composite using B4, B3 and B2 before and after registration. It is noticeable that after performing image registration the geometric confirmation is better with every terrain in consideration. The colour composites of the LISS-4 MX multispectral band data show the magnitude of co-registration errors manifested as colour artifacts throughout the images and more pronounced in feature boundaries. The registered image portions display the obvious radiometric fidelity compared to the unregistered cases where a fringe of colours that disturbs the interpretation of the features is seen. Each row in Figure 11 belongs to different data sets shown in Figure 9 and Table 2, where water bodies (second row), ocean-land boundaries (fourth row), undulating terrain (third row), and urban features (fifth row) were used to demonstrate the quality of the registration.

Conclusion

Registered multispectral remote sensing datasets are considered as bare minimum requirement for subsequent image analysis, irrespective of play of complex acquisition geometry and uncertainties in the camera geometry models and ancillary information. Satellite data supplying agencies struggle with such problems on a day-to-day basis and mainly rely on feature-based image registration tools. We could successfully employ mutual information based ASGD method to do sub-pixel level LISS-4 MX co-registration tasks by a careful choice of parameters and models related to metric, transform, optimizer, and interpolator in a rigorous and robust image registration framework which is automatic. This approach is even tolerant to a certain level of cloud presence in the images. The approach described here is not exclusive for satellite multispectral images corrected for radiometric errors

alone. The authors have tested the performance of this approach for a variety of satellite image registration tasks, irrespective of its input level of correction, and such results are pending to be reported.

1. ISRO, Resourcesat-2 Data User's Handbook, NRSA Report No. NRSC: SDAPSA: NDCNDC: DEC11-364, 2011, Hyderabad, India.
2. Manthira, M. S., Kayal, R., Ramakrishnan, R. and Srivastava, P. K., RESOURCESAT-1 LISS-4 MX bands on ground co-registration by in-flight calibration and attitude refinement. *Int. J. Appl. Earth Obs. Geoinf.*, 2008, **10**, 140–146.
3. Radhadevi, P. V., Solanki, S. S., Jyothi, M. V., Nagasubramanian, V. and Geeta, V., Automated co-registration of images from multiple bands of Liss-4 camera. *ISPRS J. Photogramm. Remote Sensing*, 2009, **64**, 17–26.
4. Pillala, S. K., Ravikanti, C., Mishra, N., Janja, S. and Geeta, V., A generalized search scheme for automatic registration of remote-sensing data. *Int. J. Remote Sensing*, 2012, **33**, 490–501.
5. Brown, L. G., A survey of image registration techniques. *ACM Comput. Surv.*, 1992, **24**, 325–376.
6. Fonseca, L. M. G. and Manjunath, B. S., Registration techniques for multisensor remotely sensed imagery. *Photogramm. Eng. Remote Sensing*, 1996, **62**, 1049–1056.
7. Zitova, B. and Flusser, J., Image registration methods: a survey. *Image Vis. Comput.*, 2003, **21**, 977–1000.
8. Maintz, J. B. A. and Viergever, M. A., A survey of medical image registration. *Med. Image Anal.*, 1998, **2**, 1–36.
9. Pluim, J. P. W., Maintz, J. B. A. and Viergever, M. A., Mutual-Information-based registration of medical images: a survey. *IEEE Trans. Med. Imaging*, 2003, **22**, 986–1004.
10. Maes, F., Collignon, A., Vandermeulen, D., Marchal, G. and Suetens, P., Multimodality image registration by maximization of mutual information. *IEEE Trans. Med. Imaging*, 1997, **16**, 187–198.
11. Viola, P. and Wells III, W. M., Alignment by maximization of mutual information. *Int. J. Comput. Vis.*, 1997, **24**, 137–154.
12. Thévenaz, P. and Unser, M., A pyramid approach to sub-pixel image fusion based on mutual information. In Proc. IEEE Int. Conf. Image Processing, Lausanne, Switzerland, 1996, vol. 265–268, pp. 16–19.
13. Cole-Rhodes, A. A., Johnson, K. L., LeMoigne, J. and Zavorin, I., Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient, *IEEE Trans. Image Process.*, 2003, **12**, 1495–1511.
14. Klein, S., Pluim, J. P. W., Staring, M. and Viergever, M. A., Adaptive stochastic gradient descent optimisation for image registration. *Int. J. Comput. Vis.*, 2009, **81**, 227–239.
15. Klein, S., Staring, M., Murphy, K., Viergever, M. A. and Pluim, J. P. W., Elastix: a toolbox for intensity-based medical image registration. *IEEE Trans. Med. Imag.*, 2010, **29**, 196–205.
16. Mattes, D., Haynor, D. R., Vesselle, H., Lewellen, T. K. and Eubank, W., PET–CT image registration in the chest using free-form deformations. *IEEE Trans. Med. Imag.*, 2003, **22**, 120–128.
17. Unser, M., Splines: a perfect fit for signal and image processing. *IEEE Signal Process. Mag.*, 1999, **16**, 22–38.
18. Rueckert, D., Sonoda, L. I., Hayes, C., Hill, D. L. G., Leach, M. O. and Hawkes, D. J., Nonrigid registration using free-form deformations: application to breast MR images. *IEEE Trans. Med. Imag.*, 1999, **18**, 712–721.
19. Goshtasby, A. A., Registration of image with geometric distortion. *IEEE Trans. Geosci. Remote Sensing*, 1988, **26**, 60–64.
20. Plakhov, A. and Cruz, P., A stochastic approximation algorithm with step size adaptation. *J. Math. Sci.*, 2004, **120**, 964–973.
21. Li, Q., Wang, G., Liu, J. and Chen, S., Robust scale-invariant feature matching for remote sensing image registration. *IEEE Geosci. Remote Sensing Lett.*, 2009, **6**, 287–291.

ACKNOWLEDGEMENTS. The authors thank Director, Space Applications Centre, ISRO, Ahmedabad who motivated us in every scientific endeavour. We thank SRM university for encouraging the external research programmes.

Received 5 January 2017; revised accepted 3 April 2017

doi: 10.18520/cs/v113/i05/877-888