Robust Automatic Registration of Multimodal Satellite Images Using CCRE With Partial Volume Interpolation

Mahmudul Hasan, *Student Member, IEEE*, Mark R. Pickering, *Member, IEEE*, and Xiuping Jia, *Senior Member, IEEE*

***Abstract*—One of the most important steps in data fusion is image registration. Automatic image-to-image registration for images captured by different sensors traditionally requires the use of information-theoretic similarity measures such as mutual information. Recently, a new similarity measure known as cross-cumulative residual entropy (CCRE) has been proposed for multimodal image registration in medical imaging applications. In this paper, we investigate the use of CCRE for multisensor registration of remote sensing imagery. In particular, we investigate the extreme case of registering synthetic aperture radar images to optical images. We also propose a novel extension to the Parzen-window optimization approach proposed by Thévenaz which involves applying partial volume interpolation in the calculation of the gradients of the similarity measure. Our experimental results show that our proposed approach which uses CCRE as the similarity measure and partial volume interpolation in the optimization procedure provides superior performance to other approaches investigated.**

***Index Terms*—Cross cumulative residual entropy (CCRE), image registration, multimodal, partial volume interpolation.**

I. Introduction

FOR REMOTE sensing of the environment, there are two popular types of sensors, namely multispectral and hyperspectral scanners, which record reflected solar irradiation in the optical wavelength range of 400 to 2400 nm, and synthetic aperture radar (SAR), which records the backscattering from the earth's surface of microwave signals sent out from the sensor. Due to the large differences in their imaging principles, these two sensors reveal very different characteristics of the earth's surface. Several studies on combining relevant information from these two types of imagery have been conducted in recent years, for example see those reported in [1], [2]. These studies aim to better interpret the imaged area for applications such as land cover mapping, change detection, and disaster management. The process of combining the use of two or more image data sets is referred to as data fusion [3], [4].

Manuscript received January 31, 2011; revised June 17, 2011 and October 21, 2011; accepted January 15, 2012. Date of publication March 21, 2012; date of current version September 21, 2012.

The authors are with the School of Engineering and Information Technology, University College, The University of New South Wales, Canberra, ACT 2600, Australia (e-mail: mahmudul.hasan@student.adfa.edu.au; m.pickering@adfa.edu.au; x.jia@adfa.edu.au).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TGRS.2012.2187456

One of the most important steps in data fusion is image registration. When image data are recorded by sensors on satellites or aircraft, they are typically distorted geometrically. The sources of this geometric distortion include the rotation of the earth, the curvature of the earth's surface, and uncontrolled variation in the position and attitude of the remote sensing platform [5]. This distortion needs to be rectified before data fusion can be conducted. When there is a corrected image available, the rectification is often performed via registering to this reference image. The main idea is to find a way to warp the distorted image so that it becomes spatially aligned with the reference image at all pixel locations.

Conventional feature-based methods of registration assuming the mapping functions used to rectify geometric distortion are polynomial and their coefficients are estimated using a number of sets of known features, called ground control points (GCPs), in both the reference image and the image to register [6]–[8]. The mapping functions may be a combination of orthogonal polynomials or sometimes higher order polynomials depending on system complexity [9]. This method is suitable for both image-to-image and image-to-map (without brightness values) registration. However, a large number of GCPs are required to avoid overfitting problems during the mapping function estimation, and the collection of GCPs is manually intensive and subject to human error. Fully automatic image registration procedures are preferable as they do not require manual selection of GCPs. Automatic image registration can be classified into two categories—feature-based image registration and intensity-based image registration. While feature-based image registration methods, such as Scale-invariant feature transform (SIFT) [10] and Speeded up robust features (SURF) [11], require automatic reliable identification of image features, such as points, corners, lines, contours, etc., intensity-based image registration makes use of the intensity relationship between the two whole images and has been investigated widely. Critical issues in this approach are the selection of the geometric mapping function and similarity measure. The choice of geometric mapping function will affect the dimensionality of the parameter search space for image registration depending on the transformation model. To cope with significant amounts of local distortion, local transformation models [12]–[15] with a large number of parameters may also be required. The simplest similarity measure that can be used is the sum of the squared intensity difference (SSD) between the two images [16]–[19]. However, when data are collected under different conditions, for example,

on different days, it is not reasonable to expect the two images to have the same brightness values due to the variations in atmospheric effects and other system noise. Correlation is a better similarity measure to use when registering images from the same sensor or a similar sensor when the wavebands are about the same [20], [21]. The limitation of this measure is that it is not suitable for registering multimodal images, such as between optical images and SAR images. This is because a pixel showing high reflectance in an optical image may not necessarily show high backscattering in a radar image, and vice versa.

Viola [22] and Collignon [23] introduced the mutual information (MI) similarity measure to overcome the problems associated with multimodal medical image registration. This measure does not require information about the surface properties of the objects, apart from their shape, and is robust with respect to variations in illumination which makes it a popular choice for multimodal image registration or image matching [13], [14], [24]–[27]. Cole-Rhodes [28], [29] and Chen [30], [31] extended the application of MI to multimodal remote sensing image registration. A good discussion on aspects of MI-based image registration is provided by Cole-Rhodes [19], [32] and Pluim [25].

To cover a wider range of transformation parameters and to speed up the registration process, a multiresolution approach, which is also known as the image pyramid approach or a wavelet-based hierarchical technique, is commonly applied during the optimization of MI [29], [33]–[36]. Commonly used search algorithms for similarity measure optimization include: a simple random search [37], genetic algorithms [38], simulated annealing [39], [40], the downhill simplex method [41], [42], Powell's conjugate direction search [23], [43], [44] and stochastic gradient maximization [22], [29].

Recently, a new similarity measure known as cross-cumulative residual entropy (CCRE) has been proposed by Wang [45], [46] for multimodal image registration in medical imaging applications. Wang [45], [47] developed a measure of the information in a random variable based on its cumulative distribution called the cumulative residual entropy (CRE) that parallels the well-known Shannon entropy. Since its definition is based on cumulative distribution functions (CDFs) rather than probability density functions and CDFs are more regular, this measure is more robust in the presence of noise. Based on the CRE, the CCRE between two random variables is defined, and it was applied to solve the image alignment problem for parameterized transformations. The key strengths of the CCRE-based registration scheme proposed by Wang [46] are: 1) it can accommodate images to be registered of varying contrast and brightness, and it is also robust in the presence of noise; 2) It was empirically shown by Wang [46] to converge faster in comparison to other registration methods that use information theory-based cost functions; 3) The cost function and its derivative share common terms, and this leads to computational savings being accrued in the numerical optimization process; 4) It is well suited for situations where the source and the target images have fields of view with large nonoverlapping regions which is quite common in practice.

Since the methods discussed above have been shown to work well for multimodal medical images, we have introduced their use with multimodal satellite images. However, the impact of the different types of noise and distortion found in remote sensing images on these new approaches is unknown. The investigation and evaluation of CCRE's performance on satellite images are one of the key contributions of this paper, and we will show that using CCRE as a similarity measure does in fact provide more robust registration performance with typical satellite imagery. We also propose a novel extension to the Parzen-window optimization approach proposed by Thévenaz [35] which involves applying partial volume interpolation in the calculation of the gradients of the similarity measure. In previous implementations which use partial volume interpolation for calculating MI [33], [44], [46], [48]–[50], it was used only in the construction of the joint histogram of the two images. In this paper, we apply partial volume interpolation in the calculation of the gradients of the similarity measure directly which enables us to implement an optimization procedure based directly on partial volume interpolation. Our experimental results show that our proposed approach which uses CCRE as the similarity measure and partial volume interpolation in the optimization procedure provides superior performance to other approaches investigated.

The remainder of the paper is organized as follows: Section II provides an overview of gradient-based multimodal image registration. In this section, different similarity measures—squared-difference (SD), squared gradient difference (SGD), MI, and CCRE—are discussed. Newton's optimization method is also summarized with attention on the required gradient and Hessian. Our proposed approach for using partial volume interpolation in the optimization of CCRE is explained in the last part of Section II. In Section III, experimental comparisons of several image registration methods applied to multimodal image pairs are provided. Finally, Section IV contains a summary of the performance of our proposed approach and our conclusions.

II. Multimodal Image Registration

Let (x', y') and (x, y) denote the coordinates of the pixels in the target image, T, (to be registered) and the reference image, R, respectively. Image registration is the process of finding the mapping functions

$$\begin{aligned} x' &= f(x, y) \\ y' &= g(x, y) \end{aligned} \tag{1}$$

which spatially align the target image with the reference image at all pixel locations. Let $t(x', y')$ and $r(x, y)$ denote the intensity values of the pixels located at (x', y') and (x, y) in images T and R, respectively. Automatic image registration is possible by making use of the intensity relationship between the pixels at corresponding locations in the two images. Let D denote the overall dissimilarity between the two images:

$$D = -S\{T, R\} = -\sum_{x,y} S\left\{t(x', y'), r(x, y)\right\} \tag{2}$$

where $S\{\ \}$ is a similarity measure. The mapping functions (in (1)) can be estimated via minimizing D as follows:

$$\min\{D\} = \min\left\{ -\sum_{x,y} S\left\{ t\left(f(x,y), g(x,y) \right), r(x,y) \right\} \right\}. \tag{3}$$

There are three issues involved in this process: 1) how to model the relationship between (x', y') and (x, y), 2) which similarity measure to use, and 3) how to implement the optimization. These three issues are discussed in the following sections.

A. Transformation Models

Several mapping functions, based on rigid, affine, projective, and nonlinear models [51], are available for selection depending on the assumed relationship between (x', y') and (x, y). A general approach is to adopt polynomials of first, second, or third degree to model the distortions contained in the target image when the explicit form of the geometric distortion is unknown [5]. The equations for second degree polynomials are given below:

$$\begin{aligned} x' &= a_0 + a_1 x + a_2 y + a_3 xy + a_4 x^2 + a_5 y^2 \\ y' &= b_0 + b_1 y + b_2 x + b_3 yx + b_4 y^2 + b_5 x^2. \end{aligned} \tag{4}$$

The first three terms in each of these equations allow shifting, scaling, shear, and rotation to be modeled. This corresponds to a pair of linear mapping functions and is also called an affine transformation model. Hence, we can rewrite (4) to model an affine transformation as follows:

$$\begin{aligned} x' &= a_0 + a_1 x + a_2 y \\ y' &= b_0 + b_1 y + b_2 x. \end{aligned} \tag{5}$$

Given that the earth's surface is usually a very large distance from the camera sensor, this affine model can be used as a good approximation of the actual geometric distortion. While higher order polynomials can model more complicated relationship between (x', y') and (x, y), the affine model has been adopted widely in practice to keep the operation manageable. Hence, we also adopt an affine model in this study to evaluate our proposed automatic registration approach.

B. Similarity Measures

Similarity measure selection is a critical step in automatic image registration. The simplest measure is the squared difference in brightness given by

$$S_{SD}\left\{ t(x', y'), r(x,y) \right\} = -\left(t(x', y') - r(x,y) \right)^2. \tag{6}$$

When the difference is small, the similarity is high. Therefore, a negative sign is used in (6). When two images are collected on different days, their brightness values may not be the same due to changes in the imaging conditions. However, they can be expected to be highly correlated. α and β are the linear regression parameters to fit the relationship between the two data sets. The similarity between the target image and the transformed reference image (via linear mapping) can be expressed as a squared intensity difference (SID)

$$S_{SID}\left\{ t(x', y'), r(x,y) \right\} = \left(t(x', y') - [\alpha \times r(x,y) + \beta] \right)^2. \tag{7}$$

For multimodal image registration, the SGD can be used:

$$\begin{aligned} &S_{SGD}\left\{ t(x', y'), r(x,y) \right\} \\ &\qquad = -\left(\sqrt{\Delta t_x^2 + \Delta t_y^2} - \sqrt{\Delta r_x^2 + \Delta r_y^2} \right)^2 \end{aligned} \tag{8}$$

where

$$\begin{aligned} \Delta t_x &= t(x', y') - t(x'+1, y') \\ \Delta t_y &= t(x', y') - t(x', y'+1) \\ \Delta r_x &= r(x,y) - r(x+1, y) \\ \Delta r_y &= r(x,y) - r(x, y+1). \end{aligned} \tag{9}$$

In this approach, the absolute gradient responses of two images are calculated individually. The SGD between the two images is determined by calculating the SSD between the two absolute gradient response images. Hence, the SGD measure is not directly dependent on intensity values of image pixels but instead is dependent on the edges in the images. As edges are commonly preserved to some extent for most multimodal registration applications, SGD is a better choice of similarity measure than the SD and SID measure.

An alternative similarity measure using MI for multimodal applications was first introduced by Viola [52] and Collignon [23] with later modifications provided by Studholme [53]. The MI similarity measure is defined as follows:

$$S_{MI}\{T, R\} = H(T) + H(R) - H(T, R) \tag{10}$$

where $H(T)$ and $H(R)$ are the Shannon entropy of images T and R, respectively, and $H(T, R)$ is the joint entropy for images T and R. Alternatively, MI is expressed in terms of probability distributions as [23], [53]

$$\begin{aligned} S_{MI}\{T, R\} = \sum_{u \in L_T} \sum_{v \in L_R} \Bigg[& P(t = u, r = v) \\ & \times \log\left\{ \frac{P(t = u, r = v)}{P_T(t = u) P_R(r = v)} \right\} \Bigg] \end{aligned} \tag{11}$$

where $P(t = u, r = v)$ is the joint probability that pixels in image T have intensity value of u and pixels at the same location in image R have an intensity value of v, P_T and P_R are the marginal probability distribution functions of the intensity values in images T and R, respectively, and L_T and L_R denote the number of discrete intensity values in images T and R, respectively.

This measure does not require information about the surface properties of the object and is robust with respect to variations of illumination. As a result the algorithms are quite general and may be used in a wide variety of imaging situations.

C. Cross-Cumulative Residual Entropy

Wang [45], [47] and Rao [54] used the cumulative distribution of a random variable to develop a novel measure of information that parallels Shannon entropy, called CRE. The key features of CRE are summarized as follows: 1) its definition is valid in both the continuous and discrete domains, 2) it is mathematically more general than the Shannon entropy, and 3) its computation from sample data is easy, and these computations converge asymptotically to the true values. The CRE of a random variable X with values denoted by x is defined in terms of the cumulative residual distribution $P(x > u)$ and is given by [45]

$$\varepsilon(X) = -\sum_{u} P(x > u) \log \left(P(x > u)\right). \tag{12}$$

The cumulative residual distribution function $P(x > u)$ describes the probability that the values of X will be greater than u. Using the definition provided by Wang [46], the CCRE between the intensity values in images T and R is given by

$$S_{CCRE}\{T, R\} = \varepsilon(T) - E\left[\varepsilon(T|R)\right] \tag{13}$$

where $\varepsilon(T|R)$ is the conditional CRE and E[ ] is the expectation operator. Following the definition provided by Wang [46], CCRE is expressed in terms of cumulative and marginal distributions as follows:

$$S_{CCRE}\{T, R\} = \sum_{u \in L_T} \sum_{v \in L_R} \left[Pt(t > u, r = v) \times \log \left\{ \frac{P(t > u, r = v)}{P_T(t > u) P_R(r = v)} \right\} \right] \tag{14}$$

where $P(t > u, r = v)$ is the joint probability that pixels in image T have intensity values greater than u and pixels at the same location in image R have an intensity value of v and $P_T(t > u)$ is the cumulative residual distribution of target image T.

Wang [46] showed that using CCRE provided more robust registration performance than MI for multimodal medical image alignment applications. The key strengths of the CCRE over using the popular MI method (based on Shannon's entropy) are that the former has significantly larger noise immunity and a much larger convergence range over the field of parameterized transformations. This new measure is a *fundamental departure* from all the existing measures of entropy in that it is based on the probability distribution of a random variable rather than its density function.

D. Optimization

The value of the dissimilarity measure for two images is expected to be a minimum when they are perfectly aligned. The accuracy and computation time required for dissimilarity-measure-based automatic registration procedures mainly depends on the choice of the search method. The main techniques used to find the transformation parameters that minimize the dissimilarity measure include: exhaustive search, heuristic search, and gradient-based search. An exhaustive search can provide parameters to obtain the exact minimum value, but it requires evaluating the dissimilarity measure for every possible combination of parameters which is impractical for image registration. Heuristic searches are experience-based techniques for solving problems. Heuristic searches are used to accelerate the process of finding a satisfactory solution, where an exhaustive search is not feasible. Gradient-based search algorithms require direct or approximate measurement of the gradient of the similarity measures. As repeated evaluation of the dissimilarity measures between the images is computationally expensive (even for heuristic search algorithms), gradient-based search algorithms (also known as gradient-descent optimization algorithms) are preferred over heuristic search algorithms when the optimum position solution is not far away from the current position of the image and the gradients are computationally feasible for a practical registration process. Several gradient-based algorithms are used for dissimilarity measure optimization, such as the Quasi-Newton method [46], Newton's method [16], and the Levenberg–Marquardt method [17], [35]. Newton's method was selected for use in the optimization procedure, since it outperforms the Quasi-Newton method and is simpler to implement than the Levenberg–Marquardt approach.

For convenience, we introduce the notation for a motion vector m, where

$$m = [a_0 a_1 a_2 \ldots b_0 b_1 b_2 \ldots]^T \tag{15}$$

and $[\,]^T$ represents the transpose operation. The elements of m are denoted by m_j for the jth element of the vector for notational consistency—e.g., in the case of an affine transformation model

$$\begin{aligned} m &= [m_1, m_2, m_3, m_4, m_5, m_6]^T \\ &= [a_0, a_1, a_2, b_0, b_1, b_2]^T \end{aligned} \tag{16}$$

where m_1 and m_4 correspond to translation, m_2 and m_5 correspond to scaling in each direction, and m_2, m_3, m_5, and m_6 together model rotation and shear.

During the registration process, these transformation parameters change to bring the images closer into alignment, so, for an affine transformation model, (4) can be rewritten as a function of the transformation parameters m in addition to position values x and y as follows:

$$\begin{aligned} x' &= f(x, y, m) = m_1 + m_2 x + m_3 y \\ y' &= g(x, y, m) = m_4 + m_5 y + m_6 x. \end{aligned} \tag{17}$$

The motion vector determines the relationship between (x', y') and (x, y) when the transform mapping function has the form of (4) and can be estimated by finding the values which minimize the dissimilarity between the two images as follows:

$$\begin{aligned} &\min\{D(m)\} \\ &= \min\left\{ -\sum_{x,y} S\left\{t(x', y'), r(x, y)\right\} \right\} \\ &= \min\left\{ -\sum_{x,y} S\left\{t\left(f(x, y, m), g(x, y, m)\right), r(x, y)\right\} \right\} \end{aligned} \tag{18}$$

where $D(m)$ is the dissimilarity measure expressed as a function of the motion parameters. Newton's method is a well-known algorithm to find the local minima of a function iteratively when the minimum is close enough. A major prerequisite of this method is that the target function must have computable gradients and Hessian. As described in [41], the geometric interpretation of Newton's method is that, at each iteration, $D(m)$ is approximated by a quadratic function around the vector of parameters m and then takes a step vector p toward the minimum of that quadratic function. This method uses a second-order Taylor series approximation as follows:

$$D(m+p) = D(m) + p^T \nabla D(m) + \frac{1}{2} p^T \nabla^2 D(m) p \quad (19)$$

where p^T represents the transpose of vector p. In the dissimilarity measure optimization procedure, the motion parameter vector m is updated at iteration number n for use in the next iteration $n+1$ using the following equation:

$$m^{n+1} = m^n - \left(\nabla^2 D(m^n)\right)^{-1} \nabla D(m^n) \quad (20)$$

where m^n represents the m vector at the nth iteration of Newton's optimization method. The optimization method ends when there is negligible change between m^n and m^{n+1}.

1) Optimization of Mutual Information: As MI in its original form is not continuous and differentiable, Thévenaz [35] applied Parzen-window estimates of the probability density function and computed tractable closed-form expression for the gradient and Hessian of MI. MI as defined in (11) can be expressed using Parzen-window estimates of the joint and marginal probability [35], [55] to transform the function in a differentiable form as follows:

$$\begin{aligned} S_{MI}(m) &= S_{MI}\{T, R\} \\ &= \sum_{u=1}^{L_T} \sum_{v=1}^{L_R} \hat{P}(u,v) \log\left(\frac{\hat{P}(u,v)}{\hat{P}_T(u)\hat{P}_R(v)}\right) \end{aligned} \quad (21)$$

where $P(t=u, r=v)$, $P_R(r=v)$, and $P_T(t=u)$ are replaced by $\hat{P}(u,v)$, $\hat{P}_R(v)$, and $\hat{P}_T(u)$ which are Parzen-window estimated distributions instead of the original distributions. These Parzen-window estimates of the joint and marginal probability distribution functions as proposed by Thévenaz [35] are given by

$$\begin{aligned} \hat{P}(u,v) &= \frac{1}{N} \sum_{i=1}^{N} \beta^3\left(u - t\left(x_i', y_i'\right)\right) \beta^3\left(v - r(x_i, y_i)\right) \\ \hat{P}_T(u) &= \sum_{v=1}^{L_R} \hat{P}(u,v) = \frac{1}{N} \sum_{i=1}^{N} \beta^3\left(u - t\left(x_i', y_i'\right)\right) \\ \hat{P}_R(v) &= \sum_{u=1}^{L_T} \hat{P}(u,v) = \frac{1}{N} \sum_{i=1}^{N} \beta^3\left(v - r(x_i, y_i)\right) \end{aligned} \quad (22)$$

where N is the total number of pixels in the overlapping area of the two images, $r(x_i, y_i)$ and $t(x_i', y_i')$ are the intensity values of the ith pixel in the overlapping area of images R and T in their respective co-ordinate systems. These estimates are defined using the cubic B-spline function $\beta^3(\ )$ which is separable, has the unit integral property, and is defined as follows:

$$\beta^3(x) = \begin{cases} \left(4 - 6|x|^2 + 3|x|^3\right)/6 & 0 \le |x| < 1 \\ \left(2 - |x|\right)^2/6 & 1 \le |x| < 2 \\ 0 & 2 \le |x| \end{cases} \quad (23)$$

The gradient vector elements and Hessian matrix elements of the MI are calculated as follows [35], [55]:

$$\begin{aligned} \frac{\partial S_{MI}(m)}{\partial m_j} &= \sum_{u=1}^{L_T} \sum_{v=1}^{L_R} \frac{\partial \hat{P}(u,v)}{\partial m_j} \log\left(\frac{\hat{P}(u,v)}{\hat{P}_T(u)\hat{P}_R(v)}\right) \\ \frac{\partial^2 S_{MI}(m)}{\partial m_j \partial m_k} &= -\sum_{u=1}^{L_T} \frac{\partial \hat{P}_T(u)}{\partial m_j} \frac{\partial \hat{P}_T(u)}{\partial m_k} \frac{1}{\hat{P}_T(u)} \\ &\quad + \sum_{u=1}^{L_T} \sum_{v=1}^{L_R} \frac{\partial \hat{P}(u,v)}{\partial m_j} \frac{\partial \hat{P}(u,v)}{\partial m_k} \frac{1}{\hat{P}(u,v)}. \end{aligned} \quad (24)$$

To evaluate these elements, we need to calculate $\partial \hat{P}(u,v)/\partial m_j$ and $\partial \hat{P}_T(u)/\partial m_j$ which can be obtained by partial differentiation of (22) with respect to m_j as follows:

$$\begin{aligned} \frac{\partial \hat{P}(u,v)}{\partial m_j} &= -\frac{1}{N} \sum_{i=1}^{N} \left\{ \beta^3\left(v - r(x_i, y_i)\right) \right. \\ &\quad \left. \times \frac{\partial \beta^3\left(u - t\left(x_i', y_i'\right)\right)}{\partial\left(u - t\left(x_i', y_i'\right)\right)} \frac{\partial t\left(x_i', y_i'\right)}{\partial m_j} \right\} \\ \frac{\partial \hat{P}_T(u)}{\partial m_j} &= -\frac{1}{N} \sum_{i=1}^{N} \frac{\partial \beta^3\left(u - t\left(x_i', y_i'\right)\right)}{\partial\left(u - t\left(x_i', y_i'\right)\right)} \frac{\partial t\left(x_i', y_i'\right)}{\partial m_j}. \end{aligned} \quad (25)$$

Here, the partial derivative of the image T with respect to the motion parameters m_j can be calculated as

$$\frac{\partial t\left(x_i', y_i'\right)}{\partial m_j} = \frac{\partial t\left(x_i', y_i'\right)}{\partial x_i'} \frac{\partial x_i'}{\partial m_j} + \frac{\partial t\left(x_i', y_i'\right)}{\partial y_i'} \frac{\partial y_i'}{\partial m_j} \quad (26)$$

where $\partial t(x_i', y_i')/\partial x_i'$ and $\partial t(x_i', y_i')/\partial y_i'$ are the simple gradients of the image in the horizontal and vertical directions of the ith pixel of the image T. $\partial x_i'/\partial m_j$ and $\partial y_i'/\partial m_j$ can be calculated using (4) or for an affine model using (5) and (17).

2) Optimization of Cross-Cumulative Residual Entropy: The CCRE defined in (14) can also be rewritten using Parzen-window estimates of the joint and marginal cumulative residual distributions as a function of motion parameters m as follows [46]:

$$\begin{aligned} S_{CCRE}(m) &= S_{CCRE}\{T, R\} \\ &= \sum_{u=1}^{L_T} \sum_{v=1}^{L_R} \hat{G}(u,v) \log\left(\frac{\hat{G}(u,v)}{\hat{G}_T(u)\hat{P}_R(v)}\right) \end{aligned} \quad (27)$$

where $P(t>u, r=v)$ and $P_T(t>u)$ are replaced by $\hat{G}(u,v)$ and $\hat{G}_T(u)$ which are Parzen-window estimated distributions

instead of the original distributions. These Parzen-window estimates are given by

$$\hat{G}(u,v) = \frac{1}{N}\sum_{i=1}^{N} \phi\left(u - t\left(x_i', y_i'\right)\right) \beta^3\left(v - r(x_i, y_i)\right)$$

$$\hat{G_T}(u) = \sum_{v=1}^{L_R} \hat{G}(u,v) = \frac{1}{N}\sum_{i=1}^{N} \phi\left(u - t\left(x_i', y_i'\right)\right) \quad (28)$$

where N is the total number of pixels in the overlapping area of the two images, $r(x_i, y_i)$ and $t(x_i', y_i')$ are the ith pixel in the overlapping area of images R and T in their respective coordinate systems, $\beta^3()$ is the cubic B-spline function, and $\phi(\ )$ is the cumulative residual of the cubic B-spline function which is defined as [55]

$$\phi(u) = \int_u^{\infty} \beta^3(\alpha)\, d\alpha. \quad (29)$$

The gradient vector elements and Hessian matrix elements of the CCRE are then calculated as follows [46], [55]:

$$\frac{\partial S_{CCRE}(m)}{\partial m_j} = \sum_{u=1}^{L_T}\sum_{v=1}^{L_R} \frac{\partial \hat{G}(u,v)}{\partial m_j} \log\left(\frac{\hat{G}(u,v)}{\hat{G}_T(u)\hat{P}_R(v)}\right)$$

$$\frac{\partial^2 S_{CCRE}(m)}{\partial m_j \partial m_k} = -\sum_{u=1}^{L_T} \frac{\partial \hat{G}_T(u)}{\partial m_j}\frac{\partial \hat{G}_T(u)}{\partial m_k}\frac{1}{\hat{G}_T(u)} + \sum_{u=1}^{L_T}\sum_{v=1}^{L_R} \frac{\partial \hat{G}(u,v)}{\partial m_j}\frac{\partial \hat{G}(u,v)}{\partial m_k}\frac{1}{\hat{G}(u,v)}. \quad (30)$$

To evaluate these elements, we need to calculate $\partial\hat{G}(u,v)/\partial m_j$ and $\partial\hat{G}_T(u)/\partial m_j$ which can be obtained by partial differentiation of (28) with respect to m_j as follows:

$$\frac{\partial \hat{G}(u,v)}{\partial m_j} = \frac{1}{N}\sum_{i=1}^{N}\left\{\beta^3\left(u - t\left(x_i', y_i'\right)\right) \times \beta^3\left(v - r(x_i, y_i)\right)\frac{\partial t\left(x_i', y_i'\right)}{\partial m_j}\right\}$$

$$\frac{\partial \hat{G}_T(u)}{\partial m_j} = \frac{1}{N}\sum_{i=1}^{N}\beta^3\left(u - t\left(x_i', y_i'\right)\right)\frac{\partial t\left(x_i', y_i'\right)}{\partial m_j} \quad (31)$$

where the partial derivative of the image T with respect to motion parameters $\partial t(x_i', y_i')/\partial m_j$ can be calculated using (26).

3) Comparing the Optimization of CCRE and MI: The performance of gradient-based optimization is mostly dependent on the characteristics of the approximated gradients of the function to be minimized around the global minimum. As both CCRE and MI are similarity measures, we are using the negative of CCRE and MI as the function to be minimized. In Newton's optimization method, at each iteration, the gradient of the target function determines the direction to proceed toward the global minimum. For a simple demonstration, consider using Fig. 4(a) as reference and its artificially distorted version as the target. The distortion was introduced with only one transformation parameter which represents translation in the horizontal direction (x-axis). The optimization process reaches its goal when it finds the value of horizontal translation which produces a value of zero for the gradient. Fig. 1(a) shows the target image T translated in the horizontal direction only with respect to the reference image R which means the m vector contains only one element m_1. Fig. 1(b) shows the values of the gradients of the two similarity measures $\partial S_{CCRE}(m)/\partial m$ and $\partial S_{MI}(m)/\partial m$ with respect to the translation in the horizontal direction between T and R. As the gradient values near the x-axis are not clearly visible due to the presence of large values in Fig. 1(b), Fig. 1(c) shows the same curves with a smaller range of values on the y-axis.

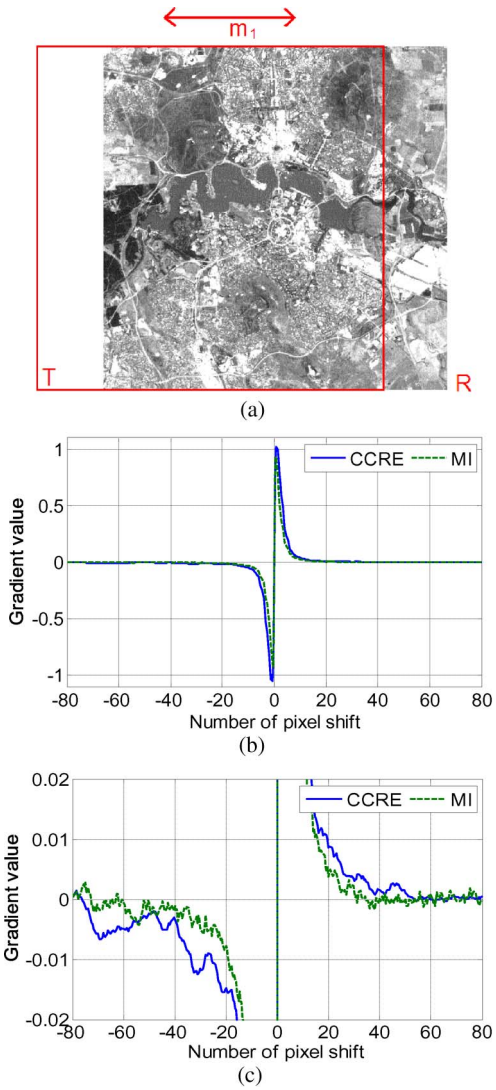


Fig. 1. (a) Target image (T) is translated in horizontal direction with respect to reference image (R). (b) Negative gradient curves of CCRE and MI with change in horizontal translation. (c) Same curves shown in (b) with smaller range of gradient values.

As both the target and reference images are the same, the correct registration result and global minima occurs at a pixel translation equal to zero $m_1 = 0$. As expected, both the gradients of CCRE and MI are seen crossing the x-axis from the negative to the positive direction representing a global minima at zero pixel translation, i.e., $m_1 = 0$. When the translation between the images is negative, a negative value of the

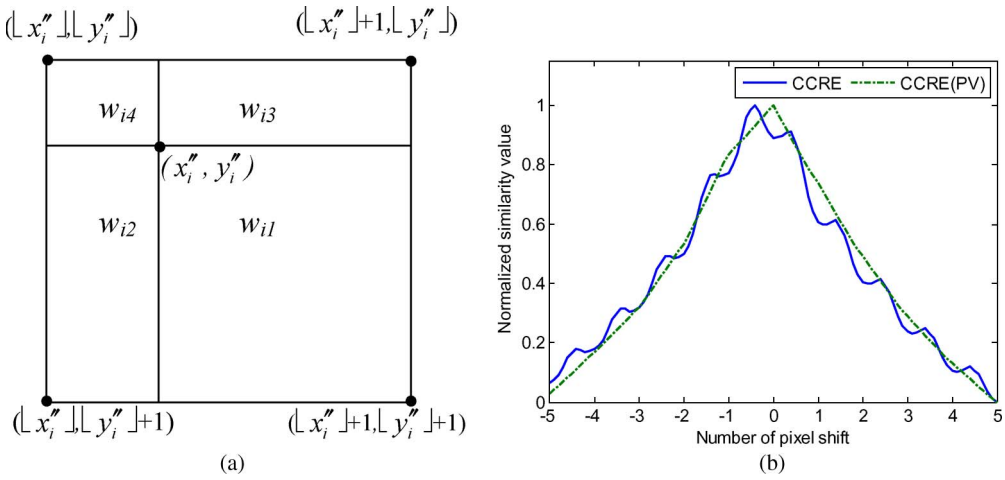


Fig. 2. (a) Partial volume interpolation weight calculation for four neighbor pixels. (b) The CCRE similarity measure as a function of horizontal translation for the images shown in Fig. 4 when using linear interpolation and partial volume interpolation.

gradient is required to move the images closer to alignment. Similarly, when the translation is positive, a positive value of the gradient is required to successfully register the images. Hence, successful registration of the two images will depend on the range of initial translations which produce a negative gradient for negative translations and a positive gradient for positive translations. For the CCRE similarity measure, initial translation between -74 pixels and $+52$ pixels will result in a successful registration. Values outside this range will produce an incorrect sign for the gradient and result in the algorithm converging to a local minimum of the similarity measure. For the MI similarity measure, the range of initial translations which produce a successful registration is between -63 pixels and $+32$ pixels which is much smaller than for CCRE. Hence, using CCRE as the similarity measure instead of MI allows a greater range of initial registration errors to be corrected.

E. Partial Volume Interpolation

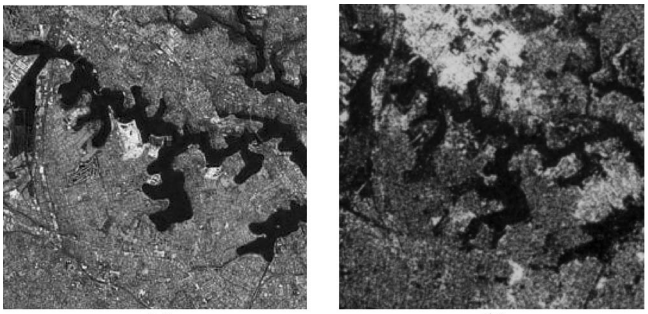
After geometrically transforming an image, the pixel locations in the transformed image will not usually coincide exactly with the pixel locations in the original image. Hence, interpolation is necessary to estimate the pixel values in the transformed image from neighboring pixel locations in the original image. Some common interpolation methods are: nearest neighbor, bilinear, and bicubic. The use of these common interpolation techniques results in spurious local minima of the dissimilarity measure when the transformation parameters are varied. Partial volume interpolation was introduced by Maes [44] and Collignon [33] to eliminate these spurious local minima caused by interpolation. This method does not create an interpolated image instead it builds the joint histogram for two images with the weighted contributions of neighboring pixels for each transformed pixel location. Similar to linear interpolation, partial volume interpolation calculates the weights of four neighboring pixels and, instead of calculating a weighted average of neighboring pixel values, it increments the joint histogram entries using the four neighboring pixel weights and greyscale values. In the different implementations of partial volume interpolation [33], [44], [46], [48]–[50], for calculating MI, it was used only in the construction of the joint histogram of two images. In this paper, we apply partial volume interpolation in the calculation of the gradients of the similarity measure directly which enables us to implement an optimization procedure based directly on partial volume interpolation.

1) Optimization of CCRE using Partial Volume Interpolation: In our proposed algorithm, we optimize CCRE using partial volume interpolation as follows: Consider a pixel at a location (x_i', y_i') in the target image T after applying a geometric transform to the original target image T_{org}. Let (x_i'', y_i'') be the corresponding location in the original target image and let its four neighboring pixels $t_{org}(\lfloor x_i'' \rfloor, \lfloor y_i'' \rfloor)$, $t_{org}(\lfloor x_i'' \rfloor + 1, \lfloor y_i'' \rfloor)$, $t_{org}(\lfloor x_i'' \rfloor, \lfloor y_i'' \rfloor + 1)$ and $t_{org}(\lfloor x_i'' \rfloor + 1, \lfloor y_i'' \rfloor + 1)$ as shown in Fig. 2(a) be denoted by $t_{org}(x_{i1}'', y_{i1}'')$, $t_{org}(x_{i2}'', y_{i2}'')$, $t_{org}(x_{i3}'', y_{i3}'')$ and $t_{org}(x_{i4}'', y_{i4}'')$. Their corresponding partial volume weights $w_{i1}, w_{i2}, w_{i3}, w_{i4}$ are calculated as

$$\begin{aligned} w_{i1} &= (\lfloor x_i'' \rfloor + 1 - x_i'')\,(\lfloor y_i'' \rfloor + 1 - y_i'') \\ w_{i2} &= (x_i'' - \lfloor x_i'' \rfloor)\,(\lfloor y_i'' \rfloor + 1 - y_i'') \\ w_{i3} &= (\lfloor x_i'' \rfloor + 1 - x_i'')\,(y_i'' - \lfloor y_i'' \rfloor) \\ w_{i4} &= (x_i'' - \lfloor x_i'' \rfloor)\,(y_i'' - \lfloor y_i'' \rfloor)\,. \end{aligned} \tag{32}$$

The Parzen-window estimates of the joint cumulative residual distribution, its gradient elements, and the gradient elements of the marginal cumulative residual distributions can be rewritten using partial volume interpolation as follows:

$$\begin{aligned} &\hat{G}(u,v) \\ &= \frac{1}{N}\sum_{a=1}^{4}\sum_{i=1}^{N}\Big\{ w_{ia}\phi\left(u - t_{org}\left(x_{ia}'', y_{ia}''\right)\right) \\ &\qquad \times \beta^3\left(v - r(x_i, y_i)\right)\Big\} \\ &\frac{\partial \hat{G}(u,v)}{\partial m_j} \\ &= \frac{1}{N}\sum_{a=1}^{4}\sum_{i=1}^{N}\Big\{ w_{ia}\beta^3\left(v - r(x_i, y_i)\right) \\ &\qquad \times \beta^3\left(u - t_{org}\left(x_{ia}'', y_{ia}''\right)\right)\frac{\partial t_{org}\left(x_{ia}'', y_{ia}''\right)}{\partial m_j}\Big\} \end{aligned}$$



(a) (b)

Fig. 3. Multimodal images used to evaluate the multimodal image registration algorithms. (a) The Google satellite image. (b) SAR image.

$$\frac{\partial \hat{G}_T(u)}{\partial m_j} = \frac{1}{N}\sum_{a=1}^{4}\sum_{i=1}^{N}\left\{ w_{ia}\beta^3\left(u - t_{org}\left(x''_{ia}, y''_{ia}\right)\right) \times \frac{\partial t_{org}\left(x''_{ia}, y''_{ia}\right)}{\partial m_j}\right\}. \quad (33)$$

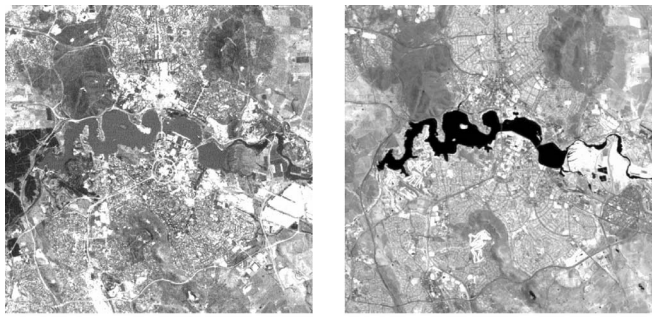
Fig. 2(b) shows a plot of the values obtained for the CCRE of the image pair shown in Fig. 4 when the target image is translated horizontally and linear interpolation is used to estimate values at subpixel locations of the original image. The curve in Fig. 2(b) shows that linear interpolation produces nonconvex behavior of the CCRE similarity measure which may cause the registration method to converge to a local maximum. Furthermore, due to interpolation artifacts, the maximum of the CCRE function values does not occur at zero translation which indicates that the registration method will align the images away from its correct position causing some inaccuracy. However, after applying partial volume interpolation as described above, the plot of the similarity measure as a function of translation shows no local maxima and reaches its global maximum at the correct translation value.

III. Experimental Results

A. Data set

Two pairs of multimodal image were used in this study. In the first pair of images, the site used is an area within Sydney, Australia. The two images used were the visible red channel $(\text{wavelength} = 0.65 \times 10^{-6}\ \text{m})$ of a Google satellite image captured in 2000 and an SIR-B SAR image of L band $(\text{wavelength} = 0.235\ \text{m})$ with an incidence angle of 16.0° captured in 1984. Both images have a size of 256×256 pixels with a spatial resolution of 30 m. These images are shown in Fig. 3.

The brightness of the two images is not correlated, and there are also differences due to the images being captured at different times. The SAR image provides information different from that of optical sensors which operate in the visible and infrared regions of the electromagnetic spectrum. SAR data consists of high-resolution backscattering of radar-frequency energy from terrain that has been illuminated by a directed beam of pulses generated by the sensor. There are large bright areas on the SAR image where the streets and houses were orientated in parallel to the satellite track direction and therefore, strong corner reflections were created. In the optical image, this effect is not present, and the reflectance of the solar spectrum from these areas was low. SSD-based image registration methods will not work for this case. As the original pair of images used in this test case were not of the same resolution, the interpolation method used to correct for the images having different initial resolutions will influence the registration result. We have chosen to use cubic interpolation as it provides the most robust performance.



(a) (b)

Fig. 4. Another set of images used to evaluate the multimodal registration algorithms. (a) Band 1 and (b) band 4 ETM+ image around Canberra.

TABLE I
List of Value Ranges for Affine Transformation Parameters

Parameter	Value range	Effects
m_2, m_5	0.9 to 1.1	up to ±10% image shrinking or stretching in the horizontal and vertical directions respectively.
m_3, m_6	-0.1 to 0.1	rotation and shearing when combined with m_2 and m_5.
m_1, m_4	band 1 to band 4: -40 to 40 SAR to optical: -26 to 26	up to ±10% translation in the horizontal and vertical directions respectively.

The other pair of images in the study shows an area of the center of Canberra, Australia. The images were 400×400 pixels in size with a spatial resolution of 30 m and are shown in Fig. 4. The images in Fig. 4(a) and (b) show the reflectance of band 1 (blue-green light) and band 4 (near infrared), respectively, from a set of Landsat Enhanced Thematic Mapper Plus (ETM+) data recorded in the year 2000. Band 1 and band 4 of this data set contain the reflectance for wavelengths in the range 450–520 nm and 760–900 nm, respectively. Although it is rare to register remote sensing images across the bands, in this study, band 1 and band 4 images are registered to demonstrate the ability of multimodal similarity measures to register images captured at different wavelengths. The content in the images will also influence the registration performance. However, a reasonable-sized remote sensing image normally contains adequate edges and spatial structures, which can be used in the proposed intensity-based registration method.

The two images were first manually registered by selecting a number of control points and estimating the geometric transform required to make these control points have the same location in both images. The manually registered images were then

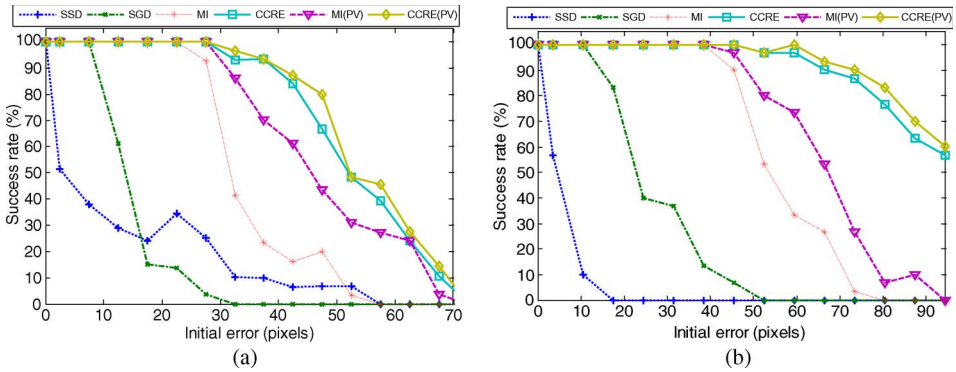


Fig. 5. Registration success rate comparison plot for different registration algorithms at different ranges of initial registration error for (a) the SAR and optical image data sets and (b) band 1 and band 4 of the Landsat ETM+ image.

used for verification in the testing of the automatic registration methods.

B. Experimental Procedure

To conduct a thorough testing of the performance of the CCRE measure when compared to other similarity measures, a set of 420 affine transformations were applied to the target image of each pair in Figs. 3 and 4. As all the intensity-based image registration methods can only work within a certain range of initial registration errors, we have simulated a range of test cases to examine each method's performance. The 420 sets of affine motion parameters, m_1 to m_6 were generated using a uniform distribution within the ranges given in Table I.

While the original two images have 256 brightness levels, they were quantized to 32 levels for two reasons. One reason is to save the computational load in the calculation of the probability distributions and cumulative distributions of the two images. The other is to avoid very sparse histograms with a low probability in each bin. The most expensive operation in the proposed procedure is the calculation of the histograms required to estimate the gradients of the similarity measure at each iterations. Reducing the number of brightness levels to 32 from the original 256 levels will save a significant computational load for these calculations. We found that the current Matlab implementation with 32 levels requires approximately 40 s run time for the image pair in Fig. 3 whereas the same implementation with 256 levels requires approximately 600 s run time. However, further reducing the number of levels to less than 32 meant that a sufficient description of the probability distribution could not be achieved.

The algorithms were applied for 130 iterations for each of the 420 artificial affine transformations on the aforementioned two data sets. The performance of the registration algorithms was evaluated by measuring the maximum registration error (MRE) at four corner points of the image, given by

$$\text{MRE} = \max_i \left(\sqrt{(x_i' - x_i)^2 + (y_i' - y_i)^2} \right) \qquad (34)$$

where x_i and y_i are the true locations on the reference image, and x_i' and y_i' are the corresponding locations on the target image for a given pixel in the scene. To evaluate the performance of the proposed method using CCRE with partial volume interpolation (CCRE(PV)), we have implemented and compared the proposed approach with algorithms using the following similarity measures: SSD, SGD, MI, MI, with partial volume interpolation (MI(PV)) and CCRE without partial volume interpolation (CCRE). The MI (PV) method implemented here is similar to the implementation proposed for CCRE(PV) earlier rather than the implementation of Maes [44] and Thévenaz [35]. In the method used in our experiments, partial volume interpolation is used in the calculation of all gradients and the Hessian.

TABLE II
TOTAL SUCCESS RATE OF THE AUTOMATIC IMAGE REGISTRATION ALGORITHM USING DIFFERENT SIMILARITY MEASURES

Similarity Measure	Figure 3 Dataset	Figure 4 Dataset
SSD	27.61 %	5.00 %
SGD	22.38 %	27.14 %
MI	50.00 %	57.61 %
CCRE	75.95 %	90.47 %
MI(PV)	67.85 %	67.61 %
CCRE(PV)	78.35 %	92.58 %

C. Results

Fig. 5 shows the success rate of these algorithms for different ranges of initial MRE. The registration method was considered to be successful if the MRE after the final iteration was less than 2 pixels. The algorithms using SSD and SGD as similarity measures had very poor performance in terms of the success rate of registration as they are not suitable for multimodal data. As MI- and CCRE-based registration algorithms are capable of handling multimodal data with nonlinear relationships they produced a much better registration success rate.

Table II shows the total success rate of the proposed automatic image registration algorithm using different similarity measures for the aforementioned two-image data sets. The total success rate was defined as the percentage of all 420 initial affine transformations that were successfully registered by each registration method. The total success rate when using CCRE is 25.95% and 32.86% higher for the two data sets, respectively, when compared to using MI as the similarity measure. Introducing PV interpolation according to our proposed approach increases the success rate by 17.85% and 10.00% when using MI as the similarity measure and also increases the success

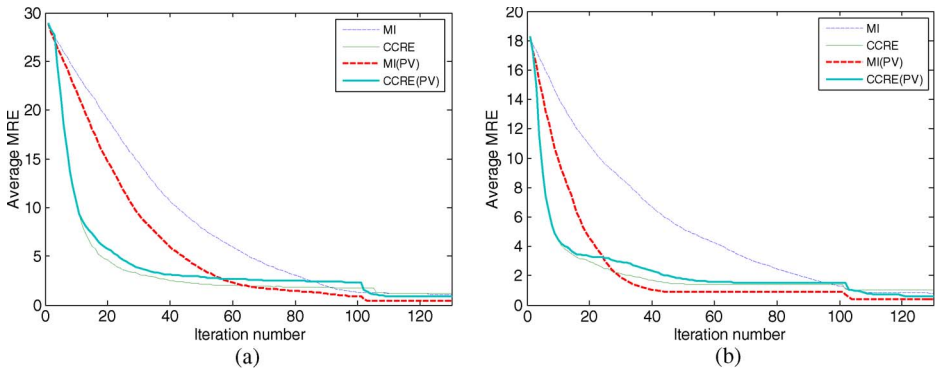


Fig. 6. Average MRE at each iteration of different registration algorithms (a) for SAR and optical image data set over successful cases (b) for band 1 and band 4 Landsat ETM+ image over successful cases.

rate by 2.40% and 2.11% when using CCRE on the two sets of data, respectively. These results show that our proposed approach using partial volume interpolation effectively removes the distortion created by interpolation artifacts for both MI and CCRE similarity measures. However, the improvement is more for MI when compared to CCRE because CCRE is already more robust.

The results plotted in Fig. 5 show that the success rate of the multimodal algorithms are approximately equal for small initial disparities but for larger disparities the CCRE(PV)-based algorithm has a significantly greater chance of successfully converging to the correct set of transform parameters. The CCRE(PV) provides a success rate of 100% for initial MRE of up to 28 pixels for the first data set and 45 pixels for the second data set.

The average MRE evaluated at each iteration over all the successful test cases for all the registration methods having more than 50% success rate was calculated and plotted in Fig. 6. In the unsuccessful test cases, the registration algorithm might align the images into any arbitrary position which could produce very large MRE values, so these cases are discarded while calculating the average MRE. To make a fair comparison of convergence rate and accuracy among these methods, we need to consider only the cases when all these methods are successful, i.e., all methods are starting from the same average initial registration error. However, the success rate of the SSD- and SGD-based methods for the image set in Fig. 3 and the image set in Fig. 4 are too small to make a good comparison, so we only considered the successful cases for the other four methods. The average MRE value for these measures at each iteration is plotted for the 210 successful cases for the image pair in Fig. 3 and the 241 successful cases for the image pair in Fig. 4. As shown in Fig. 6(a) and (b), the MI-based image registration method is the slowest to reach the solution whereas CCRE(PV) and CCRE converge to the solution very quickly—they converge to within 4 pixels error after only 25 iterations on average, which implies that using CCRE and CCRE(PV) will reduce the computational time by at least 30% by effectively finding the solution in a fewer number of iterations. The final average registration errors of the MI, CCRE, MI(PV), and CCRE(PV)-based algorithms are 0.90, 1.03, 0.41, 0.56 pixels and 0.86, 0.96, 0.30, 0.65 pixels as shown by Fig. 6(a) and (b). We can see from this graph that the CCRE(PV)- and MI(PV)-based algorithms' final registration errors are the smallest and within sub pixel accuracy. MI(PV) provides slightly better performance in terms of the final registration accuracy (by 0.15 and 0.35 pixels, respectively). This is because the CCRE measure is less affected by large image gradients causing it to be more robust but slightly less accurate than MI [46].

However, the main result we are investigating is the robustness of each of the automatic registration techniques to large initial distortion errors along with pixel intensity variations between the images. Fig. 5 clearly shows that for the reasons we have explained CCRE(PV) is much more robust when compared to MI(PV) in this respect. Furthermore, the results in Figs. 5 and 6 show that applying partial volume interpolation during the optimization procedure greatly improves the success rate, convergence, and accuracy of both the MI- and CCRE-based registration algorithms.

IV. Conclusion

The most important criteria to consider when comparing registration algorithms are success rate and registration accuracy. Several automatic multimodal registration algorithms using different similarity measures for an affine transformation model were investigated in this paper. The experimental results show that our proposed approach of using CCRE as a similarity measure and adopting partial volume interpolation when calculating the gradients for the optimization procedure provided a significantly higher success rate than the other algorithms tested. The results also showed that using partial volume interpolation in the optimization process significantly improved the performance of CCRE and MI-based algorithms both in terms of registration success rate and accuracy.

References

[1] L. Paolini, F. Grings, J. Sobrino, J. Jiménez Muñoz, and H. Karszenbaum, "Radiometric correction effects in landsat multi-date/multi-sensor change detection studies," *Int. J. Remote Sens.*, vol. 27, no. 4, pp. 685–704, Feb. 2006.

[2] J.-Y. Rau, L.-C. Chen, J.-K. Liu, and T.-H. Wu, "Dynamics monitoring and disaster assessment for watershed management using time-series satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 6, pp. 1641–1649, Jun. 2007.

[3] L. A. Klein, "Sensor and data fusion: A tool for information assessment and decision making," in *Proc. SPIE*, 2004, pp. 8–10.

[4] I. R. Farah, W. Boulila, K. S. Ettabaâ, B. Solaiman, and M. B. Ahmed, "Interpretation of multisensor remote sensing images: Multiapproach fusion of uncertain information," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 12, pp. 4142–4152, Dec. 2008.

[5] J. A. Richards and X. Jia, *Remote Sensing Digital Image Analysis.* 4th ed. Berlin, Germany: Springer-Verlag, 2006, pp. 56–58.

[6] X. Jia, "Automatic ground control points refinement for remote sensing imagery registration," in *Proc. 2nd Int. Conf. Intell. Sensors, Sensor Netwo. Inf. Process.*, Melbourne, Australia, 2005, pp. 145–149.

[7] X. Jia, "On error correction and accuracy assessment of satellite imagery registration," *J. Aust. Map Circle Inc.*, no. 54, pp. 35–38, Jan. 2003, The Globe, No. 54.

[8] Z. Wangfei, G. Jianguo, X. Tianshu, and H. Yanru, "The selection of ground control points in a remote sensing image correction based on weighted voronoi diagram," in *Proc. Int. Conf. Inf. Technol. Comput. Sci.*, Kiev, Ukraine, 2009, vol. 2, pp. 326–329.

[9] A. J. De Leeuw, L. M. M. Veugen, and H. T. C. Van Stokkom, "Geometric correction of remotely-sensed imagery using ground control points and orthogonal polynomials," *Int. J. Remote Sens.*, vol. 9, no. 10/11, pp. 1751–1759, 1988.

[10] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, Nov. 2004.

[11] H. Bay, A. Ess, T. Tuytelaars, and L. Van Gool, "Surf: Speeded up robust features," *Comput. Vis. Image Understand.*, vol. 110, no. 3, pp. 346–359, 2008.

[12] D. Liu, P. Gong, M. Kelly, and Q. Guo, "Automatic registration of airborne images with complex local distortion," *Photogramm. Eng. Remote Sens.*, vol. 72, no. 9, pp. 1049–1059, Sep. 2006.

[13] S. Suri and P. Reinartz, "Mutual-information-based registration of TerraSAR-X and Ikonos imagery in urban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 2, pp. 939–949, Feb. 2010.

[14] V. Arévalo and J. González, "Improving piecewise linear registration of high-resolution satellite images through mesh optimization," *IEEE Trans. Geosci. Remote Sens.*, vol. 46, no. 11, pp. 3792–3803, Nov. 2008.

[15] A. Borzi, M. di Bisceglie, C. Galdi, and G. Giangregorio, "Robust registration of satellite images with local distortions," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Cape Town, South Africa, Jul. 2009, vol. 3, pp. 251–254.

[16] S. Baker and I. Matthews, "Lucas-Kanade 20 years on: A unifying framework," *Int. J. Comput. Vis.*, vol. 56, no. 3, pp. 221–255, 2004.

[17] J. Nocedal and S. Wright, *Numerical Optimization*. New York: Springer-Verlag, 1999, pp. 262–266.

[18] S. AbdelSayed, D. Ionescu, and D. Goodenough, "Matching and registration method for remote sensing images," in *Proc. Int. Geosci. Remote Sens. Symp.*, Firenze, Italy, 1995, vol. 2, pp. 1029–1031.

[19] A. A. Cole-Rhodes and R. D. Eastman, "Gradient descent approaches to image registration," in *Image Registration for Remote Sensing*, J. L. Moigne, N. S. Netanyahu, and R. D. Eastman, Eds. Cambridge: Cambridge Univ. Press, 2011.

[20] W. K. Pratt, "Correlation techniques of image registration," *IEEE Trans. Aerosp. Electron. Syst.*, vol. AES-10, no. 3, pp. 353–358, May 1974.

[21] J. Kim and J. A. Fessler, "Intensity-based image registration using robust correlation coefficients," *IEEE Trans. Med. Imag.*, vol. 23, no. 11, pp. 1430–1444, Nov. 2004.

[22] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," *Int. J. Comput. Vis.*, vol. 24, no. 2, pp. 137–154, Sep. 1997.

[23] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, and G. Marchal, "Automated multimodality image registration based on information theory," in *Proc. 14th Int. Conf. Inf. Process. Med. Imag.*, France, 1995, pp. 263–274.

[24] J. P. Kern and M. S. Pattichis, "Robust multispectral image registration using mutual-information models," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1494–1505, May 2007.

[25] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Mutual-information-based registration of medical images: A survey," *IEEE Trans. Med. Imag.*, vol. 22, no. 8, pp. 986–1004, Aug. 2003.

[26] S. Klein, M. Staring, and J. P. W. Pluim, "Evaluation of optimization methods for nonrigid medical image registration using mutual information and b-splines," *IEEE Trans. Image Process.*, vol. 16, no. 12, pp. 2879–2890, Dec. 2007.

[27] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Image registration by maximization of combined mutual information and gradient information," *IEEE Trans. Med. Imag.*, vol. 19, no. 8, pp. 809–814, Aug. 2000.

[28] A. Cole-Rhodes, K. Johnson, and J. L. Moigne, "Multiresolution registration of remote sensing images using stochastic gradient," in *Proc. SPIE Conf. Wavelet Ind. Compon. Anal. Appl. IX*, Orlando, FL, 2002, vol. 4738, pp. 44–55.

[29] A. A. Cole-Rhodes, K. L. Johnson, J. LeMoigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1495–1511, Dec. 2003.

[30] H.-M. Chen, P. K. Varshney, and M. K. Arora, "Performance of mutual information similarity measure for registration of multitemporal remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 11, pp. 2445–2454, Nov. 2003.

[31] H. M. Chen, M. K. Arora, and P. K. Varshney, "Mutual information-based image registration for remote sensing data," *Int. J. Remote Sens.*, vol. 24, no. 18, pp. 3701–3706, 2003.

[32] A. A. Cole-Rhodes and P. K. Varshney, "Image registration using mutual information," in *Image Registration for Remote Sensing*, J. L. Moigne, N. S. Netanyahu, and R. D. Eastman, Eds. Cambridge: Cambridge Univ. Press, 2011.

[33] A. Collignon, "Multi-modality medical image registration by maximization of mutual information," Ph.D. Thesis, Catholic Univ. Leuven, Leuven, Belgium, 1998.

[34] I. Zavorin and J. Le Moigne, "Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery," *IEEE Trans. Image Process.*, vol. 14, no. 6, pp. 770–782, Jun. 2005.

[35] P. Thévenaz and M. Unser, "Optimization of mutual information for multiresolution image registration," *IEEE Trans. Image Process.*, vol. 9, no. 12, pp. 2083–2099, Dec. 2000.

[36] I. Sarkar and B. Manu, "A wavelet-based multiresolution approach to solve the stereo correspondence problem using mutual information," *IEEE Trans. Syst., Man, Cybern. B, Cybern.*, vol. 37, no. 4, pp. 1009–1014, Aug. 2007.

[37] F. J. Solis and R. J.-B. Wets, "Minimization by random search techniques," *Math. Oper. Res.*, vol. 6, no. 1, pp. 19–30, Feb. 1981.

[38] J. M. Rouet, J. J. Jacq, and C. Roux, "Genetic algorithms for a robust 3-D MR-CT registration," *IEEE Trans. Inf. Technol. Biomed.*, vol. 4, no. 2, pp. 126–136, Jun. 2000.

[39] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi, "Optimization by simulated annealing," *Science*, vol. 220, no. 4598, pp. 671–680, May 1983.

[40] N. Ritter, R. Owens, J. Cooper, R. H. Eikelboom, and P. P. Van Saarloos, "Registration of stereo and temporal images of the retina," *IEEE Trans. Med. Imag.*, vol. 18, no. 5, pp. 404–418, May 1999.

[41] J. A. Nelder and R. Mead, "A simplex method for function minimization," *Comput. J.*, vol. 7, no. 4, pp. 308–313, 1965.

[42] V. Zagrodsky, R. Shekhar, and J. F. Cornhill, "Multifunction extension of simplex optimization method for mutual information-based registration of ultrasound volumes," in *Proc. SPIE Conf. Med. Imag.—Image Process.*, San Diego, CA, 2001, vol. 4322, pp. 508–515.

[43] M. J. D. Powell, "An efficient method for finding the minimum of a function of several variables without calculating derivatives," *Comput. J.*, vol. 7, no. 2, pp. 155–162, 1964.

[44] F. Maes, A. Collignon, D. Vandermeulen, G. Marchal, and P. Suetens, "Multimodality image registration by maximization of mutual information," *IEEE Trans. Med. Imag.*, vol. 16, no. 2, pp. 187–198, Apr. 1997.

[45] F. Wang, B. C. Vemuri, M. Rao, and Y. Chen, "A new & robust information theoretic measure and its application to image alignment," in *Inf. Process. Med. Imag.*, vol. 2732, C. Taylor and J. A. Noble, Eds. Berlin/Heidelberg, Germany: Springer-Verlag, 2003, ser. Lecture Notes in Computer Science, pp. 388–400.

[46] F. Wang and B. C. Vemuri, "Non-rigid multi-modal image registration using cross-cumulative residual entropy," *Int. J. Comput. Vis.*, vol. 74, no. 2, pp. 201–215, Aug. 2007.

[47] F. Wang, B. C. Vemuri, M. Rao, and Y. Chen, "Cumulative residual entropy, a new measure of information & its application to image alignment," in *Proc. 9th IEEE Int. Conf. Comput. Vis.*, Nice, France, 2003, vol. 1, pp. 548–553.

[48] J. P. W. Pluim, J. B. A. Maintz, and M. A. Viergever, "Interpolation artefacts in mutual information-based image registration," *Comput. Vis. Image Understand.*, vol. 77, no. 2, pp. 211–232, Feb. 2000.

[49] G. K. Rohde, A. Aldroubi, and D. M. Healy, "Interpolation artifacts in sub-pixel image registration," *IEEE Trans. Image Process.*, vol. 18, no. 2, pp. 333–345, Feb. 2009.

[50] J. Inglada, V. Muron, D. Pichard, and T. Feuvrier, "Analysis of artifacts in subpixel remote sensing image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 1, pp. 254–264, Jan. 2007.

[51] A. A. Goshtasby, *2-D and 3-D Image registration: For medical, remote sensing, and industrial applications*. Hoboken, NJ: Wiley, 2005.

[52] P. Viola and W. M. Wells, III, "Alignment by maximization of mutual information," in *Proc. 5th Int. Conf. Comput. Vis.*, Cambridge, MA, 1995, pp. 16–23.

[53] C. Studholme, D. L. G. Hill, and D. J. Hawkes, "Automated 3-D registration of MR and CT images of the head," *Med. Image Anal.*, vol. 1, no. 2, pp. 163–175, 1996.
[54] M. Rao, Y. Chen, B. C. Vemuri, and F. Wang, "Cumulative residual entropy: A new measure of information," *IEEE Trans. Inf. Theory*, vol. 50, no. 6, pp. 1220–1228, Jun. 2004.
[55] M. R. Pickering, Y. Xiao, and X. Jia, "Registration of multi-sensor remote sensing imagery by gradient-based optimization of cross-cumulative residual entropy," in *Proc. SPIE Conf. Algorithms Technol. Multispectral, Hyperspectral, and Ultraspectral Imagery XIV*, Orlando, FL, 2008, vol. 6966, pp. 69 660U1–69 660U10.

Mahmudul Hasan (S'10) received the B.Sc. degree in computer science and engineering from Rajshahi University of Engineering and Technology, Rajshahi, Bangladesh, in 2005. Currently, he is working toward the Ph.D. degree in electrical engineering from the School of Engineering and Information Technology, University of New South Wales, Australian Defence Force Academy, Canberra, Australia.

His current research interests include remote sensing image processing and image registration.

Mark R. Pickering (S'92–M'95) received the B.E. degree in electrical engineering from the Capricornia Institute of Advanced Education, Rockhampton, Australia, in 1988, and the M.E. and Ph.D. degrees in electrical engineering from the University of New South Wales, Australian Defense Force Academy, (UNSW@ADFA), Canberra, Australia, in 1991 and 1995, respectively.

He was a Lecturer from 1996 to 1999 and a Senior Lecturer from 2000 to 2009 with the School of Electrical Engineering and Information Technology, UNSW@ADFA. He is currently an Associate Professor with the School of Engineering and Information Technology, UNSW@ADFA. His current research interests include image registration, image and video coding, information security, pattern recognition, and error-resilient data transmission.

Xiuping Jia (M'93–SM'03) received the B.Eng. degree from the Beijing University of Posts and Telecommunications, Beijing, China, in 1982, and the Ph.D. degree in electrical engineering from The University of New South Wales, Canberra, Australia, in 1996.

Since 1988, she has been with the School of Information Technology and Electrical Engineering, University College, The University of New South Wales, Canberra, where she is currently a Senior Lecturer. She is also a Guest Professor with Harbin Engineering University, China. She is the coauthor of the remote sensing textbook titled *Remote Sensing Digital Image Analysis* [Springer-Verlag, 3rd (1999) and 4th eds. (2006)]. Her research interests include remote sensing and imaging spectrometry.

Dr. Jia is an Associate Editor for the IEEE TRANSACTION ON GEOSCIENCE AND REMOTE SENSING, and for CPGIS Annals of GIS.