Fast phase-based registration of multimodal image data

Alexander Wong *, Paul Fieguth

Systems Design Engineering, University of Waterloo, 200 University Avenue West, Waterloo, Ontario, Canada N2L 3G1

ARTICLE INFO

ABSTRACT

An interesting problem in pattern recognition is that of image registration, which plays an important role in many vision-based recognition and motion analysis applications. Of particular interest among registration problems are multimodal registration problems, where the images exist in different feature spaces. State-of-the-art phased-based approaches to multimodal image registration methods have provided good accuracy but have high computational cost. This paper presents a fast phase-based approach to registering multimodal images for the purpose of initial coarse-grained registration. This is accomplished by simultaneously performing both globally exhaustive dynamic phase sub-cloud matching and polynomial feature space transformation estimation in the frequency domain using the fast Fourier transform (FFT). A multiscale phase-based feature extraction method is proposed that determines both the location and size of the dynamic sub-clouds being extracted. A simple outlier pruning based on resampling is used to remove false keypoint matches. The proposed phase-based approach to registration can be performed very efficiently without the need for initial estimates or equivalent keypoints from both images. Experimental results show that the proposed method can provide accuracies comparable to the state-of-the-art phase-based image registration methods for the purpose of initial coarse-grained registration while being much faster to compute.

© 2008 Elsevier B.V. All rights reserved.

1. Introduction

Image registration is the process of matching points in one image to their corresponding points in another image. The problem of image registration plays a very important role in many visual and object recognition and motion analysis applications. Some of these applications include visual motion estimation [1,2], vision-based content-based retrieval [3,4], image registration [5–7,9], and biometric authentication [10]. In the best case scenario, the images exist at the same scale, in the same orientation, as well as represented in the same feature space. However, this is not the case in most real-world applications. There are many situations where the images exist in different feature spaces. This particular problem will be referred to as the multimodal registration problem and is a particularly difficult problem to solve. Examples of this problem in real-world situations include medical image registration and tracking of MRI/CT/PET data [11] and building modeling and visualization using LIDAR and optical data [12,13].

There are several important issues that make multimodal registration a difficult problem to solve. First, many registration algorithms require that equivalent keypoints be identified within each image. However, given the differences between feature spaces in which the images exist, it is often a very difficult task. The significant differences between feature spaces also make it impractical to perform direct intensity matching between the two images. In recent years, an effective approach to

---

* Corresponding author. Tel.: +1 647 280 1947; fax: +1 519 746 4791.
E-mail addresses: alexanderwong@einfodaily.com,
a28wong@engmail.uwaterloo.ca (A. Wong), pfieguth@uwaterloo.ca
(P. Fieguth).

multimodal registration has been proposed that utilizes local phase [14,15]. This state-of-the-art approach evaluates the mutual information between the local phase of two images to determine the optimal alignment and has been shown to be very effective at matching multimodal medical image data, outperforming existing multimodal registration methods [14,15]. However, this approach is computationally expensive ($O(N^6)$ for the mutual information evaluation process). As such, a registration method that is able to take advantage of local phase information to determine point correspondences between images while being computationally efficient is highly desired for the purpose of initial coarse-grained registration.

The main contribution of this paper is fast phase-based registration algorithm for aligning multimodal images. The proposed method is designed to provide a fast alternative to the phase-based registration algorithm proposed by Mellor et al. [15]. It is important to note that the main contributions of this paper reside in the methods for keypoint detection and dynamic sub-cloud extraction, as well as the method for simultaneous phase correspondence evaluation and feature transformation estimation, not in the outlier pruning scheme. Furthermore, the proposed method is designed for fast initial coarse-grained matching and by no means guarantee the smoothness of the global data correspondence problem. A fine-grained matching method can be used after the initial matching to provide improved alignment based on global smoothness constraints.

2. Multimodal registration problem

The multimodal registration problem can be defined in the following manner. Suppose there exist two images f and g, where points in f and g are represented using two different feature spaces, respectively. For every point in f, the goal of registration is to determine a corresponding point in g such that the highest degree of correspondence can be found between f and g. This problem can be alternatively be formulated as finding the optimal transformation T that maps all points from f to the points from g such that the highest degree of correspondence can be achieved. The relationship between f and g can be defined as

$$g(\underline{x}_g) = f(T(\underline{x}_f)), \tag{1}$$

where $\underline{x}_f$ and $\underline{x}_g$ are coordinate vectors corresponding to f and g, respectively, and T is a transformation that maps points from f to g.

Based on the above relationship, the multimodal registration problem can be formulated as a minimum distance optimization problem, with the distance representing the degree of data correspondence between two images expressed as

$$T = \underset{T}{\operatorname{argmin}}[D(g(\underline{x}_g), f(T(\underline{x}_f)))], \tag{2}$$

where D is the distance function that is inversely proportional to the degree of correspondence between feature points. Low values of D indicate a high level of correspondence between the images.

3. Previous work

A large number of methods have been proposed for the purpose of image registration. In general, current methods can be grouped into four main types:

(1) Methods based on relative distances [16–19].
(2) Methods in the frequency domain [20–22].
(3) Methods based on direct comparisons [6–9,23–27].
(4) Methods based on extracted features [14,15,28–36].

Methods based on relative distances exploit the spatial relationships between neighboring pixels within an image to determine the best match between two points. These algorithms are based on the assumption that if a point in image f, $p_{f,0}$, has a corresponding point in image g, $p_{g,0}$, then there exist other points in f, $\{p_{f,1}, p_{f,2}, \ldots, p_{f,n}\}$, that have a corresponding points in g, $\{p_{g,1}, v_{g,2}, \ldots, p_{g,n}\}$, such that the distance between $p_{f,0}$ and point $p_{f,k}$ is equal to the distance between $p_{g,0}$ and point $p_{g,k}$. Methods based on relative distances are primarily useful for situations where the transformation between the images consists only of translations and rotations.

Methods in the frequency domain [20–22] exploit the frequency characteristics such as phase to estimate the transformation between two images. A common frequency domain registration method is phase correlation, where the Fourier coefficients calculated from image f are divided by that calculated from image g. Performing the inverse transform on the result yields a single peak indicating the translation that matches the two images. This technique has been extended to account for global rotations and scale by Reddy et al. [21]. As such, frequency domain methods are only suited for globally rigid point correspondences.

Methods based on direct comparisons [6–9,23–27] attempt to find data point correspondences between two images by performing point matches directly in their respective feature spaces. Many techniques in this group make use of feature information from neighboring data points to determine the similarity between two data points. Some common similarity metrics used in direct comparison methods include maximum likelihood [5], correlation [23,26], and mutual information [6–9]. Of particular interest in recent years are techniques based on mutual information, which attempt to match data points by finding the mutual dependence between the images. The key advantage of techniques based on mutual information is that it allows images existing in different feature spaces to be compared in a direct manner. However, such techniques also require good initial match estimates to produce accurate results as they are very under-constrained in nature. Furthermore, techniques based on mutual information are computationally expensive and may not be practical for certain situations where computational speed is important. A comparison between the computational complexity of mutual information-based techniques and the proposed method will be discussed later on in the paper.

In methods based on extracted features [14,15,28–36], the images are compared indirectly using extracted features that exist in a common feature space. Common

features used in such methods include contours [29,30], invariant moments [34,35], orientation [33,36], and shape properties [28]. By comparing methods using derived features in a common feature space, this allows such techniques to match images existing in different feature spaces in cases where similar features can be extracted from both images.

In recent years, an effective approach has been proposed for the purpose of multimodal registration that evaluates the mutual information between the local phase of images to determine optimal point correspondences [14,15]. This state-of-the-art approach improves mutual information performance by reducing feature non-homogeneities and accentuating structural information, and has been shown to be very effective at finding correspondences between medical imaging data acquired using different medical imaging modalities. However, this approach suffers from the same computational cost issues associated with mutual information-based methods. The primary goal of this paper is to provide a fast approach to phase-based image registration that provides comparable accuracy at a significantly reduced computational cost.

4. Proposed registration algorithm

The proposed approach to phase-based image registration can be divided into three main parts. First, keypoints are selected and sub-clouds are extracted using a multi-scale phase-based keypoint extraction method, as described in Sections 4.1 and 4.2. A simultaneous globally exhaustive phase correspondence evaluation and feature space transformation estimation scheme is performed between the phase sub-clouds from one image over all translations and local rotations of the other image, as described in Sections 4.3 and 4.4. A simple outlier pruning scheme based on resampling is used to remove erroneous matches and the transformation that maps data points from one image to the other, as described in Section 4.5. The main contributions of the proposed framework is a novel multiscale phase-based keypoint and sub-cloud extraction scheme, as well as a novel simultaneous globally exhaustive phase sub-cloud correspondence evaluation and feature space transformation estimation scheme.

4.1. Keypoint detection and sub-cloud size estimation

In conventional keypoint-based registration algorithms, a set of keypoints $\{p_{g,10}, p_{g,20}, \ldots, p_{g,n0}\}$ are extracted from each image to construct sub-clouds $\{C(p_{g,10}), C(p_{g,20}), \ldots, C(p_{g,n0})\}$ of a certain size [35]. This is done for several reasons. First, it is computationally expensive to match every data point in one image to those in the other image. As such, it is much more computationally efficient to select a subset of points within each image that can be used to estimate the overall mapping between the two images. Second, it can be said that a majority of points within a image cannot be uniquely distinguished from other points based on its neighboring points. As such, these non-unique points can

result in false point correspondences. Therefore, it is intuitive that only points that are distinctive be considered in the registration process.

One of the issues encountered in many registration methods is that they require that equivalent keypoints be identified within each image. What this means is that for every keypoint selected from image f, there must exist a corresponding keypoint selected from image g for a valid keypoint match to occur. While it is relatively easy to determine equivalent keypoints in cases where the images exist in the same feature space, it is considerably more difficult to do so in cases where the images exist in significantly different feature spaces. If equivalent keypoints are not selected in both images, correct keypoint matches will not be found. The proposed algorithm removes the need for equivalent keypoints by looking at the problem from a different perspective. Suppose that a distinctive keypoint p_g is detected in image g. Ideally, there exists a corresponding keypoint p_f in image f that is distinctive based on its neighboring points. In such a situation, it is intuitive that p_f will likely be found if p_g is compared with all points in f. Therefore, if every point in f is selected as a keypoint, a correct keypoint match will theoretically be found for every distinctive keypoint in g. In the proposed algorithm, every point in f is selected as a keypoint.

There are several advantages to this approach. First, it ensures that a correct keypoint match for a keypoint in g is possible. This is in contrast to methods that attempt to determine keypoints in both images independently, which does not guarantee a correct match is possible. Second, since only distinctive keypoints are selected from g, the presence of non-distinctive keypoints in f will have a lesser impact on the accuracy of keypoint matches. Finally, since only a small set of points are selected as keypoints from g, the computational complexity is significantly lower than an exhaustive evaluation between all possible point pairs between the two images.

Another issue that must be dealt with during the keypoint detection stage is determining the size of individual sub-clouds centered around each keypoint. The size of a sub-cloud should be adjusted in an adaptive manner to account for the feature characteristics of the underlying image. For example, smaller sub-clouds should be used in situations where the feature characteristics around a keypoint are most distinctive at a finer scale. Conversely, larger sub-clouds should be used in situations characterized by feature characteristics that are most distinctive at a larger, coarser scale. Therefore, a method for determining the size of individual sub-clouds is desired.

Various algorithms have been proposed for the purpose of keypoint detection [37,39–44]. However, there are drawbacks to these techniques. Most commonly used keypoint detection algorithms such as that proposed by Harris et al. [41], the DoG maxima method used by Lowe et al. [37], and the wavelet-based method proposed by Fauqueur et al. [39] are highly sensitive to data non-homogeneities, where the same data content is represented by different data point values due to certain conditions. For example, in the case of MRI data volumes, RF inhomogeneity conditions can result in significant data

non-homogeneities in the constructed data volume [38]. Furthermore, many of these techniques utilize Gaussian pre-filtering to suppress noise in the image, which can also significantly reduce the distinctiveness of data characteristics for the purpose of keypoint detection.

A better approach to keypoint detection that addresses the issue associated with data non-homogeneities can be developed based on the feature significance measure introduced by Kovesi [44], which uses local phase characteristics obtained from complex-valued wavelets. By utilizing only the phase information, this measure of feature significance is largely invariant to data non-homogeneities. Furthermore, this measure is highly robust to noise without the need for pre-filtering, provides improved localization over existing measures, and accounts for variations due to orientation. The proposed method builds upon this measure of feature significance to create a novel keypoint detection and sub-cloud size estimation algorithm.

The proposed keypoint detection and sub-cloud size estimation scheme can be described as follows. The local amplitude and phase at each point in image g is computed over multiple scales and orientations using complex-valued wavelets such as logarithmic Gabor wavelets. The local amplitude and phase at a particular point $\underline{x}$ at wavelet scale n is computed based on a pair of even-symmetric and odd-symmetric wavelets H_n^e and H_n^o at scale n,

$$A_n(\underline{x}) = \sqrt{(g(\underline{x}) * H_n^e)^2 + (g(\underline{x}) * H_n^o)^2}, \tag{3}$$

$$\phi_n(\underline{x}) = \tan^{-1}\left(\frac{(g(\underline{x}) * H_n^e)}{(g(\underline{x}) * H_n^o)}\right). \tag{4}$$

The local phase coherence at point $\underline{x}$, orientation θ, and over a range of N scales can then be computed as

$$\rho(\underline{x}, \theta) = \frac{\sum_n^N W(\underline{x}, \theta) \lfloor A_n(\underline{x}, \theta) \Delta\Phi(\underline{x}, \theta) - T \rfloor}{\sum_n A_n(\underline{x}, \theta) + \varepsilon}, \tag{5}$$

$$\Delta\Phi(\underline{x}, \theta) = \cos(\phi_n(\underline{x}, \theta) - \bar{\phi}(\underline{x}, \theta)) \\ - |\sin(\phi_n(\underline{x}, \theta) - \bar{\phi}(\underline{x}, \theta))|, \tag{6}$$

where W represents the frequency spread weighting factor, A_n and ϕ_n represent the amplitude and phase at wavelet scale n, respectively, $\bar{\phi}$ represents the weighted mean phase, T represents the noise threshold and ε is a small constant used to avoid division by zero. The parameters used to compute local phase coherence is the same as that described in [44]. The feature significance at point $\underline{x}$ can then be computed as the minimum moment of local phase coherence $\mu(\underline{x})$ based on the local phase coherence $\rho(\underline{x}, \theta)$ at different orientations $\{\theta_1, \theta_2, \ldots, \theta_l\}$,

$$\mu(\underline{x}) = \frac{1}{2}\sum_\theta \rho(\underline{x}, \theta)^2 \\ + \frac{1}{2}\sqrt{4\left(\sum_\theta (\rho(\underline{x}, \theta)\sin(\theta))(\rho(\underline{x}, \theta)\cos(\theta))\right)^2 \\ - \left(\sum_\theta [(\rho(\underline{x}, \theta)\cos(\theta))^2 - (\rho(\underline{x}, \theta)\sin(\theta))^2]\right)^2}. \tag{7}$$

The feature significance increases with the increase of μ. Once the minimum moment of local phase coherence has been determined, the keypoint locations of g are determined at the local minimum moment maxima using non-maximal suppression.

Once the keypoints of g have been selected, it is necessary to identify the dominant scale, s, at which the keypoint is most distinctive. This is achieved by computing the minimum moment of local phase coherence μ_s at each keypoint $p_{g,i0}$ over a fixed range of scales $\{s_1, s_2 \ldots, s_z\}$ and finding the maxima of minimum moment of local phase coherence over the scales,

$$s(p_{g,i0}) = \underset{s}{\operatorname{argmax}}(\mu_s(p_{g,i0})). \tag{8}$$

The radius of an individual sub-cloud centered at keypoint $p_{g,i0}$ can then be determined as

$$r(p_{g,i0}) = \lambda^{s(p_{g,i0})-1} r_{\min}, \tag{9}$$

where λ is the scale size and $r_{\min}$ is the minimum radius size. During testing, $\lambda = 1.5$ and $r_{\min} = 10$ was found to provide good results and so are used for testing purposes. However, these parameters can be tuned specifically based on a particular application for improved performance.

4.2. Phase sub-cloud extraction

Once the keypoints $\{p_{g,10}, p_{g,20}, \ldots, p_{g,n0}\}$ have been detected from g, a set of phase sub-clouds $\{C(p_{g,10}), C(p_{g,20}), \ldots, C(p_{g,n0})\}$ of neighboring points is extracted around the keypoints. This is based on the assumption that if a keypoint in g, $p_{g,i0}$, corresponds to a keypoint in f, $p_{f,i0}$, then its neighboring data points in g will have corresponding points in f such that the distances between $p_{g,i0}$ and its neighboring points $\{p_{g,i1}, \ldots, p_{g,im}\}$ will be equal to the distances between $p_{f,i0}$ and the corresponding points $\{p_{f,i0}, \ldots, p_{f,im}\}$, where m is the number of neighboring points. This assumption is very similar to that made in methods based on relative distances. However, by selecting a phase sub-cloud of neighboring points, the assumption is only made from a local perspective. This leads to a second assumption:

The relative distances between a keypoint $p_{g,i0}$ and its neighboring points $\{p_{g,i1}, \ldots, p_{g,im}\}$ within a sub-cloud in g is approximately equivalent to the relative distances between its corresponding keypoint $p_{f,i0}$ and its neighboring points $\{p_{f,i1}, \ldots, p_{f,im}\}$ within a sub-cloud in f. However, the relative distances between a keypoint and all other points in g is not equivalent to the relative distances between its corresponding keypoint and all other points in f.

The second assumption implies that, under geometric distortions, the spatial relationships between local neighboring points are largely maintained but variations in global spatial relationships between points can still occur. Without this assumption, it would not be feasible to match keypoints based on its neighboring data points.

For the proposed algorithm, the intermediate phase sub-cloud $E(p_{g,00})$ of a keypoint $p_{g,00}$ is determined as all points $\{p_{g,i1}, \ldots, p_{g,im}\}$ that is bounded by a circle centered

at $p_{g,00}$ and radius $r(p_{g,00})$ (as determined during the keypoint detection and sub-cloud size estimation process),

$$r(p_{g,00})^2 = (x - x_{p_{g,00}})^2 + (y - y_{p_{g,00}})^2. \tag{10}$$

To take better advantage of the spatial relationships between neighboring points while reducing the amount of non-distinctive points within a phase sub-cloud, the proposed algorithm constructs dynamically shaped phase sub-clouds $\{C(p_{g,10}), C(p_{g,20}), \ldots, C(p_{g,n0})\}$ based on keypoint triplets. Suppose that a keypoint triplet is constructed from an arbitrary keypoint in g, $p_{g,10}$, and its two nearest neighbor keypoints $p_{g,20}$ and $p_{g,30}$. A larger phase sub-cloud $C(p_{g,10})$ may be formed from the keypoint triplet by combining the individual phase sub-clouds $\{E(p_{g,10}), E(p_{g,20}), E(p_{g,n0})\}$ formed around each of the three keypoints. Therefore, the combined phase sub-cloud $C(p_{g,10})$ can be expressed as

$$C(p_{g,10}) = \bigcup_{j=1}^{3} E(p_{g,j0}). \tag{11}$$

An example of this combined sub-cloud is illustrated in Fig. 1. This combined phase sub-cloud will be referred to as a triplet phase sub-cloud. The key advantage of this triplet phase sub-cloud is that it preserves the distinctiveness of the individual phase sub-clouds while taking better advantage of the spatial relationships between keypoints. As such, a triplet phase sub-cloud is extracted for each keypoint. It is important to note that the overall shape of a triplet phase sub-cloud may vary from keypoint to keypoint as it depends on the spatial relationship between a keypoint and its two nearest neighbor keypoints.

4.3. Simultaneous sub-cloud matching and feature space transformation estimation

Once a set of triplet phase sub-clouds $\{C(p_{g,10}), C(p_{g,20}), \ldots, C(p_{g,n0})\}$ have been extracted, sub-cloud matching is performed between each keypoint p_g in g and every point p_f in f. Since triplet phase sub-clouds were not explicitly extracted from f, it is intuitive to vary the triplet phase sub-cloud $C(p_f)$ for a point in f based on the shape of the triplet phase sub-cloud $C(p_g)$ in g with which it is being evaluated against. Therefore, a different triplet phase sub-cloud $C_i(p_{f,j})$ is dynamically formed for the point in f for each triplet phase sub-cloud $C(p_{g,i})$ it is matched against in g. This concept is illustrated in Fig. 2.

As stated earlier, the use of mutual information to evaluate the similarity of local phase information, as proposed by Mellor et al. [14,15], has very high computational complexity and is impractical in certain situations. Therefore, a faster approach that produces comparable results is highly desired.

Instead of using mutual information, the proposed algorithm evaluates the similarity between triplet phase sub-clouds using a simple sum of squared distance metric. The similarity between a triplet phase sub-cloud $C(p_{g,i})$ at $\underline{p}_{g,i}$ and a triplet phase sub-cloud $C_i(p_{f,j})$ at $\underline{p}_{f,j} = (\underline{p}_{g,i} - \underline{t})$ may be expressed as the sum of squared distances between the local phase ϕ_f and ϕ_g within
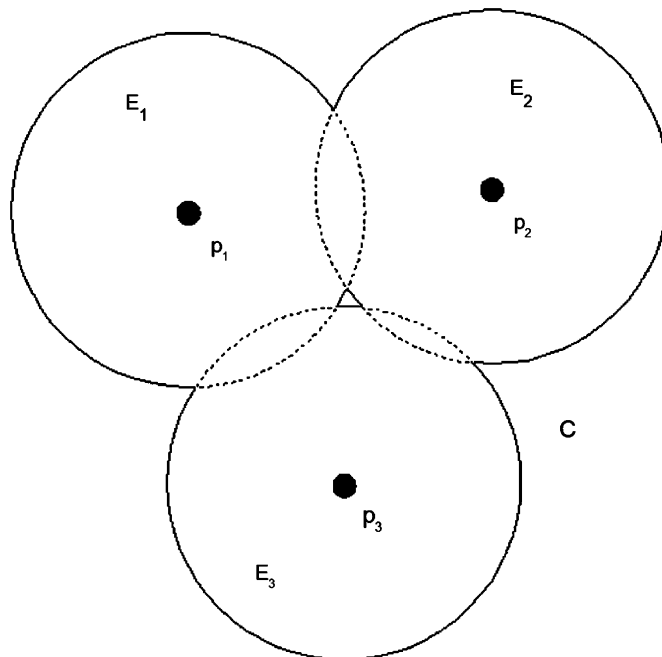

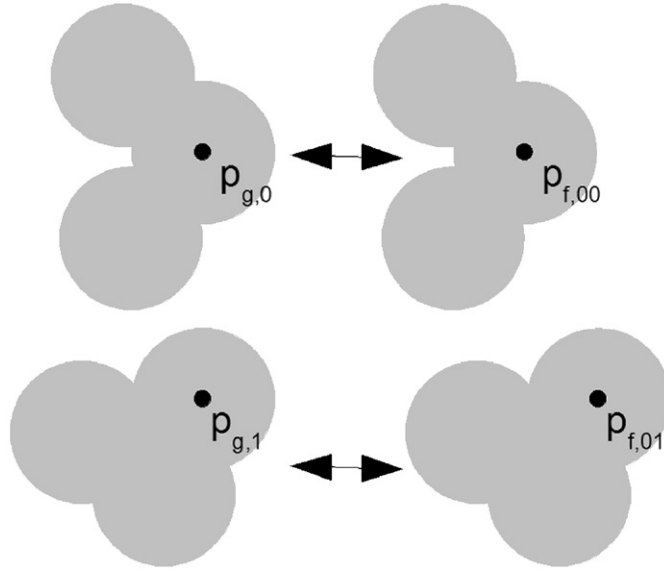**Fig. 1.** Combined phase sub-cloud from keypoint triplets.

Fig. 2. Triplet phase sub-cloud matching: note that the phase sub-cloud at data point $p_{f,0}$ changes depending on the phase sub-cloud it is being matched against in g.

the sub-clouds,

$$D(\underline{t}) = \sum_{\underline{x}} (\phi_f(\underline{x} - \underline{t}) - \phi_g(\underline{x}))^2 \kappa(\underline{x}), \tag{12}$$

where

$$\kappa(\underline{x}) = \begin{cases} \dfrac{1}{8\pi(r(p_{\min,g}))^2} \exp\left[\dfrac{1}{2}\left(\left(\dfrac{x - x_{(p_{\min,g})}}{2(r(p_{\min,g}))}\right)^2 \right.\right. \\ \left.\left. + \left(\dfrac{y - y_{(p_{\min,g})}}{2(r(p_{\min,g}))}\right)^2\right)\right], & \underline{x} \in C(\underline{p}_{g,j}), \\ 0, & \underline{x} \notin C(\underline{p}_{g,j}), \end{cases} \tag{13}$$

where $p_{\min,g}$ is the keypoint in the triplet closest to $\underline{x}$. Therefore, the minimum distance optimization problem between a triplet phase sub-cloud $C(\underline{p}_{g,i})$ over all triplet phase sub-clouds $\{C_i(\underline{p}_{f,j}) = C_i(\underline{p}_{g,i} - \underline{t}_j)\}$ can be expressed as

$$\underline{t}_i = \underset{\underline{t}}{\text{argmin}} \left[\sum_{\underline{x}} (\phi_f(\underline{x} - \underline{t}) - \phi_g(\underline{x}))^2 C_i(\underline{x}) \right], \tag{14}$$

where t_i is an optimal translation vector between keypoint $\underline{p}_{g,i}$ and its corresponding point $\underline{p}_{f,i}$. By finding t_i, the location of $\underline{p}_{f,i}$ can be determined as

$$\underline{x}_{f,i} = \underline{x}_{g,i} - \underline{t}_i. \tag{15}$$

The biggest issue with this basic similarity evaluation formulation between triplet phase sub-clouds is that, while using local phase alleviates much of the problems associated with data non-homogeneities inherit within an image, the local phase values of images with significantly different feature spaces are often not directly comparable. This is because the local phase of features in significantly different feature spaces can exist in different feature spaces themselves. Therefore, the basic formulation is poorly suited for situations characterized by images with significantly different feature spaces. Mellor et al. address this issue through the use of mutual information, which can be seen as a statistical approach to implicit feature space transformation, as the feature space changes depending on the alignment of the images. However, this approach is computationally expensive, particularly for large images.

In the proposed algorithm, a fast polynomial approximation approach to implicit feature space transformation is used instead to find the correspondence between phase sub-clouds. Rather than dynamically changing the feature spaces of both phase sub-clouds, the proposed algorithm attempts to determine a feature space transformation function that transforms the feature space of a triplet phase sub-cloud in f to that of a triplet phase sub-cloud it is being evaluated against in g. As such, a different feature space transformation function is implicitly computed for each pair of phase sub-clouds being evaluated.

The proposed algorithm models the feature space transformation function between two phase sub-clouds as a polynomial function,

$$W(\underline{x}, a_0, \ldots, a_n) = \sum_{i=0}^{n} a_i \phi_f(\underline{x})^i, \tag{16}$$

where a_i is the ith coefficient of the polynomial. Integrating the nth-order feature space transformation function into the sum of squared distance similarity function yields the following expression:

$$D_i(\underline{t}, a_0, \ldots, a_n) = \sum_{\underline{x}} (W(\underline{x} - \underline{t}) - \phi_g(\underline{x}))^2 C_i(\underline{x}). \tag{17}$$

Given this new combined phase sub-cloud similarity evaluation and feature space transformation estimation

function, the minimum squared distance problem must also account for the coefficients $\{a_0, \ldots, a_n\}$. The modified minimum distance optimization problem between a triplet phase sub-cloud $C(\underline{p}_{g,i})$ over all triplet phase sub-clouds $\{C_i(\underline{p}_{f,j}) = C_i(\underline{p}_{g,i} - \underline{t}_j)\}$ can be expressed as

$$\{\underline{t}_i, a_{i0}, \ldots, a_{in}\} = \underset{\{\underline{t}, a_0, \ldots, a_n\}}{\operatorname{argmin}}[D_i(\underline{t}, a_0, \ldots, a_n)]. \tag{18}$$

By combining the phase sub-cloud similarity metric and feature space transformation model into a single combined metric, both phase sub-cloud matching and feature space transformation estimation processes are evaluated in a simultaneous manner.

One aspect that has not been accounted so far is the effect of local rotations on phase sub-cloud matches. While the assumptions made about local spatial relationships still hold true, the coordinates of the data points being compared within each phase sub-cloud are no longer the same. One method of improving the robustness of the proposed algorithm to local rotations is to introduce additional sub-clouds by rotating the sub-cloud boundary for a sub-cloud at different angles. The sub-cloud boundary of the new sub-clouds at keypoint $p_{g,i}$ can be expressed as

$$C_{i,\underline{\theta}}(\underline{x}) = R_\theta C_i(\underline{x}), \tag{19}$$

where R_θ is the rotation matrix for $\underline{\theta}$. What this means is that a triplet phase sub-cloud within f must also be compared to the newly introduced phase sub-clouds within g as well. Integrating local rotations into (18) produces the following problem formulation:

$$\{\underline{t}_i, \underline{\theta}_i, a_{i0}, \ldots, a_{in}\} = \underset{\{\underline{t}, \underline{\theta}, a_0, \ldots, a_n\}}{\operatorname{argmin}}[D_{i,\underline{\theta}}(\underline{t}, a_0, \ldots, a_n)], \tag{20}$$

where $D_{i,\underline{\theta}}$ is the similarity between a sub-cloud in f and a sub-cloud in g created by rotating its sub-cloud boundary by $\underline{\theta}$. For testing purposes, a first-order polynomial feature transformation model and a discrete sub-sampled set of local rotations was used as it was found to provide good accuracy for the testing data under evaluation. The experimental results using this configuration can be seen in Section 6. These parameters can be increased depending on the type of images being evaluated.

4.4. Solving the simultaneous matching and feature space transformation estimation problem in the frequency domain

As the underlying goal of the proposed method is to present a faster alternative to the phase mutual information method proposed by Mellor et al. [14,15], a very important factor that must be considered in solving the simultaneous matching and feature space transformation estimation problem is performance. While the underlying sum of squared distance formulation is simple and fast to compute compared to mutual information, the fact that the problem must be minimized over translation, rotation, and polynomial coefficient parameters can make it computationally expensive to solve in a direct manner. Therefore, it is necessary to further reduce the computational complexity of the problem so that an efficient implementation can be achieved.

To significantly reduce the computational complexity of the proposed combined phase sub-cloud matching and feature space transformation estimation problem, we propose re-formulating the problem in a way that can be evaluated efficiently in the frequency domain. This approach has been used in other forms to reduce the computational complexity of computing correlation-based distances [23,26,45]. We extend this basic approach so that it can be used to evaluate our combined sub-cloud matching and feature space transformation estimation problem.

Suppose that a first-order polynomial model is used to estimate the feature space transformation. At an arbitrary rotation $\underline{\theta}$, the similarity between a triplet phase sub-cloud $C(\underline{p}_{g,i})$ over all triplet phase sub-clouds $\{C_i(\underline{p}_{f,j}) = C_i(\underline{p}_{g,i} - \underline{t}_j)\}$ can be expressed as

$$D_{i,\theta}(\underline{t}, a_0, a_1) = \sum_{\underline{x}} (a_1 \phi_f + a_0 - \phi_g)^2 C_i, \tag{21}$$

where ϕ_f represents $\phi_f(\underline{x} - \underline{t})$, ϕ_g represents $\phi_g(\underline{x})$, and C_i represents $C_i(\underline{x})$. The similarity equation in (21) can be rewritten such that all terms involving $\underline{t}$ can be expressed as convolutions with respect to $\underline{t}$,

$$\begin{aligned} D_{i,\theta}(\underline{t}, a_0, a_1) &= a_1^2 (\overline{\phi}_f^2 * C_i)(\underline{t}) + 2a_0 a_1 (\overline{\phi}_f * C_i)(\underline{t}) \\ &\quad - 2a_1 (\overline{\phi}_f * (\phi_g C_i))(\underline{t}) + a_0^2 \sum_{\underline{x}} C_i \\ &\quad - 2a_0 \sum_{\underline{x}} \phi_g C_i + \sum_{\underline{x}} \phi_g^2 C_i, \end{aligned} \tag{22}$$

where $\overline{\phi}_f(\underline{x}) = \phi_f(-\underline{x})$. The convolution terms can be solved efficiently in the frequency domain as multiplications while the non-convolution terms can be solved directly in an efficient manner. Therefore, the final similarity equation becomes

$$\begin{aligned} D_{i,\theta}(\underline{t}, a_0, a_1) &= a_1^2 F^{-1}(F(\overline{\phi}_f^2)F(C_i))(\underline{t}) + 2a_0 a_1 (F(\overline{\phi}_f)F(C_i))(\underline{t}) \\ &\quad - 2a_1 F^{-1}(F(\overline{\phi}_f)F(\phi_g C_i))(\underline{t}) \\ &\quad + a_0^2 \sum_{\underline{x}} C_i - 2a_0 \sum_{\underline{x}} \phi_g C_i + \sum_{\underline{x}} \phi_g^2 C_i, \end{aligned} \tag{23}$$

where F and F^{-1} is the fast Fourier transform (FFT) and IFFT, respectively. This solves the similarity and feature space transformation estimation function for all translations simultaneously and is significantly faster than solving it in a direct fashion. Furthermore, it is important to note that the values $F(\overline{\phi}_f^2)$ and $F(\overline{\phi}_f)$ are independent of $\underline{\theta}$ and C_i. Therefore, both of these values need only be calculated once during the entire phase sub-cloud matching and feature space estimation process. By caching the results of $F(\overline{\phi}_f^2)$ and $F(\overline{\phi}_f)$ for subsequent uses, computational complexity is significantly reduced as the number of forward FFTs that need to be performed for all matches and feature space transformation estimations beyond the initial match is reduced by half. The polynomial coefficients a_0 and a_1 can then be determined efficiently by solving for the minimum of the parabolic form analytically for each translation.

4.5. Outlier pruning through resampling

The sub-cloud matching and feature space transformation estimation process produces a set of keypoint matches between the images f and g. However, due to the difficulty of matching sub-clouds that are represented by different feature spaces, it is inevitable that false keypoint matches will occur. These false keypoint matches can significantly reduce the performance of the point transformation parameter estimation. The popular least squares parameter estimation methods are particularly sensitive to outliers. To improve the robustness of the parameter estimation process, it is necessary to remove potential outliers from the set of keypoint matches.

The outlier pruning problem can be viewed as a classification problem involving two classes: (i) inliers and (ii) outliers. In general, the sample distribution for the outlier and inlier classes are not known a priori. Furthermore, it is often the case that the sample distribution of the classes cannot be well characterized using known distributions such as a Gaussian distribution. As such, statistical methods such as maximum a posteriori (MAP) and parametric methods such as generalized Euclidean distance (GED) classifiers are not well suited for this type of problem. One approach that is useful for this situation is to construct a classifier ensemble through resampling. Popular robust statistics techniques that utilize resampling include the random sample consensus (RANSAC) algorithm [47] and least median of squares (LMS) [48]. There are several advantages to using resampling methods. First, no assumptions are made about the class distributions, but are based on actual data samples. As such, it is very robust in situations where parametric assumptions about the data distributions cannot be made. Furthermore, resampling methods are conceptually simple in nature and can be implemented effectively.

It is important to first define the classifiers used in the classifier ensemble. Given a set of N matched keypoint pairs, suppose that each matched keypoint pair consists of two data points $\underline{x}_g$ and $\underline{x}_f$. Let $\hat{T}$ denote an estimated transformation matrix produced using a subset of matched keypoint pairs such that

$$\hat{\underline{x}}_g = \hat{T}\underline{x}_f. \tag{24}$$

Assuming that $\hat{T}$ is the estimated transformation matrix for all keypoint pairs within the sample set, a keypoint pair may be considered an outlier based on the distance between $\hat{\underline{x}}_g$ and $\underline{x}_g$. Therefore, a distance-based classifier can be defined by the following likelihood function:

$$l(\underline{x}_f, \underline{x}_g) = (\hat{T}\underline{x}_f - \underline{x}_g)^2 \begin{cases} > \varepsilon & outlier, \\ < \varepsilon & inlier, \end{cases} \tag{25}$$

where ε is a threshold value separating outliers and inliers. As such, training the above classifier is equivalent to estimating the transformation matrix $\hat{T}$ based on a subset of matched keypoint pairs.

Given N keypoint pairs, a classification ensemble can be constructed using the above distance-based classifier in the following manner:

(1) By drawing randomly (with replacement) from the N keypoint pairs, create K data sets consisting of R keypoint pairs, where R is the number of keypoint pairs needed to estimate $\hat{T}$.
(2) Learn a classifier for each data set to produce K classifiers.
(3) Classify the N keypoint pairs using the K classifiers.
(4) Select the top M classifiers with the highest number of inliers.
(5) If the number of pairs identified as inliers for each of the selected classifiers is greater than ρ or L iterations has been reached, then a final set of inliers is determined as

$$I_{final} = \bigcup_{j \in M'} I_j, \tag{26}$$

where I_j is the set of pairs classified as inliers by classifier j, and M' is a subset of classifiers in M that satisfies the following:

$$\sup\{n_{inliers,m} : m \in M\} - n_{inliers,k} < \tau, \tag{27}$$

where $n_{inliers,k}$ is the number of pairs classified as inliers by classifier k, and τ is a threshold used to prevent poor classifiers from being used in the combination process.
(6) If the number of pairs classified as inliers for each of the selected classifiers is less than ρ, go to Step 1 and repeat the process.

The value of ρ should be selected based on the estimated probability that a keypoint pair in the total set of matched keypoint pairs is an inlier. For the purpose of testing, $\rho = 0.85N$, $\tau = 0.05 \sup\{n_{inliers,m} : m \in M\}$, $K = 20$, $M = 2$, and $L = 100$.

Once the final set of inlier matched keypoint pairs have been found, point transformation parameter estimation is performed to determine a transformation function that maps the data points in f to their corresponding data points in g. This can be accomplished using least squares transformation estimation techniques such as those described in [49–51].

4.6. Algorithm outline

Based on the theory presented, the registration algorithm can be outlined as follows:

(1) Given two images f and g:
(2) Extract a set of keypoints $\{p_{g,10}, p_{g,20}, \ldots, p_{g,n0}\}$ from g using the keypoint detection scheme proposed in Section 4.1.
(3) Extract a set of phase sub-clouds $\{C(p_{g,10}), C(p_{g,20}), \ldots, C(p_{g,n0})\}$ for each keypoint in g using the scheme proposed in Section 4.2.
(4) For each phase sub-cloud $C(p_{g,i0})$ detected from g, perform the simultaneous sub-cloud matching and feature space transformation estimation scheme proposed in Section 4.4 to obtain a set of keypoint pairs.
(5) Perform outlier rejection on the set of sub-cloud matches using the resampling method described in Section 4.5.

(6) Using the final set of keypoint pairs, perform point transformation parameter estimation to find the alignment between the two images.

Table 1
RMSE results for multimodal image data.

Test set	RMSE			
	Manual selection	NMI	PMI [15]	Proposed
Test 1	2.6562	10.8123	3.0959	2.6720
Test 2	3.1804	7.0775	3.5697	3.2097
Test 3	2.7876	8.8763	3.6461	2.8215
Test 4	2.0163	3.2437	3.0860	2.7925
Test 5	1.2538	2.8418	2.3655	1.5963
Test 6	2.5170	2.9920	2.6068	2.5342
Test 7	2.6403	3.1478	2.8311	2.8915

The RMSE is computed as the average of 10 test trials.

5. Computational complexity analysis

It is helpful to put the computational complexity reduction gained from using the proposed method into perspective. Supposed that the two images f and g being matched are each of size $N \times N$. As a simplification, suppose that the orientation of both images are the same. If the two images were compared in a direct manner for a single translation, the computational complexity of the comparison is $O(N^2)$. Since there are N^2 possible translations, the computational complexity to perform an exhaustive comparison is $O(N^4)$. Taking the two polynomial coefficients a_0 and a_1 from (21) into account, the computational complexity grows to $O(N^6)$.

If the two images were compared using the efficient approach presented in Section 4.4, the computational complexity of performing an exhaustive comparison for
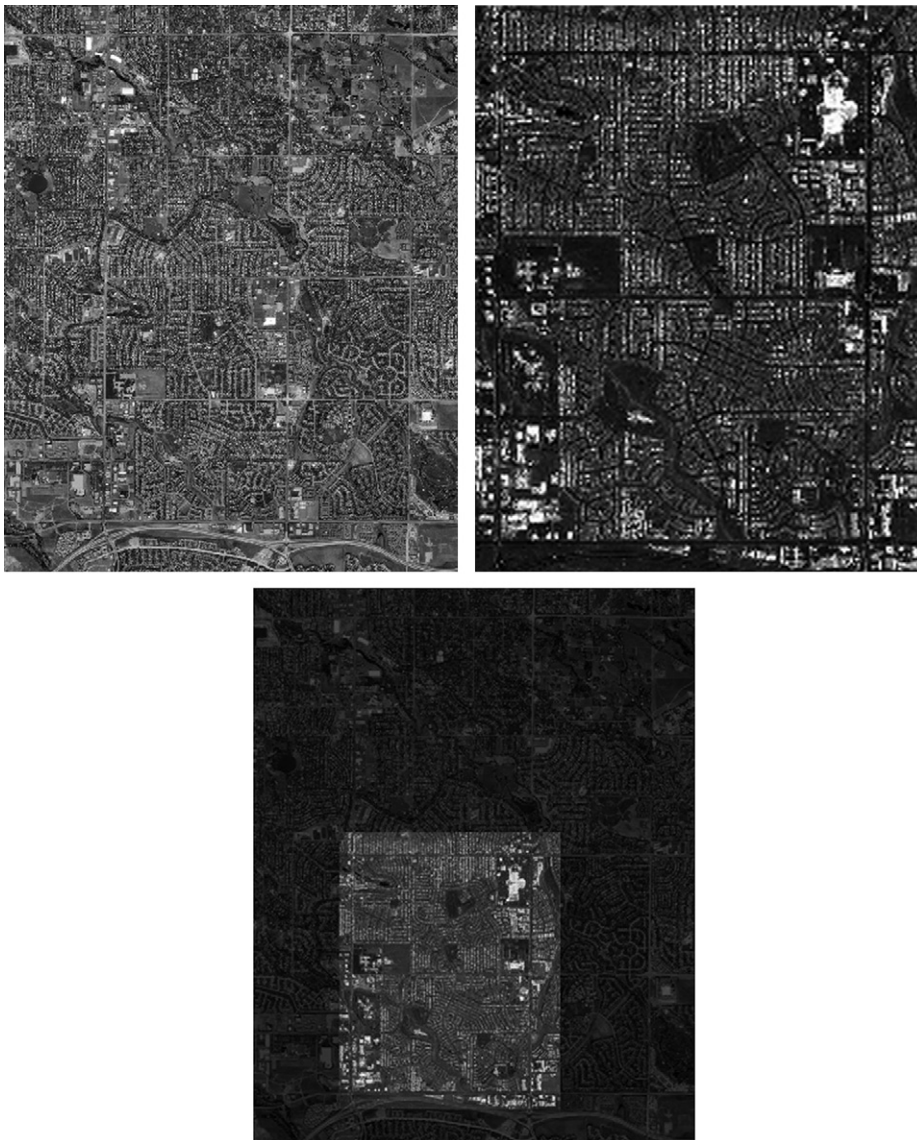

Fig. 3. Registration results from Test 2 test set; Top-left: optical data; Top-right: LIDAR data; Bottom: registered result.

all possible integer translations is reduced to $O(N^2 \log N)$. The computation of the polynomial coefficients a_0 and a_1 using the efficient approach has a computational complexity of $O(N^2)$. Therefore, the overall computational complexity of the efficient approach is $O(N^2 \log N)$, compared to $O(N^6)$ for the direct approach. To put the computational complexity differences between the proposed method and the phase mutual information method proposed by Mellor and Brady [15] into perspective, the computational complexity of the mutual information evaluation process for all N^2 possible translations is $O(N^6)$ if computed using a direct approach. With the use of a fast mutual information estimation method such as that proposed in [46], the computational complexity of the mutual information evaluation process is reduced to $O(N^4)$, which is still significantly higher than the proposed method. As such, the proposed method has significantly lower computational complexity, which is very important in situations where large images needs to be compared.

6. Experimental results

The effectiveness of the proposed algorithm was tested by applying it to a range of multimodal image data. The sub-cloud matching problem has a total of five degrees of freedom (two for translation, one for rotation, and two for the polynomial coefficients for the first-order polynomial feature space transformation model). The proposed algorithm was tested using real-world multimodal data sets obtained from Intermap Technologies Inc., the Visible Human Project, and the U.S. Geological Survey (USGS) Global Visualization Viewer project. A description of each test set is described below.

(1) *Test* 1: Pair of LIDAR DEM and orthorectified air-photo data of Highlands Ranch, CO, NW quad, 1 m resolution. This test set was provided by Intermap Technologies Inc.
(2) *Test* 2: Pair of LIDAR DEM and orthorectified air-photo data of Highlands Ranch, CO, NE quad, 1 m resolution. This test set was provided by Intermap Technologies Inc.
(3) *Test* 3: Set of satellite imaging data pair from USGS project, 240 m resolution. Data 1: Sensor: Landsat 7 ETM+, Band: 3. Data 2: Sensor: Landsat 4–5 TM, Band: 5.
(4) *Test* 4: Set of medical imaging data pair from Visible Human project, 1 mm resolution. Data 1: Sensor: T2-weighted MRI, Axial Cranial Slice. Data 2: Sensor: CT, Axial Cranial Slice.
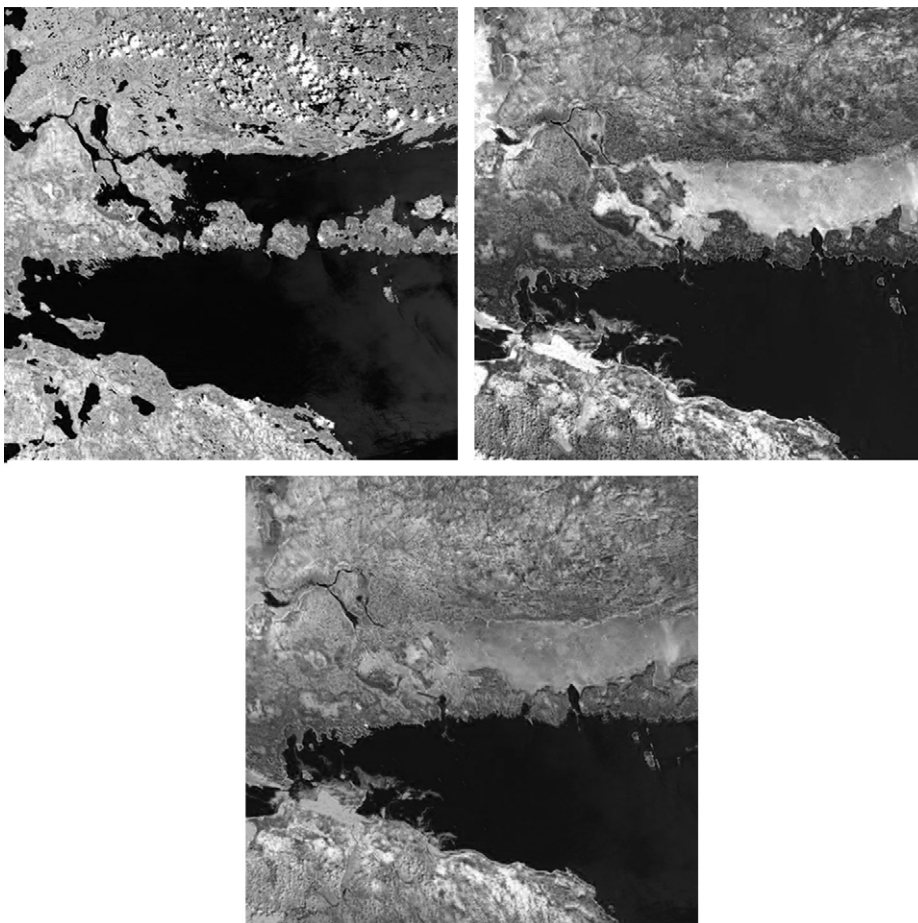

Fig. 4. Registration results from Test 3 test set; Top-left: Landsat 7 ETM + data; Top-right: Landsat 4–5 TM data; Bottom: registered result.
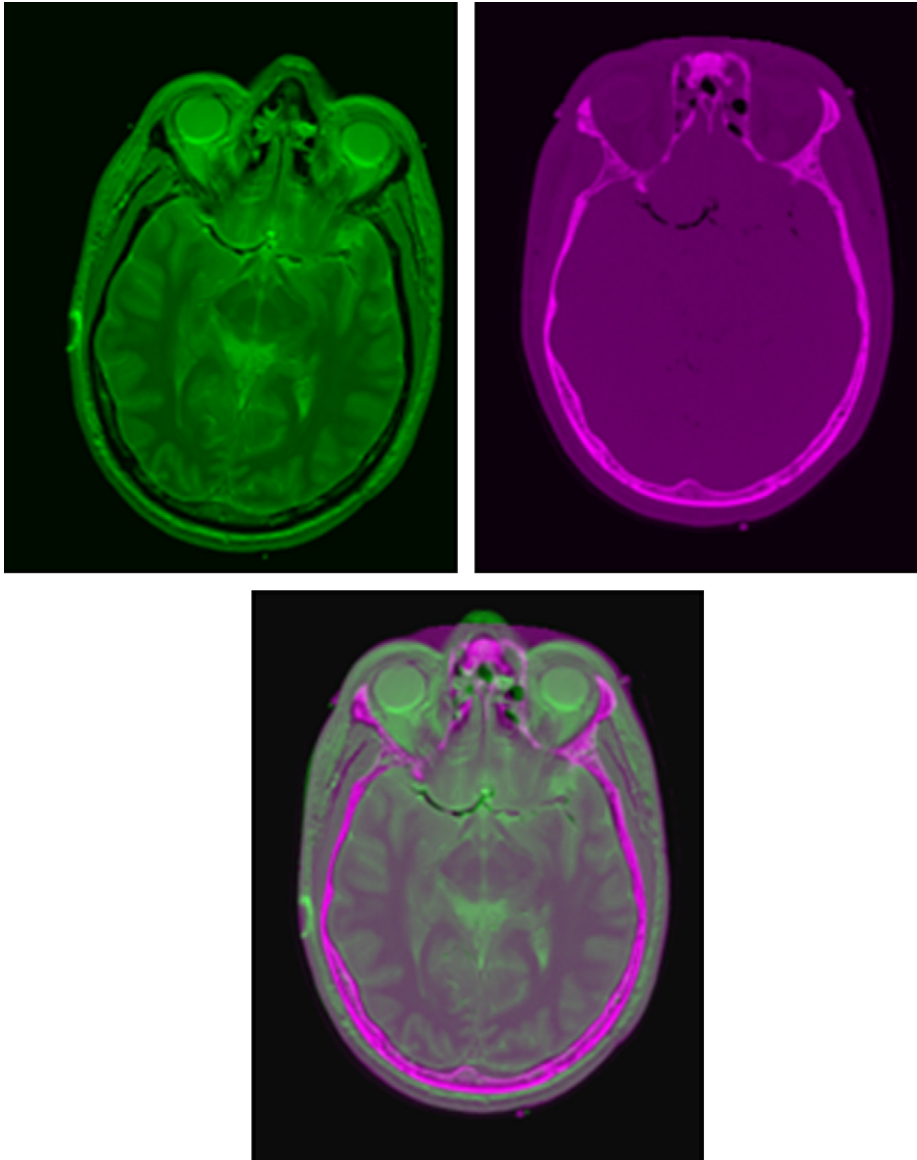
Fig. 5. Registration results from Test 4 test set; Top-left: T2-weighted MRI data; Top-right: CT data; Bottom: registered result.

(5) *Test* 5: Set of medical imaging data pair from Visible Human project, 1 mm resolution. Data 1: Sensor: T1-weighted MRI, Axial Cranial Slice. Data 2: Sensor: CT, Axial Cranial Slice.

(6) *Test* 6: Set of medical imaging data pair from Visible Human project, 1.875 mm resolution. Data 1: Sensor: T2-weighted MRI, Coronal Torso Slice. Data 2: Sensor: PD-weighted MRI, Coronal Torso Slice.

(7) *Test* 7: Set of medical imaging data pair from Visible Human project, 1.875 mm resolution. Data 1: Sensor: T2-weighted MRI, Coronal Pelvis Slice. Data 2: Sensor: PD-weighted MRI, Coronal Pelvis Slice.

Of particular difficulty are the test sets Tests 1 and 2 involving pairs of LIDAR and optical data, as the feature space used to represent LIDAR data is significantly different than that used to represent optical data. For all test sets, a total of 100 keypoints was automatically selected from image f and the range of rotations is set to iterations of $5°$. To quantitatively evaluate the performance of the proposed algorithm, each test set is subjected to 10 random affine transformations and the average root mean squared error (RMSE) is calculated on a unit distance scale based on 20 gold standard keypoint pairs manually selected from the images. For example, if the data is at a 1 m resolution, the unit distance is 1 m. As a comparison, another 20 gold standard keypoint pairs were manually selected from the images and used to estimate a transformation function. Furthermore, an implementation of image registration by maximization
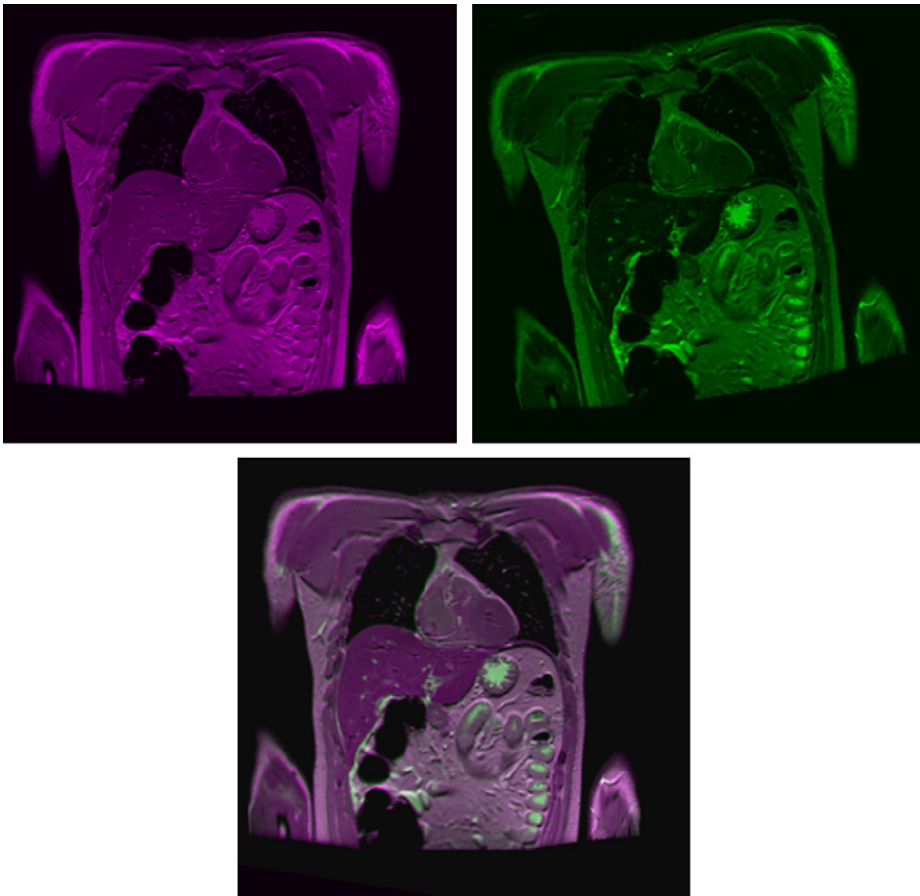
Fig. 6. Registration results from Test 6 test set; Top-left: T2-weighted MRI data; Top-right: PD-weighted MRI data; Bottom: registered result.

of normalized mutual information, as well as the state-of-the-art phase mutual information method proposed by Mellor et al. [14,15] was also used for comparison. It is important to note that all the algorithms are set up for coarse-grained registration purposes for a fair comparison since the purpose of the proposed method is for performing coarse-grained registration. As such, all algorithms are evaluated on a pixel level basis as opposed to a sub-pixel basis, thereby significantly reduces computational complexity. In practice, the results obtained from coarse-grained registration is used to initialize the fine-grain registration process to obtain the final registration between two images.

The RMSE for each test set is summarized in Table 1. It can be observed that the average RMSE realized using the proposed algorithm is lower than that achieved using the maximization of mutual information method for all test sets. Furthermore, it can be observed that the average RMSE realized using the proposed algorithm is comparable to that achieved by the transformation function produced using gold standard keypoints for all test sets. This is important as the manual selection of keypoints is a very complex and tedious process, and so is highly undesired when compared to automated methods that provide comparable registration performance. More importantly, it can be observed that the average RMSE

realized using the proposed algorithm is comparable to that achieved using Mellor's method in Tests 6–7 and noticeably lower in Tests 1–5, while having significantly lower computational complexity. The proposed method is particularly effective compared to the other methods in Tests 1–3, which consists of images with highly complex characteristics. It can also be observed that the accuracy of the proposed method has not been reduced despite the complexity of these test cases compared to the other test cases. Examples of images are presented in Figs. 3–6. It can be seen from the results that the proposed algorithm provides a fast approach to phase-based matching of multimodal images while achieving good accuracy.

7. Conclusions and future work

In this paper, we have introduced a fast approach to phase-based registration algorithm for multimodal image data. A novel multiscale phase-based feature extraction method was proposed for determining both the location and size of the dynamic phase sub-clouds being extracted. Furthermore, a fast method for simultaneously performing both globally exhaustive dynamic phase sub-cloud matching and polynomial feature space transformation estimation in the frequency domain using the fast Fourier

transform (FFT) was introduced. Experimental results showed that good registration accuracy between images existing in different feature spaces can be achieved when compared with the state-of-the-art phase mutual information method, while having significantly lower computational complexity. Future work includes investigating the effect of different sub-cloud extraction methodologies and alternative similarity metrics on the registration performance of the proposed algorithm. Furthermore, since the proposed framework is designed primarily for the purpose of initial coarse-grained registration, it would be of interest to combine the proposed method with a fine-grained registration approach to achieve improved registration accuracy.

Acknowledgments

References

[1] M. Pickering, J. Arnold, M. Frater, An adaptive search algorithm for block matching motion estimation, IEEE Transactions on Circuits and Systems for Video Technology 7 (6) (1997) 906–912.

[2] C. Cheung, L. Po, Adjustable partial distortion search algorithm for fast block motion estimation, IEEE Transactions on Circuits and Systems for Video Technology 13 (1) (2003) 100–110.

[3] C. Hsu, Y. Wu, A. Chen, Content-based image retrieval by feature point matching, in: Proceedings of SPIE, Storage and Retrieval for Media Databases, vol. 4315, 2001, pp. 39–49.

[4] C. Hsu, M. Shih, Content-based image retrieval by interest-point matching and geometric hashing, in: Proceedings of SPIE, Electronic Imaging and Multimedia Technology III, vol. 4925, September 2002, pp. 80–90.

[5] W. Li, H. Leung, A maximum likelihood approach for image registration using control point and intensity, IEEE Transactions on Image Processing 13 (8) (2004) 1115–1127.

[6] A. Collignon, F. Maes, D. Delaere, D. Vandermeulen, P. Suetens, G. Marchal, Automated multi-modality image registration based on information theory, in: Proceedings of the IPMI, 1995, pp. 263–274.

[7] P. Viola, W. Wells, Alignment by maximization of mutual information, International Journal of Computer Vision 24 (2) (1997) 137–154.

[8] C. Studholme, D. Hill, D. Hawkes, An overlap invariant entropy measure of 3D medical image alignment, Pattern Recognition 32 (1) (1999) 71–86.

[9] H. Chen, P. Varshney, M. Arora, Mutual information based image registration for remote sensing data, International Journal of Remote Sensing 24 (18) (2003) 3701–3706.

[10] L. Hong, A. Jain, Integrating faces and fingerprints for personal identification, IEEE Transactions on Pattern Analysis and Machine Intelligence 20 (12) (1998) 1295–1307.

[11] D. Mattes, D.R. Haynor, H. Vesselle, T.K. Lewellen, W. Eubank, PET-CT image registration in the chest using free-form deformations, IEEE Transactions on Medical Imaging 22 (1) (2003) 120–128.

[12] L. Chen, T. Teo, J. Rau, J. Liu, W. Hsu, Building reconstruction from LIDAR data and aerial imagery, in: Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, 2005, pp. 2846–2849.

[13] G. Vosselman, Fusion of laser scanning data, maps and aerial photographs for building reconstruction, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium 1 (2002) 85–88.

[14] M. Mellor, M. Brady, Non-rigid multimodal image registration using local phase, Lecture Notes in Computer Science: Medical Image Computing and Computer-Assisted Intervention, 2004, pp. 789–796.

[15] M. Mellor, M. Brady, Phase mutual information as a similarity measure for registration, MedIA 9 (2005) 330–343.

[16] C. Wang, H. Sun, S. Yada, A. Rosenfeld, Some experiments in relaxation image matching using corner features, Pattern Recognition 16 (1983) 167–182.

[17] J. Ton, A. Jain, Registering Landsat images by point matching, IEEE Transactions on Geoscience and Remote Sensing 27 (5) (1989) 642–651.

[18] Y. Chen, G. Medioni, Object modelling by registration of multiple range images, Image Vision Computing 10 (3) (1992) 145–155.

[19] T. Jost, H. Hugli, A multi-resolution ICP with heuristic closest point search for fast and robust 3d registration of range images, in: Proceedings of the International Conference of 3D Digital Imaging and Modeling, 2003.

[20] E. Castro, C. Morandi, Registration of translated and rotated images using finite Fourier transforms, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (5) (1987) 700–703.

[21] B. Reddy, B. Chatterji, An FFT: based technique for translation, rotation and scale invariant image registration, IEEE Transactions on Image Processing 5 (8) (1996) 1266–1271.

[22] I. Zavorin, J. Le Moigne, Use of multiresolution wavelet feature pyramids for automatic registration of multisensor imagery, IEEE Transactions on Image Processing 14 (6) (2005) 770–782.

[23] A. Fitch, A. Kadyrov, W. Christmas, J. Kittler, Fast robust correlation, IEEE Transactions on Image Processing 14 (8) (2005) 1063–1073.

[24] J. Kybic, M. Unser, Fast parametric elastic image registration, IEEE Transactions on Image Processing 12 (11) (2003) 1427–1442.

[25] Z. Li, Z. Bao, H. Li, G. Liao, Image autocoregistration and InSAR interferogram estimation using joint subspace projection, IEEE Transactions on Image Processing 44 (2) (2006) 288–297.

[26] J. Orchard, Efficient global weighted least-squares translation registration in the frequency domain, in: International Conference on Image Analysis and Recognition, 2005, pp. 116–124.

[27] A. Refice, F. Bovenga, R. Nutricato, MST-based stepwise connection strategies for multipass radar data, with application to coregistration and equalization, IEEE Transactions on Geoscience and Remote Sensing 44 (8) (2006) 2029–2040.

[28] M. Ali, D. Clausi, Automatic registration of SAR and visible band remote sensing images, IEEE International Geoscience and Remote Sensing Symposium 3 (2002) 1331–1333.

[29] F. Eugenio, F. Marques, J. Marcello, A contour-based approach to automatic and accurate registration of multitemporal and multi-sensor satellite imagery, Proceedings of the IEEE International Geoscience and Remote Sensing Symposium 6 (2002) 3390–3392.

[30] V. Govindu, C. Shekhar, Alignment using distributions of local geometric properties, IEEE Transactions on Pattern Analysis and Machine Intelligence 21 (10) (1999) 1031–1043.

[31] H. Li, B. Manjunath, S. Mitra, A contour-based approach to multisensor image registration, IEEE Transactions on Image Processing 4 (3) (1995) 320–334.

[32] N. Netanyahu, J. Le Moigne, J. Masek, Georegistration of Landsat data via robust matching of multiresolution features, IEEE Transactions on Geoscience and Remote Sensing 42 (7) (2004) 1586–1600.

[33] A. Wong, W. Bishop, J. Orchard, Efficient multi-modal least-squares alignment of medical images using quasi-orientation maps, in: Proceedings of the International Conference on Image Processing, Computer Vision, and Pattern Recognition, June 2006.

[34] D. Xiaolong, S. Khorram, A feature-based image registration algorithm using improved chain-code representation combined with invariant moments, IEEE Transactions on Geoscience and Remote Sensing 37 (5, Part 2) (1999) 2351–2362.

[35] A. Wong, D. Clausi, ARRSI: automatic registration of remote sensing images, IEEE Transactions on Geoscience and Remote Sensing 45 (5, Part II) (2007) 1483–1493.

[36] M. Brown, D. Lowe, Automatic panoramic image stitching using invariant features, International Journal of Computer Vision 74 (1) (2007) 59–73.

[37] D. Lowe, Distinctive image features from scale-invariant keypoints, International Journal of Computer Vision 60 (2) (2004) 91–110.

[38] M. Oghabian, S. Mehdipour, N. Alam, The impact of RF inhomogeneity on MR image non-uniformity, in: Proceedings of the Image and Vision Computing New Zealand, 2003.

[39] J. Fauqueur, N. Kingsbury, R. Anderson, Multiscale keypoint detection using the dual-tree complex wavelet transform, in: Proceedings of the IEEE Conference on Image Processing, 2006.

[40] S. Smith, J. Brady, SUSAN—a new approach to low level image processing, International Journal of Computer Vision 23 (1997) 45–78.

[41] C. Harris, M. Plessey, A combined corner and edge detector, in: Proceedings of the 4th Alvey Vision Conference, 1988, pp. 147–151.

[42] M. Trajkovic, M. Hedley, Fast corner detection, Image and Vision Computing 16 (2) (1998) 75–87.

[43] A. Noble, Descriptions of Image Surfaces, Ph.D. Thesis, Oxford University, 1989.

[44] P. Kovesi, Phase congruency detects corners and edges, in: Proceedings of the Australian Pattern Recognition Society Conference, 2003, pp. 309–318.

[45] F. Essannouni, R. Oulad Haj Thami, A. Salam, D. Aboutajdine, An efficient fast full search block matching algorithm using FFT algorithms, International Journal of Computer Science and Network Security 6 (3B) (2006) 130–133.

[46] S. Heldmann, O. Mahnke, D. Potts, J. Modersitzki, B. Fischer, Fast computation of mutual information in a variational image registration approach, in: Proceedings of the BVM, 2004, pp. 448–452.

[47] M. Fischler, R. Bolles, Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, Communications of the ACM 24 (1981) 381–395.

[48] C. Stewart, Robust parameter estimation in computer vision, SIAM Review 41 (3) (1999) 513–537.

[49] K. Arun, T. Huang, S. Blostein, Least-squares fitting of two 3-D point sets, IEEE Transactions on Pattern Analysis and Machine Intelligence 9 (5) (1987) 698–700.

[50] S. Umeyama, Least-squares estimation of transformation parameters between two point patterns, IEEE Transactions on Pattern Analysis and Machine Intelligence 13 (4) (1991) 376–380.

[51] R. Hartley, A. Zisserman, Multiple View Geometry in Computer Vision, Cambridge University Press, Cambridge, 2001.