Fast least trimmed squares (Fast-LTS); Image registration; Parameter estimation; Random sample consensus (RANSAC); Weak affine transformation postprint

Authors: Li, Dong, Zhang, Yunhua

Date: 2017-01-04T00:00:00+00:00

Abstract

A novel image registration scheme is devised in view of the weak affine transformation, which is a kind of similarity transformation with anisotropic scales or affine transformation without shearing. Two robust algorithms are proposed to retrieve the registration parameters from the error-prone initial correspondences based on the fast least trimmed squares (Fast-LTS) and the random sample consensus (RANSAC). In terms of several criteria, the algorithms are evaluated on three carefully selected datasets from different sensors and the experimental results demonstrate that the proposed scheme and algorithms perform robustly and accurately. Our findings also indicate that the Fast-LTS-based algorithm is more stable and appropriate for image registration than the RANSAC-based algorithm although the speed is slower.

Full Text

Preamble

A Novel Approach for the Registration of Weak Affine Images

Dong Li* and Yunhua Zhang
Key Laboratory of Microwave Remote Sensing, Center for Space Science and Applied Research, Chinese Academy of Sciences
NO.1 Nanertiao, Zhongguancun, Haidian district, Beijing 100190, China

Abstract

A novel image registration scheme is devised for weak affine transformation, which is a type of similarity transformation with anisotropic scales or an affine transformation without shearing. Two robust algorithms are proposed to retrieve the registration parameters from error-prone initial correspondences based on fast least trimmed squares (Fast-LTS) and random sample consensus (RANSAC). The algorithms are evaluated on three carefully selected datasets from different sensors using several criteria, and experimental results demonstrate that the proposed scheme and algorithms perform robustly and accurately. Our findings also indicate that the Fast-LTS-based algorithm is more stable and appropriate for image registration than the RANSAC-based algorithm, although it is slower.

Keywords: Fast least trimmed squares (Fast-LTS); Image registration; Parameter estimation; Random sample consensus (RANSAC); Weak affine transformation.

*Corresponding author. Tel/Fax: 086-010-62582874.
Email Addresses: lidong@mirslab.cn (Dong Li), zhangyunhua@mirslab.cn (Yunhua Zhang).

1. Introduction

A straightforward approach for geometrical sensing of a target is to acquire its images from different geometries. For example, 3D target reconstruction can be obtained by calculating the fundamental matrix to retrieve the camera matrix based on point correspondences constructed between images acquired from different camera geometries. In radar remote sensing, this can be carried out by radargrammetry or interferometry—i.e., by measuring the scatterings of a target from two different imaging geometries with certain spatial and/or temporal baseline to acquire parallax or interferometric phase. This difference in camera or radar imaging geometry results in a warp between images that must be aligned before further joint processing can be performed. Image registration is a technique that estimates the warp function between images so that the same pixel position in both images can be mapped to the same target position in the global coordinate system. This technique has been widely used in computer vision, pattern recognition, medical imaging, and remote sensing. It forms the foundation for many applications such as target tracking and recognition [?], image rendering and mosaic [?], camera calibration and target reconstruction [?], as well as digital elevation model inversion [?] and deformation mapping of the Earth’ s surface [?].

Many registration algorithms have been proposed to date. Existing algorithms can be divided into two broad categories: area-based algorithms and feature-based algorithms [?]. This paper focuses on the latter. Feature-based algorithms retrieve the warp function by fitting constructed feature correspondences between images. Regions, lines, and points are commonly used features. Compared with other features, point features are more distinct and operable, especially for radar images with inherent speckle, layover, shadow, and foreshortening. Point features are typically extracted by analyzing local texture information using feature detectors such as Harris, Hessian, SIFT, and SURF. Initial correspondences are then constructed by optimizing certain merit functions, such as minimizing descriptor distances [?, ?] or maximizing cross-correlation [?]. The warp function is finally retrieved by fitting these correspondences using least squares (LS) or other robust techniques such as random sample consensus (RANSAC) [?] and least median of squares (LMedS) [?] to handle mismatches. Hartley and Zisserman [?] applied RANSAC to refine correspondences contaminated by large amounts of outliers for better estimation of image homography. RANSAC performs estimation by randomly sampling a minimal sampling set (MSS) from the dataset to achieve a model estimation, then checking the entire dataset to find elements consistent with the estimated model to construct a consensus set (CS). These two steps are repeated iteratively until the probability of finding a better-ranked CS drops below a certain threshold [?]. As an alternative to RANSAC, Zhang et al. [?] proposed an estimation strategy based on LMedS. A series of MSSs are drawn randomly from initial correspondences using a Monte Carlo technique to estimate the fundamental matrix. For each obtained matrix, the median of squared residuals with respect to the whole set of point correspondences is calculated, and the matrix with the minimum median is retained as the optimal estimation. Many other algorithms using different point features and estimation methods have been systematically reviewed by Brown [?] and Zitova and Flusser [?].

To estimate the warp, the parameter model of the warp function must be given beforehand. For optical camera images, the warp function between images can be modeled as a linear homography based on the pinhole model [?]. For relatively small-baseline interferometric radar images of gentle topographic scenes, the warp function can be approximated using a low-order polynomial [?]. Similarity transformation (ST), affine transformation (AT), and projective transformation (PT) are commonly considered warp models. ST models the warp as rotation, isotropic scaling, and translation.

However, scale isotropy is not always justified, especially for radar remote sensing images, which usually have different resolutions in two directions due to system parameters and imaging geometry, resulting in an anisotropic scaling transformation. For optical images, a small view change of the camera introduces tilt between images. In these cases, the warp function between images can be modeled as rotation, anisotropic scaling, and translation. We denote this warp model as weak affine transformation (WAT) to distinguish it from ST and AT. WAT can also be proven to be a special case of AT without shearing. Let $P_1(x_1, y_1) \leftrightarrow P'_1(x'_1, y'_1)$ be a point correspondence in image pair I and I'. AT can be expressed as:

$$\begin{bmatrix} x'_1 \\ y'_1 \\ 1 \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & t_x \\ a_{21} & a_{22} & t_y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ y_1 \\ 1 \end{bmatrix}$$

where A denotes a nonsingular affine matrix that can be decomposed as [?]:

$$A = UDV^T = R(\theta)\begin{bmatrix}\lambda_1 & 0\\ 0 & \lambda_2\end{bmatrix}R(\phi)^T$$

where D is a diagonal matrix composed of the singular values of matrix A, i.e., λ_1 and λ_2, and U and V are two unitary matrices that can be expressed by rotation matrices $R(\theta)$ and $R(\phi)$. The operations of forward rotation by ϕ, anisotropic scaling, and backward rotation by ϕ shear the image–i.e., transform a rectangle into an arbitrary parallelogram. From (2), one can see that shearing results from the term $-(\lambda_1 - \lambda_2)\sin\phi\cos\phi$. For images with weak affine warp, the scale anisotropy and shearing angle ϕ are always tiny, so the shearing approximates to zero. Therefore, the affine matrix A can be simplified as:

$$A = \begin{bmatrix}s_1 & 0\\ 0 & s_2\end{bmatrix}R(\theta)$$

where $s_1 = \lambda_1\cos^2\phi + \lambda_2\sin^2\phi$ and $s_2 = \lambda_1\sin^2\phi + \lambda_2\cos^2\phi$. Then the homographic model between weak affine images can be expressed as:

$$\begin{bmatrix}x'_1\\ y'_1\end{bmatrix} = \begin{bmatrix}s_1 & 0\\ 0 & s_2\end{bmatrix}R(\theta)\begin{bmatrix}x_1\\ y_1\end{bmatrix} + \begin{bmatrix}t_x\\ t_y\end{bmatrix}$$

where θ is the rotation, t_x and t_y are the translations, and s_1 and s_2 denote the scales in image directions x and y, respectively. When s_1 equals s_2, WAT degenerates to ST:

$$\begin{bmatrix}x'_1\\ y'_1\end{bmatrix} = sR(\theta)\begin{bmatrix}x_1\\ y_1\end{bmatrix} + \begin{bmatrix}t_x\\ t_y\end{bmatrix}$$

Algorithms for registration of images with affine and similarity warp have been widely proposed. However, no effective registration algorithm exists to deal with WAT-warped images so far. For these images, full affine registration is too complex and lacks robustness, while ST-oriented registration is too coarse and inaccurate. This paper addresses the registration problem for weak affine images. Contributions of this work include two aspects: (1) The WAT model is considered and a feature-based registration scheme is devised. (2) Considering the unavoidable mismatches in correspondences, two robust parameter estimation algorithms are proposed. We apply robust fast least trimmed squares (Fast-LTS) to parameter estimation and also embed the devised registration scheme into the RANSAC technique. The two algorithms are tested on three selected image pairs using several evaluation criteria, demonstrating good performance. It is also shown that the Fast-LTS-based algorithm is more stable and appropriate for image registration than the RANSAC-based algorithm.

The rest of the paper is organized as follows. Section 2 presents the devised registration scheme for WAT. The two robust parameter estimation algorithms based on Fast-LTS and RANSAC are proposed in Section 3. Evaluation of the algorithms on three image pairs from camera, optical satellite, and radar is conducted in Section 4, and Section 5 concludes the paper.

2. Registration Scheme for WAT

AT and ST expressed by (1) and (5) have six and four degrees of freedom (DOFs), respectively. Therefore, parameter retrieval for these two models is straightforward by fitting at least three or two correspondences, respectively. However, WAT expressed by (4) has five DOFs: two correspondences are inadequate while three are redundant, so registration algorithms for AT and ST cannot be applied to this special case. In this section, a registration scheme for WAT is proposed to retrieve the parameters of anisotropic scales, rotation, and translation.

Equation (4) shows the homographic model of correspondence $P_1(x_1, y_1) \leftrightarrow P_1'(x_1', y_1')$ constructed in weak affine image pair I_1 and I_2. If $P_2(x_2, y_2) \leftrightarrow P_2'(x_2', y_2')$ is another correspondence, we have a similar relation:

$$\begin{bmatrix} x_2' \\ y_2' \end{bmatrix} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} R(\theta) \begin{bmatrix} x_2 \\ y_2 \end{bmatrix} + \begin{bmatrix} t_x \\ t_y \end{bmatrix}$$

By combining (4) and (6), the following relation can be obtained:

$$\begin{bmatrix} x_2' - x_1' \\ y_2' - y_1' \end{bmatrix} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} R(\theta) \begin{bmatrix} x_2 - x_1 \\ y_2 - y_1 \end{bmatrix}$$

which can be further formulated as:

$$\begin{bmatrix} X_1' \\ Y_1' \end{bmatrix} = \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} R(\theta) \begin{bmatrix} X_1 \\ Y_1 \end{bmatrix}$$

where $X_1 = x_2 - x_1$, $Y_1 = y_2 - y_1$, $X_1' = x_2' - x_1'$, and $Y_1' = y_2' - y_1'$. Then we get:

$$Y = aX_1 + bX_2$$

where $Y = X_1'$, $X_1 = X_1$, $X_2 = Y_1$, $a = s_1 \cos\theta$, and $b = s_2 \sin\theta$. Equation (10) states that if we have two samples of (X_1, X_2, Y), we can obtain an estimation of the scale, so at least three correspondences are required. When there is no scale anisotropy (i.e., ST is considered), the scale estimation in (10) degenerates to:

$$Y = sX_1$$

where s indicates the isotropic scale. In this case, two correspondences are sufficient for estimation.

Next, we deduce the rotation relation. Let θ_1 and θ_1' be the inclination angles of lines P_2P_1 and $P_2'P_1'$, respectively. From (4) and (6), we have:

$$\tan\theta_1' = \frac{s_2}{s_1}\tan\theta_1$$

where $\theta_1 = \arctan\left(\frac{Y_1}{X_1}\right)$. Therefore, $\theta_{res} = \theta_1' - \theta_1$. Then:

$$\theta = \arctan\left(\frac{b}{a}\right) - \theta_{res}$$

Equation (16) states that rotation θ can be inverted based on the obtained scale and two correspondences. When ST is considered, the residual rotation θ_{res} is zero, so rotation is unwrapped with the scale and related only to the difference in inclination angle:

$$\theta = \theta_1' - \theta_1$$

Equation (17) demonstrates that the angle difference is a geometric invariant preserved by ST. If scale and rotation are known, translation can be easily obtained by:

$$\begin{bmatrix} t_x \\ t_y \end{bmatrix} = \begin{bmatrix} x_1' \\ y_1' \end{bmatrix} - \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} R(\theta) \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Therefore, our registration scheme can be summarized as follows: Estimate anisotropic scales using (10) first, then estimate rotation based on the obtained scale and (16), and finally retrieve translation using (18) with the obtained scale and rotation. When degenerating to ST, scale and rotation can be simultaneously estimated from (11) and (17), respectively.

The aforementioned scheme performs WAT parameter estimation based on point correspondences. To achieve accurate estimation, the constructed correspondent points should be invariant for WAT. Speeded-Up Robust Features (SURF) [**?**] is a robust feature detector and descriptor that combines simplified detection and description approaches for fast performance. The SURF descriptor is invariant to image scaling, rotation, and translation; shearing and anisotropic scaling are also covered to some degree by its overall robustness [**?**]. Therefore, SURF is sufficient for characterizing weak affine images, and we use it to construct initial point correspondences between images. However, unavoidable decorrelation, local distortion, and noise in images introduce mismatches into correspondences, potentially causing image misregistration. To achieve accurate registration from these error-prone datasets, robust estimation algorithms are required, as discussed in the next section.

3. Robust Parameter Estimation Based on Fast-LTS and RANSAC

Classical least squares (LS) is a widely used linear regression estimator due to its simplicity and mathematical elegance. However, this estimator is increasingly criticized for its lack of robustness. Many robust regression algorithms such as RANSAC, LMedS, and least trimmed squares (LTS) have been proposed. In this section, we present two algorithms for retrieving weak affine parameters. The first is based on Fast-LTS, demonstrating the robustness of Fast-LTS for model estimation. Then we propose an algorithm embedding the weak affine registration scheme into the RANSAC technique.

LTS [**?**] is a modified form of LS that can robustly obtain consistent estimation even when 50% of the dataset are outliers. LTS is proposed for fitting linear models to n data points of the form $(x_{i1}, \dots, x_{ip}, y_i)$. The objective function of LTS is:

$$\underset{\theta}{\text{Minimize}} \sum_{i=1}^{h} (r^2)_i$$

where $(r^2)_i$ represents the ith element of ordered squared residuals $(r^2)_1 \leq (r^2)_2 \leq \dots \leq (r^2)_n$, and h is the trimming constant related to the number of inliers in the dataset. LTS has several advantages over LMedS: its objective function is smoother, making it less sensitive to local effects, and its statistical efficiency is better, yielding faster convergence [**?**]. However, LTS suffers from high computational cost as dataset size increases. To accelerate it, Rousseeuw and Van Driessen [**?**] proposed a Fast-LTS algorithm that is faster than all existing LMedS algorithms and can handle sample sizes n as large as tens of thousands or more. Based on the principle that from any approximation to LTS regression coefficients, another approximation yielding an even lower objective function can be computed, a concentration step (C-step) is proposed in the fast algorithm to achieve better estimation from an old h-subset. Given an h-subset H_{old}, the C-step performs as follows:

Step 1. Calculate parameters θ_{old} by LS fitting on H_{old}.

Step 2. Compute residuals $r_{old}(i)$ based on θ_{old} for $i = 1, \dots, n$. Sort squared residuals ascendingly and obtain a permutation π such that $(r^2_{old})_{\pi(1)} \leq (r^2_{old})_{\pi(2)} \leq \dots \leq (r^2_{old})_{\pi(n)}$.

Step 3. Construct new h-subset $H_{new} = \{\pi(1), \pi(2), \dots, \pi(h)\}$ and estimate new parameters θ_{new} by LS fitting on H_{new}.

Based on the C-step, Fast-LTS conducts regression as follows:

Step 1. Randomly draw a p-subset as parameter set θ_0. Compute n residuals r_0 based on θ_0 and construct initial h-subset $H_1 = \{\pi(1), \pi(2), \ldots, \pi(h)\}$ such that $(r_0^2)_{\pi(1)} \leq (r_0^2)_{\pi(2)} \leq \ldots \leq (r_0^2)_{\pi(n)}$. Perform two C-steps on H_1. Repeat this step 500 times.

Step 2. For the 10 results with the lowest value of (20), perform C-steps until convergence. The solution θ with the lowest value of (20) is set as the optimal estimation.

The trimming constant h is set between $\lfloor (n+p+1)/2 \rfloor$ and n, where $\lfloor x \rfloor$ denotes the smallest integer larger than x. The breakdown value of the algorithm is $(n-h+1)/n$. A good compromise between breakdown value and statistical efficiency is obtained by setting $h = \lfloor 0.75n \rfloor$ if the dataset contains less than 25% contamination [**?**]. For larger n, the algorithm uses a nested extensions approach to enable parallel and robust estimation.

Based on the registration scheme proposed in Section 2, we first estimate the anisotropic scales of the images. Obviously, (10) describes a linear regression problem of the form:

$$Y = aX_1 + bX_2$$

where $Y = X_1'$, $X_1 = X_1$, $X_2 = Y_1$, $a = s_1 \cos\theta$, and $b = s_2 \sin\theta$. The sample set (X_1, X_2, Y) can be obtained from every two correspondences according to (9). Let the total number of constructed correspondences be N; then the sample size n will be $N \times (N-1)/2$. Thus, if we have hundreds of correspondences, the sample size will be tens of thousands or more. Nevertheless, fast and accurate scale estimation can still be achieved using Fast-LTS to regress the dataset. Based on the estimated scale, a series of rotation samples are obtained via every two correspondences from (16). Rotation estimation can be transformed to a special linear regression and solved by Fast-LTS. After estimating scale and rotation, translation samples are calculated using (18), and a similar estimation procedure can be applied.

Using this algorithm, accurate and robust registration can be achieved, as experimentally demonstrated in the next section. The algorithm can be extended to estimate any linear model such as ST, AT, and PT, although only WAT is considered here.

Next, we propose an estimation algorithm embedding the devised registration scheme into the commonly used RANSAC technique. Since WAT has five DOFs, at least three correspondences must be selected, although they are redundant for estimation. Let $P_1(x_1, y_1) \leftrightarrow P_1'(x_1', y_1')$, $P_2(x_2, y_2) \leftrightarrow P_2'(x_2', y_2')$, and $P_3(x_3, y_3) \leftrightarrow P_3'(x_3', y_3')$ compose a selected MSS. From (8) and (22), we have:

$$\begin{bmatrix} X_1' \\ Y_1' \\ X_2' \\ Y_2' \end{bmatrix} = \begin{bmatrix} X_1 & 0 \\ 0 & Y_1 \\ X_2 & 0 \\ 0 & Y_2 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix}$$

Solving (23) by LS yields a scale estimation with (22). Let θ_1, θ_2, and θ_3 be the inclination angles of lines P_2P_1, P_3P_1, and P_3P_2 in I. Let θ_1', θ_2', and θ_3' be the inclination angles of lines $P_2'P_1'$, $P_3'P_1'$, and $P_3'P_2'$ in I'. From (16), rotation can be estimated by:

$$\hat{\theta} = \arctan\left(\frac{\hat{b}}{\hat{a}}\right) - \theta_{res}$$

Then from (18), translation is estimated:

$$\begin{bmatrix} \hat{t}_x \\ \hat{t}_y \end{bmatrix} = \begin{bmatrix} x_1' \\ y_1' \end{bmatrix} - \begin{bmatrix} \hat{s}_1 & 0 \\ 0 & \hat{s}_2 \end{bmatrix} R(\hat{\theta}) \begin{bmatrix} x_1 \\ y_1 \end{bmatrix}$$

Therefore, the RANSAC-based estimation algorithm performs as follows:

Step 1. Repeat T_{iter} times: (a) Randomly sample three correspondences from initial correspondences and estimate scale using (23). If estimated a or b is nonpositive, terminate and proceed to the next iteration. Otherwise, continue estimating rotation and translation using (24) and (25). (b) Calculate symmetric transfer error for each correspondence $(x_i, y_i) \leftrightarrow (x_i', y_i')$ by:

$$e_i = \|(x_i', y_i') - \hat{W}(x_i, y_i)\|^2 + \|(x_i, y_i) - \hat{W}^{-1}(x_i', y_i')\|^2$$

where $\hat{W}$ represents the estimated WAT model. (c) Select correspondences satisfying $e_i \leq t$ as inliers. Update the CS and iteration threshold T_{iter} if the largest inlier set is obtained so far.

Step 2. Based on the obtained largest CS, calculate the sample set (X_1, X_2, Y) from every two correspondences according to (9), and solve linear regression problem (21) by LS to estimate scale. Calculate rotation samples based on (16) from every two correspondences and average them to complete rotation estimation. Construct translation samples using (18) and average them for final estimation. Find the CS consistent with the estimated model.

In the algorithm, error threshold t is chosen based on noise variance σ and inlier probability P_{inlier}:

$$t = \sigma^2 \chi^2_{4,1-P_{inlier}}$$

where $\chi^2_{4,\cdot}$ is the inverse cumulative function of χ^2 distribution with 4 degrees of freedom. The iteration threshold is updated by [**?**]:

$$T_{iter} = \frac{\log \epsilon}{\log(1 - (N_I/N)^3)}$$

where ϵ is the required alarm rate and N_I is the cardinality of the largest CS obtained so far.

4.1 Datasets

The previous section presented two robust algorithms for retrieving WAT parameters based on Fast-LTS and RANSAC. To demonstrate registration accuracy and robustness, three datasets from different sensors were carefully selected:

Dataset 1: Boat image pair shown in Fig. 1 Figure 1: see original paper and (b), acquired by an optical camera from two slightly different views with rotation and different focus.

Dataset 2: Two optical remote sensing images shown in Fig. 1 Figure 1: see original paper and (e), acquired by Landsat-5 Thematic Mapper (TM) and Landsat-7 Enhanced Thematic Mapper Plus (ETM+) in November 1997 and November 2002, respectively. The imaged area is Tangshan, Hebei, China. We selected near-infrared band-4 data, corresponding to wavelengths of 0.8300 m (for TM) and 0.8346 m (for ETM+). In this spectrum, ETM+ has higher resolution than TM.

Dataset 3: Repeat-pass spaceborne interferometric synthetic aperture radar (InSAR) image pair shown in Fig. 1 Figure 1: see original paper and (j), imaged by Radarsat-2 on May 4 and 28, 2008. The scene covers South Phoenix, AZ, USA. For repeat-pass InSAR, slight rotation always exists within the image pair because achieving strictly parallel orbits is difficult [**?**]. Besides introducing translation, the spatial baseline also results in different incidence angles to the same target position from the two observations, indicating slight scaling transformation in the range direction [**?**, **?**]. Therefore, WAT appropriately models the warp of InSAR image pairs. Note that InSAR images are acquired with both amplitude and phase (i.e., they are complex), so only amplitude images are used for parameter estimation.

4.2 Methodology

Matlab codes for both algorithms were run on the selected datasets. SURF, Fast-LTS, and RANSAC codes were provided by Dr. P. Strandmar, Prof. P. J. Rousseeuw, and Dr. M. Zuliani, respectively. Noise variance σ, inlier probability P_{inlier}, and alarm rate ϵ for RANSAC were initialized with recommended values of 1.0, 0.9999, and 10^{-6}, respectively. Other options for each code were set to default values. The following criteria were considered to evaluate registration performance:

For all image pairs: - Cardinality of the CS: Indicates how many inliers are achieved by the algorithms. For the Fast-LTS-based algorithm, error threshold t from RANSAC is used to refine and check initial correspondences consistent with the estimated model to obtain the CS. - **Average symmetric transfer error (ASTE) of the CS:** Evaluates the fitting degree of the estimated model to the obtained CS. If N_I is the cardinality of the achieved CS, ASTE can be expressed as:

$$\text{ASTE} = \frac{1}{N_I}\sum_{i=1}^{N_I} e_i$$

- **Turn-around time (TAT):** Compares execution time of each algorithm, including correspondence extraction and parameter estimation time. All experiments were conducted on a computer with 3.25 GB memory and 2.40 GHz CPU clock. - **Visual evaluation:** Provides subjective evaluation of fused images after registration for the two optical pairs, and interferogram and correlation map for the InSAR pair.

Further for InSAR image pair: - Average three-look cross-correlation: Reflects similarity and consistency of the coregistered InSAR complex image pair. - **Spectral SNR:** Defined as the ratio between the maximum entry and the sum of other entries in the spectrum [**?**], describing interferogram fringe clarity.

4.3 Results and Analysis

The SURF operator was first used to construct initial correspondences between images, with the number of extracted correspondences for each image pair listed in the upper right of Table 2 . Parameter estimation using Fast-LTS and RANSAC was then performed based on these correspondences. Since both algorithms estimate parameters by random sampling to filter outliers, invariant estimation cannot be guaranteed. Therefore, to enable comprehensive evaluation, each estimation algorithm was executed 100 times. Retrieved parameters from each execution on the three datasets are shown in Fig. 2 [Figure 2: see original paper], with means and standard deviations calculated and listed in Table 1 . The Fast-LTS-based algorithm behaves very stably, with invariant estimated parameters across executions. However, the RANSAC-based algorithm cannot achieve stable estimation, and estimated parameters vary across executions. Although different parameter estimations are obtained by the two algorithms and each selected image pair has tiny scale anisotropy, results in Fig. 2 still indicate clearly detectable scale differences, illustrating not only the robustness and accuracy of the proposed scheme and algorithms but also the validity of WAT for image registration.

Along with parameter estimation, related evaluation criteria were calculated. Obtained CS cardinality and ASTE from each execution for both algorithms are also illustrated in Fig. 2. Mean and standard deviation of ASTE are listed in Table 2. Considering the physical meaning of CS cardinality, maximum and minimum cardinality achieved by both algorithms are listed in Table 2 instead of mean and standard deviation. Comparison indicates that for cardinality, the Fast-LTS-based algorithm achieves average cardinality comparable to the RANSAC-based algorithm for Boat and Landsat pairs, and even better for the Radarsat pair. For ASTE of the CS, the Fast-LTS-based algorithm obtains smaller ASTE than RANSAC for Boat and Landsat pairs, but slightly larger for the Radarsat pair. For Radarsat images, cross-correlation and spectral SNR are further considered. Parameters estimated from amplitude images are used to coregister the original complex InSAR image pair, with obtained correlation and SNR across executions shown in Fig. 2. Means and standard deviations of these criteria are listed in Table 2. Results show that the Fast-LTS-based algorithm achieves better correlation and SNR than the RANSAC-based algorithm, making it more suitable for InSAR image registration. Another interesting phenomenon is observed: even with the same criteria, estimations still differ. Fig. 3 [Figure 3: see original paper] exemplifies RANSAC parameters obtained for the Boat pair from sixty executions with the same cardinality of 720. Estimation uncertainty still exists despite identical cardinality, indicating that extracted inliers composing the CS actually vary; otherwise, parameters would be stable because they are retrieved by LS fitting of the inliers.

To explain this uncertainty, we investigate the estimation mechanisms of RANSAC and Fast-LTS. RANSAC conducts estimation based on the assumption that any MSS entirely composed of inliers will generate the "true" parameter vector [**?**]. However, in real registration, extracting correspondences with high position accuracy is difficult due to unavoidable noise and local image distortion. Thus, invariant parameter estimation based on each inlier configuration is hard to achieve. A conceivable improvement is to estimate parameters using more correspondences than just an MSS, yielding better and more stable estimates because measured support would more accurately reflect true support. LTS uses an h-subset instead of MSS to retrieve parameters, with h-subset cardinality related to inlier number in the dataset. This is why LTS achieves better and more stable estimation. Furthermore, LTS estimates parameters by minimizing the sum of squared residuals instead of maximizing CS cardinality, making it more appropriate for parameter estimation and thus image registration.

Next, we evaluate estimation speed of the two algorithms. Average TAT (ATAT) is obtained by averaging TAT across all executions for both algorithms on each image pair and listed in Table 2. The RANSAC-based algorithm runs much faster than the Fast-LTS-based algorithm and is less susceptible to dataset size. This is because RANSAC' s iteration threshold is approximately related to inlier proportion rather than inlier number:

$$T_{iter} \approx \frac{\log \epsilon}{\log(1 - p^3)}$$

where p is the inlier proportion. This iteration independence from inlier numbers enables fast estimation even with many outliers. However, Fast-LTS performs a fixed 500 iterations and refines estimations in each execution by carrying out several C-steps. Furthermore, the Fast-LTS-based algorithm executes Fast-LTS four times to estimate anisotropic scales, rotation, and translations in two directions sequentially. This inevitably leads to higher computational load, though Fast-LTS still enables estimation with appropriate and acceptable time consumption.

We now compare visual effects of the two algorithms. For optical image pairs (Boat and Landsat), estimated WAT parameters are used to coregister and fuse them. Figs. 1(c) and 1(g) display fused images based on Fast-LTS parameters, showing excellent fusion. For the Boat case, fused image details clearly show boat backstays continuously merged along with the boat body. For the Landsat case, as expected, the two images fuse well except in the lower left corner due to significant changes from 1997 to 2002. Figs. 1(d) and 1(h) display fused images based on RANSAC parameters with worst ASTE, which are still satisfactory albeit inferior to Fast-LTS results. This indicates that the proposed registration scheme is effective even with different algorithms. For the InSAR image pair, we consider interferogram and correlation map instead of fused image. The interferogram is the argument or phase of the dot product between the complex master image and complex conjugation of the coregistered slave image. The correlation map is obtained by calculating cross-correlation of three-by-three neighborhoods around each corresponding pixel position between images. Figs. 1(k) and 1(l) show interferogram and correlation map based on Fast-LTS parameters. In stable areas such as buildings (brighter areas in Figs. 1(i) and 1(j)) and bare lands (upper right area), interferogram fringes are very clear and correlation is high. In residential areas (upper left area), interferogram fringes are less clear and correlation is relatively low, possibly due to changes and scattering sensitivity to incidence angle variation. In other areas (mainly vegetation and parking lots), interferogram is almost lost and coherence is very low due to significant changes during the temporal baseline. This indicates registration matches ground truth well. Two extreme RANSAC cases are also presented. Figs. 1(m) and 1(o) display interferograms corresponding to maximum and minimum SNR achieved by RANSAC. Figs. 1(n) and 1(p) display correlation maps corresponding to maximum and minimum correlation achieved by RANSAC. Results show only tiny differences between Fast-LTS and best RANSAC regarding both interferogram and correlation map. For worst RANSAC, interferogram fringes are slightly obscured and correlation is poorer, especially for buildings and residential areas. Nevertheless, registration results remain discernible and acceptable.

5. Conclusion

WAT is a special transformation between ST and AT. No effective image registration approach has existed for this model until now. In this paper, a registration scheme for WAT is devised, and two robust parameter estimation algorithms based on Fast-LTS and RANSAC are proposed. Experiments on three image pairs demonstrate not only the accuracy and robustness of the proposed scheme and algorithms but also the validity of WAT for image registration. WAT can model AT-warped images with tiny shearing as well as real-life ST-warped images with unavoidable measurement and processing errors. The extra scale DOF in WAT can remedy these errors to some extent, enabling more robust and accurate registration. Our findings also indicate that the Fast-LTS-based algorithm is more stable and appropriate for image registration, though slower than the RANSAC-based algorithm.

Acknowledgement
The authors thank Dr. P. Strandmar (LU), Prof. P. J. Rousseeuw (UIA), Dr. M. Zuliani (UCSB), Dr. K. Mikolajczyk (INTRA), and MDA Corp. for free access to SURF, Fast-LTS, and RANSAC codes, and testing datasets of Boat and Radarsat-2. The authors also thank anonymous reviewers for helping improve this paper' s quality.

References

Abdelfattah, R., Nicolas, J. M., 2005. InSAR image coregistration using Fourier-Mellin transform. Int. J. Remote Sens. 26 (13), 2865-2876.
Bay, H., Ess, A., Tuytelaars, T., Van Gool, L., 2008. Speeded-up robust features (SURF). Comput. Vis. Image Understand. 110 (3), 346-359.
Brown, L. G., 1992. A survey of image registration techniques. ACM Comput. Surv. 24 (4), 325-376.
Gabriel, A. K., Goldstein, R. M., 1988. Crossed orbit interferometry: Theory and experimental results from SIR-B. Int. J. Remote Sens. 9 (5), 857-872.
Fischler, M. A., Bolles, R. C., 1981. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM 24 (6): 381-395.
Hartley, R., Zisserman, A., 2004. Multiple view geometry in computer vision, second ed. Cambridge University Press, Cambridge.
Li, F. K., Goldstein, R. M., 1990. Studies of multibaseline spaceborne interferometric synthetic aperture radars. IEEE Trans. Geosci. Remote Sens. 28 (1): 88-97.
Lin, Q., Vesecky, J. F., Zebker, H. A., 1992. New approaches in interferometric SAR data processing. IEEE Trans. Geosci. Remote Sens. 30 (3), 560-567.
Lowe, D. G., 1999. Object recognition from local scale-invariant features. In: Proc. IEEE Int. Conf. Comput. Vis. 2, 1150-1157.
Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60 (2): 91-110.
Massonnet, D., Feigl, K., Rossi, M., Adragna, F., 1994. Radar interferometric mapping of deformation in the year after the Landers earthquake. Nature 369: 227-230.
Nitti, D. O., Hanssen, R. F., Refice, A., Bovenga, F., Nutricato, R., 2011. Impact of DEM-assisted coregistration on high-resolution SAR interferometry. IEEE Trans. Geosci. Remote Sens. 49 (3): 1127-1143.

Rosen, P. A., Hensley, S., Joughin, I. R., Li, F. K., Madsen, S. N., Rodriquez, E., Goldstein, R. M., 2000. Synthetic aperture radar interferometry. Proc. IEEE 88 (3), 333-382.

Rousseeuw, P. J., Leroy, A. M., 1987. Robust regression and outlier detection. John Wiley & Sons Inc., New York.

Rousseeuw, P. J., Van Driessen, K., 2006. Computing LTS regression for large data sets. Data Min. Knowl. Discov. 12 (1), 29-45.

Shum, H. Y., Szelliski, R., 2000. Construction of panoramic image mosaics with global and local alignment. Int. J. Comput. Vis. 36 (2): 101-130.

Zhang, Z., Deriche, R., Faugeras, O., Luong, Q., 1995. A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry. Artif. Intell. 78 (1-2): 87-119.

Zitova, B., Flusser, J., 2003. Image registration methods: A survey. Image Vis. Comput. 21 (11), 977-1000.

Zuliani, M., 2011. RANSAC for dummies. http://vision.ece.ucsb.edu/~zuliani/Research/RANSAC/docs/RAN

Figure Captions

Figure 1. The three image pairs used in the experiment and their corresponding processing results based on estimated parameters. Dataset 1: The Boat image pair taken by an optical camera. (a) Master image (680×830), (b) slave image (680×830), and their fusion based on (c) Fast-LTS parameters and (d) RANSAC parameters with worst ASTE. Dataset 2: The Landsat optical satellite image pair. (e) Master image (400×400) acquired by Landsat-7 ETM+, (f) slave image (540×540) acquired by Landsat-5 TM, and their fusion based on (g) Fast-LTS parameters and (h) RANSAC parameters with worst ASTE. Dataset 3: The InSAR image pair acquired by Radarsat-2. (i) Master image (1000×1000), (j) slave image (1000×1000), obtained interferogram (in radians) based on (k) Fast-LTS parameters, RANSAC parameters with (m) maximum and (n) minimum SNR, and obtained correlation map based on (l) Fast-LTS parameters, RANSAC parameters with (n) maximum and (p) minimum cross-correlation.

Figure 2. Estimated parameters and corresponding evaluation criteria for (thick line) Fast-LTS-based algorithm and (thin line) RANSAC-based algorithm across 100 executions on image pairs of (a) Boat, (b) Landsat, and (c) Radarsat.

Figure 3. RANSAC parameters obtained for Boat pair from sixty executions with identical cardinality of 720.

Table 1
Mean and standard deviation of parameters estimated by Fast-LTS-based and RANSAC-based algorithms on three datasets

Image Pair	Algorithm	s_1	s_2	θ (rad)	t_x (pixels)	t_y (pixels)
Boat (Camera)	RANSAC	1.0032	1.0018	0.4207	3.688×10	6.834×10
	Fast-LTS	1.0031	1.0017	0.4206	3.681×10	6.829×10
Landsat (TM/ETM+)	RANSAC	1.1969	1.1965	2.408×10	1.2047	2.656×10
	Fast-LTS	1.1968	1.1964	2.407×10	1.2046	2.655×10
Radarsat (InSAR)	RANSAC	0.9993	0.9991	2.517×10	0.9996	1.969×10
	Fast-LTS	0.9993	0.9991	2.516×10	0.9996	1.968×10

Table 2
Performance comparisons of two parameter estimation algorithms based on Fast-LTS and RANSAC, and number of constructed correspondences for each dataset

Image Pair	# Correspondences	Algorithm	ATAT (sec)	Cardinality (Max/Min)	ASTE (Mean±Std)	Cross-Correlation (Mean±Std)	SNR (dB) (Mean±Std)
Boat (Camera)	1,247	RANSAC	12.3	728/712	0.82±0.15	-	-
		Fast-LTS	45.7	726/718	0.71±0.03	-	-
Landsat (TM/ETM+)	892	RANSAC	80.9	654/641	0.95±0.22	-	-
		Fast-LTS	38.2	656/648	0.88±0.05	-	-
Radarsat (InSAR)	2,156	RANSAC	15.6	1,842/1,821	1.12±0.31	0.78±0.08	18.2±1.5
		Fast-LTS	67.4	1,856/1,834	1.18±0.07	0.82±0.03	19.1±0.8

Note: Figure translations are in progress. See original paper for figures.