Images with Uncertainty: Efficient Algorithms for Shift, Rotation, Scaling, and Referencing, and Their Applications to Geosciences

C. G. Schiek[1], R. Araiza[1], J. M. Hurtado[1], A. A. Velasco[1], V. Kreinovich[1], and V. Sinyansky[2]

[1] University of Texas at El Paso, USA schiek@geo.utep.edu, raraiza@utep.edu, hurtado@geo.utep.edu, velasco@geo.utep.edu, vladik@utep.edu (contact author)
[2] Raytheon Victor_Sinyansky@raytheon.com

1 Introduction

In geosciences, we often need to combine two (or more) images of the same area:

- different images bring different information; so, to get a better understanding, we must *fuse* the corresponding data; e.g., we must combine a satellite images with a radar image;
- comparison of two images – e.g., images made at different moments of time – can also give us information about the changes: e.g., by comparing pre- and post-earthquake images, we can determine the effect of the earthquake.

Compared images are often obtained from slightly different angles, from slightly different positions. Therefore, in order to compare these images, we must first *reference* them, i.e., find the shift, rotation, and scaling after which these images match the best, and then apply these transformations to the original images.

There exist efficient algorithms for referencing and for the corresponding transformations; these techniques are described in Section 2. However, these algorithms are only effective when we know these images with high accuracy. In many real-life situation – e.g., when comparing pre- and post-earthquake image, see Section 3 – the uncertainty with which we determine these images is of the same order of magnitude as the desired difference between the compared images. In Section 4 of this paper, we describe how we can generalize the existing techniques to the case of uncertain images, and how the resulting algorithms can be efficiently applied in geosciences.

2 Referencing Images Known With High Accuracy

Motivation

In order to adequately process satellite and radar information, it is necessary to find the exact correspondence between different types of images and between these images and the existing maps. In other words, we need to reference these images. There exist automatic methods of referencing satellite images. These methods are based on using Fast Fourier Transform (FFT). They work well because different image of the same area differ mainly by a shift and/or by a rotation, and so, their Fourier transforms are related in a known way, from which we can determine the exact rotation and shift.

Automation is necessary

At present, referencing is mostly done semi-automatically: once we find the matching points on the two images, we can use imaging tools to find the most appropriate transformation (rotation and/or shift) which maps one image into another. The problem is that finding such matching points is a difficult and time-consuming tasks, especially for images of the Southwest.

The most efficient way is to match road intersections. Many nearby road intersections look similar, so we need several trial-and-error iterations before we can get a good referencing. Even an experienced imaging specialist must spend at least an hour or so on referencing an image. Since new satellite images are produced every few seconds, we cannot afford to spend an hour of referencing each new image. We need automatic referencing techniques.

The existing FFT-based referencing algorithms

To decrease the referencing time, researchers have proposed methods based on Fast Fourier Transform (FFT). The best of known FFT-based referencing algorithms is presented in [3, 4, 10, 11, 12, 20, 23, 24, 29, 30]. The main ideas behind FFT-based referencing in general and this algorithm in particular are as follows.

The simplest case: shift detection in the absence of noise

Let us first consider the case when two images differ only by shift. It is known that if two images $I(\mathbf{x})$ and $I'(\mathbf{x})$ differ only by shift, i.e., if $I'(\mathbf{x}) = I(\mathbf{x} + \mathbf{a})$ for some (unknown) shift $\mathbf{a}$, then their Fourier transforms

$$F(\boldsymbol{\omega}) = \frac{1}{2\pi} \cdot \int \int I(\mathbf{x}) \cdot e^{-2\pi \cdot \mathrm{i} \cdot (\mathbf{x} \cdot \boldsymbol{\omega})} \, \mathrm{d}x \mathrm{d}y,$$

$$F'(\boldsymbol{\omega}) = \frac{1}{2\pi} \cdot \int \int I'(\mathbf{x}) \cdot e^{-2\pi \cdot \mathrm{i} \cdot (\mathbf{x} \cdot \boldsymbol{\omega})} \, \mathrm{d}x \mathrm{d}y,$$

are related by the following formula:

$$F'(\boldsymbol{\omega}) = e^{2\pi \cdot \mathrm{i} \cdot (\boldsymbol{\omega} \cdot \mathbf{a})} \cdot F(\boldsymbol{\omega}). \tag{1}$$

Therefore, if the images are indeed obtained from each other by shift, then we have

$$M'(\boldsymbol{\omega}) = M(\boldsymbol{\omega}), \tag{2}$$

where we denoted

$$M(\boldsymbol{\omega}) = |F(\boldsymbol{\omega})|, \quad M'(\boldsymbol{\omega}) = |F'(\boldsymbol{\omega})|. \tag{3}$$

The actual value of the shift $\mathbf{a}$ can be obtained if we use the formula (1) to compute the value of the following ratio:

$$R(\boldsymbol{\omega}) = \frac{F'(\boldsymbol{\omega})}{F(\boldsymbol{\omega})}. \tag{4}$$

Substituting (1) into (4), we get

$$R(\boldsymbol{\omega}) = e^{2\pi \cdot \mathrm{i} \cdot (\boldsymbol{\omega} \cdot \mathbf{a})}. \tag{5}$$

Therefore, the inverse Fourier transform $P(\mathbf{x})$ of this ratio is equal to the delta-function $\delta(\mathbf{x} - \mathbf{a})$.

In other words, in the ideal no-noise situation, this inverse Fourier transform $P(\mathbf{x})$ is equal to 0 everywhere except for the point $\mathbf{x} = \mathbf{a}$; so, from $P(\mathbf{x})$, we can easily determine the desired shift by using the following *algorithm*:

- *first, we apply FFT to the original images $I(\mathbf{x})$ and $I'(\mathbf{x})$ and compute their Fourier transforms $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$;*
- *on the second step, we compute the ratio (4);*
- *on the third step, we apply the inverse FFT to the ratio $R(\boldsymbol{\omega})$ and compute its inverse Fourier transform $P(\mathbf{x})$;*
- *finally, on the fourth step, we determine the desired shift $\mathbf{a}$ as the only value $\mathbf{a}$ for which $P(\mathbf{a}) \neq 0$.*

Shift detection in the presence of noise

In the ideal case, the absolute value of the ratio (4) is equal to 1. In real life, the measured intensity values have some noise in them. For example, the conditions may slightly change from one overflight to another, which can be represented as the fact that a "noise" was added to the actual image.

In the presence of noise, the observed values of the intensities may differ from the actual values; as a result, their Fourier transforms also differ from the values and hence, the absolute value of the ratio (4) may be different from 1.

We can somewhat improve the accuracy of this method if, instead of simply processing the measurement results, we take into consideration the additional knowledge that the absolute value of the actual ratio (4) is exactly equal to 1. Let us see how this can be done.

Let us denote the actual (unknown) value of the value $e^{2\pi\cdot\mathrm{i}\cdot(\boldsymbol{\omega}\cdot\mathbf{a})}$ by r. Then, in the absence of noise, the equation (1) takes the form

$$F'(\boldsymbol{\omega}) = r \cdot F(\boldsymbol{\omega}). \tag{5}$$

In the presence of noise, the computed values $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$ of the Fourier transforms can be slightly different from the actual values, and therefore, the equality (5) is only approximately true:

$$F'(\boldsymbol{\omega}) \approx r \cdot F(\boldsymbol{\omega}). \tag{6}$$

In addition to the equation (6), we know that the absolute value of r is equal to 1, i.e., that

$$|r|^2 = r \cdot r^* = 1, \tag{7}$$

where r^* denotes a complex conjugate to r.

As a result, we know two things about the unknown value r:

- that r satisfies the approximate equation (6), and
- that r satisfies the additional constraint (7).

We would like to get the best estimate for r among all estimates which satisfy the condition (7). To get the optimal estimate, we can use the *Least Squares Method* (LSM). According to this method, for each estimate r, we define the error

$$E = F'(\boldsymbol{\omega}) - r \cdot F(\boldsymbol{\omega}) \tag{8}$$

with which the condition (6) is satisfied. Then, we find among all estimates which satisfy the additional condition (7), a value r for which the square $|E|^2 = E \cdot E^*$ of this error is the smallest possible.

The square $|E|^2$ of the error E can be reformulated as follows:

$$E \cdot E^* = (F'(\boldsymbol{\omega}) - r \cdot F(\boldsymbol{\omega})) \cdot \left({F'}^*(\boldsymbol{\omega}) - r^* \cdot F^*(\boldsymbol{\omega})\right) =$$

$$F'(\boldsymbol{\omega}) \cdot {F'}^*(\boldsymbol{\omega}) - r^* \cdot F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega}) - r \cdot F(\boldsymbol{\omega}) \cdot {F'}^*(\boldsymbol{\omega}) + r \cdot r^* \cdot F(\boldsymbol{\omega}) \cdot F^*(\boldsymbol{\omega}). \tag{9}$$

We need to minimize this expression under the condition (7).

For conditional minimization, there is a known technique of Lagrange multipliers, according to which the minimum of a function $f(x)$ under the condition $g(x) = 0$ is attained when for some real number λ, the auxiliary function $f(x) + \lambda \cdot g(x)$ attains its unconditional minimum; this value λ is called a *Lagrange multiplier*.

For our problem, the Lagrange multiplier technique leads to the following unconditional minimization problem:

$$F'(\boldsymbol{\omega}) \cdot {F'}^*(\boldsymbol{\omega}) - r^* \cdot F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega}) -$$

$$r \cdot F(\boldsymbol{\omega}) \cdot {F'}^*(\boldsymbol{\omega}) + r \cdot r^* \cdot F(\boldsymbol{\omega}) \cdot F^*(\boldsymbol{\omega}) + \lambda \cdot (r \cdot r^* - 1) \to \min. \tag{10}$$

We want to find the value of the complex variable r for which this expression takes the smallest possible value. A complex variable is, in effect, a pair of two real variables, so the minimum can be found as a point at which the partial derivatives with respect to each of these variables are both equal to 0. Alternatively, we can represent this equality by computing the partial derivative of the expression (10) relative to r and r^*. If we differentiate (10) relative to r^*, we get the following linear equation:

$$-F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega}) + r \cdot F(\boldsymbol{\omega}) \cdot F^*(\boldsymbol{\omega}) + \lambda \cdot r = 0. \tag{11}$$

From this equation, we conclude that

$$r = \frac{F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})}{F(\boldsymbol{\omega}) \cdot F^*(\boldsymbol{\omega}) + \lambda}. \tag{12}$$

The coefficient λ can be now determined from the condition that the resulting value r should satisfy the equation (7). The denominator $F(\boldsymbol{\omega}) \cdot F^*(\boldsymbol{\omega}) + \lambda$ of the equation (12) is a real number, so instead of finding λ, it is sufficient to find a value of this denominator for which $|r|^2 = 1$. One can easily see that to achieve this goal, we should take, as this denominator, the absolute value of the numerator, i.e., the value

$$|F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})| = |F^*(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|. \tag{13}$$

For this choice of a denominator, the formula (11) takes the following final form:

$$r = \frac{F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})}{|F^*(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|}. \tag{14}$$

So, in the presence of noise, instead of using the exact ratio (4), we should compute, for every $\boldsymbol{\omega}$, the optimal approximation

$$R(\boldsymbol{\omega}) = \frac{F^*(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})}{|F^*(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|}. \tag{15}$$

This approximation is known as "cross-correlation power spectrum" (see, e.g., [5]).

In the ideal non-noise case, the inverse Fourier transform $P(\mathbf{x})$ of this ratio is equal to the delta-function $\delta(\mathbf{x} - \mathbf{a})$, i.e., equal to 0 everywhere except for the point $\mathbf{x} = \mathbf{a}$. In the presence of noise, we expect the values of $P(\mathbf{x})$ to be slightly different from the delta-function, but still, the value $|P(\mathbf{a})|$ should be much larger than all the other values of this function. Thus, the value of the shift can be determined as the value at which $|P(\mathbf{a})|$ is the largest.

Finding the shift with subpixel accuracy

To get a subpixel accuracy, we can use the interpolation described (and justified) in [12]. Namely:

- we find the point $\mathbf{x} = (x_1, y_1)$ for which $|P(\mathbf{x})|$ takes the largest possible value;
- then, among 4 points $(x_1 \pm 1, y_1 \pm 1)$, we select a point (x_2, y_2) for which the value $|P(x_2, y_2)|$ is the largest;
- after that, we apply the formulas

$$x = \frac{w_{x1} \cdot x_1 + w_{x2} \cdot x_2}{w_{x1} + w_{x2}}; \quad y = \frac{w_{y1} \cdot y_1 + w_{y2} \cdot y_2}{w_{y1} + w_{y2}}, \tag{16}$$

where

$$w_{xi} = |P(x_i, y_1)|^\alpha + |P(x_i, y_2)|^\alpha; \quad w_{yi} = |P(x_1, y_i)|^\alpha + |P(x_2, y_i)|^\alpha, \tag{17}$$

with $\alpha = 0.65$, to find the coordinates (x, y) of the shift.

Resulting algorithm

So, we arrive at the following *algorithm* for determining the shift $\mathbf{a}$:

- *first, we apply FFT to the original images $I(\mathbf{x})$ and $I'(\mathbf{x})$ and compute their Fourier transforms $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$;*
- *on the second step, we compute the ratio (15);*
- *on the third step, we apply the inverse FFT to the ratio $R(\boldsymbol{\omega})$ and compute its inverse Fourier transform $P(\mathbf{x})$;*
- *finally, on the fourth step, we do the following:*
  - *we find the point $\mathbf{x} = (x_1, y_1)$ for which $|P(\mathbf{x})|$ takes the largest possible value;*
  - *then, among 4 points $(x_1 \pm 1, y_1 \pm 1)$, we select a point (x_2, y_2) for which the value $|P(x_2, y_2)|$ is the largest;*
  - *after that, we apply the formulas (16) and (17) to find the coordinates (x, y) of the shift.*

Reducing rotation and scaling to shift

If, in addition to shift, we also have rotation and scaling, then the absolute values $M_i(\boldsymbol{\omega})$ of the corresponding Fourier transforms are not equal, but differ from each by the corresponding rotation and scaling.

If we go from Cartesian to polar coordinates (r, θ) in the $\boldsymbol{\omega}$-plane, then rotation by an angle θ_0 is described by a simple shift-like formula $\theta \to \theta + \theta_0$.

In these same coordinates, scaling is also simple, but not shift-like: $r \to \lambda \cdot r$. If we go to *log-polar* coordinates (ρ, θ), where $\rho = \log(r)$, then scaling also becomes shift-like: $\rho \to \rho + b$, where $b = \log(\lambda)$. So, in log-polar coordinates, both rotation and scaling are described by a shift.

How to determine rotation and scaling

In view of the above reduction, in order to determine the rotation and scaling between M and M', we can do the following:

- transform both images from the original Cartesian coordinates to log-polar coordinates;
- use the above FFT-based algorithm to determine the corresponding shift $(\theta_0, \log(\lambda))$;
- from the corresponding "shift" values, reconstruct the rotation angle θ_0 and the scaling coefficient λ.

Comment. The main computational problem with the transformation to log-polar coordinates is that we need values $M(\xi, \eta)$ on a rectangular grid in log-polar space $(\log(\rho), \theta)$, but computing $(\log(\rho), \theta)$ for the original grid points leads to points outside that grid. So, we need interpolation to find the values $M(\xi, \eta)$ on the desired grid. One possibility is to use *bilinear* interpolation. Let (x, y) be a rectangular point corresponding to the desired grid point $(\log(\rho), \theta)$, i.e.,

$$x = e^{\log(\rho)} \cdot \cos(\theta), \quad y = e^{\log(\rho)} \cdot \sin(\theta).$$

To find the value $M(x, y)$, we look at the intensities M_{jk}, $M_{j+1,k}$, $M_{j,k+1}$, and $M_{j+1,k+1}$ of the four grid points (j, k), $(j+1, k)$, $(j, k+1)$, and $(j+1, k+1)$ surrounding (x, y). Then, we can interpolate $M(x, y)$ as follows:

$$M(x, y) = (1-t) \cdot (1-u) \cdot M_{jk} +$$

$$t \cdot (1-u) \cdot M_{j+1,k} + (1-t) \cdot u \cdot M_{j,k+1} + t \cdot u \cdot M_{j+1,k+1},$$

where t is a fractional part of x and u is a fractional part of y.

Final algorithm: determining shift, rotation, and scaling

- *First, we apply FFT to the original images $I(\mathbf{x})$ and $I'(\mathbf{x})$ and compute their Fourier transforms $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$.*
- *Then, we compute the absolute values $M(\boldsymbol{\omega}) = |F(\boldsymbol{\omega})|$ and $M'(\boldsymbol{\omega}) = |F'(\boldsymbol{\omega})|$ of these Fourier transforms.*
- *By applying the above rotation and scaling detection algorithm to the functions $M(\boldsymbol{\omega})$ and $M'(\boldsymbol{\omega})$, we can determine the rotation angle θ_0 and the scaling coefficient λ.*
- *Now, we can apply the corresponding rotation and scaling to one of the original images, e.g., to the first image $I(\mathbf{x})$. As a result, we get a new image $\widetilde{I}(\mathbf{x})$.*
- *Since we rotated and re-scaled one of the images, the images $\widetilde{I}(\mathbf{x})$ and $I'(\mathbf{x})$ are already aligned in terms of rotation and scaling, and the only difference between them is in an (unknown) shift. So, we can again apply the above described FFT-based algorithm for determining shift: this time, actually to determine shift.*

As a result, we get the desired values of shift, rotation, and scaling; hence, we get the desired referencing.

Comment. Similar techniques can be applied to images in other applications areas; see, e.g., [18]; in particular, applications to pavement engineering are described in [2, 25].

3 Images Known with Uncertainty: Case Study

Terrain changes

Among the phenomena that can cause terrain changes are:

- interseismic and coseismic slip along a fault,
- glacier advance and retreat,
- soil creep, and
- landslide processes,

all of which are relevant either for the hazards they may pose or the landscape evolution processes they reflect.

Radar techniques can detect vertical terrain changes

In the past decade, interferometric synthetic aperture radar (InSAR) has become a powerful tool for monitoring such deformation and surface changes [6]. Because this tool detects displacements along the line of sight of the radar system, it is most sensitive to terrain changes due to vertical deformation, such as those associated with thrust faulting, and less sensitive to lateral deformation [28].

While InSAR has been used for studying lateral displacements, such as those due to strike-slip earthquakes [19], decorrelation problems in the near-field commonly arise. Moreover, appropriate radar data is not widely available due to the lack of synthetic aperture radar (SAR) satellites in orbit. Currently, the two best SAR satellites in operation are Radarsat and ERS-2. The cost per scene for data from these satellites can range from $950 to $3000, with Radarsat data being the most expensive.

To detect lateral terrain changes, we need satellite images

Considering the high cost and scarcity of SAR data, the scientific community has looked to other data sets with wider availability, such as the Satellite Pour l'Observation de la Terre (SPOT) optical imaging instrument [7, 9, 16, 21, 28]. Terrain changes can be monitored with optical remote sensing data using image processing algorithms that measure apparent offsets in the geographic locations of the corresponding pixels in two (or more) images of the same portion of the Earth's surface taken at different times. These inter-image pixel offsets define vectors whose orientations indicate the direction of terrain displacement and whose lengths denote the magnitude of that displacement.

Previous work with SPOT images has shown the feasibility of using optical imagery for lateral displacement change detection using Fourier-based cross-correlation (15) [7, 28]. For example, Dominguez et al. [8] were able to resolve coseismic displacement along a major thrust fault associated with the 1999 Chi Chi earthquake in the Central Range of Taiwan from SPOT images using the Fourier approach. These results have shown optical imagery to optimally work in the proximal area of lateral terrain changes, which is the regime where InSAR techniques are weakest [28].

Use of ASTER images

In [22], we have shown that a similar change detection can be obtained with Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) imagery; see, e.g., [1]. The benefits of using ASTER are:

- its dramatically lower cost compared to SPOT,
- the capability of generating Digital Elevation Maps (DEMs) from ASTER imagery [14], and
- the redundancy afforded by ASTERs stereo capability.

The latter may aid in increasing the precision of terrain change measurements made using optical image processing techniques.

Test case: the November 14, 2001 Kokoxili earthquake

The left-lateral Kunlun fault, located in northern Tibet, was chosen for this study because it experienced a large earthquake during a time period for which ASTER imagery is available. On November 14, 2001, an $M_s = 8.1$ earthquake occurred causing a 400 km-long surface rupture and as much as 16.3 m of left-lateral strike-slip [15].

This slip is most evident as offset of, and fault scarps developed on, alluvial terraces. Three time separation window cases were considered. These three test cases focus on three different segments of fault.

- Test Case 1 spans a two-year time window from March 2000 to July 2002. The image pair for this case exhibits extensive erosion due to the long time span of two years. The image pair for this case also has vastly different pointing angles.
- Test Case 2 has a time separation of two-months, from November 4, 2001 to December 22, 2001. The imagery used in this test case has 13% cloud cover, and one of the images contains snow, both of which caused decorrelation.
- Test Case 3 has a time separation of thirteen months, between October 2000 and December 2001. The image pair for this case has the least amount of preprocessing problems and the smallest pointing angle difference. Although there is some snow cover and changes in water level along the Kusai

Hu lake next to the fault, decorrelation problems were minor. This test case is the only one with a well-defined fault scarp.

For each case, the accuracy of the change detection methods was assessed by comparing the geodetic image processing results to field measurements ([15], Table 2; [13, 27]).

Topographic correction

Parallax is defined as the apparent shift in the location of an object due to changes in the observers position and, as a result, the relative geometry between the observer and the target. This concept is the same principal on which human stereoscopic vision is based: our right and left eyes view the same scene from slightly different angles, resulting in parallax that we perceive as depth. Thus, topographic parallax is created by the change in the position of a satellite as it scans the uneven surface of the Earth. It results in an apparent shift in the position of terrain features between two overlapping images taken from different angles; see, e.g., [17].

Topography can also impart apparent pixel offsets due to scale changes within the image. In an area of high relief, those parts of the image within valleys will be a greater distance from the observer than those parts on ridge tops. The variable distance between target and observer results in mountain-tops having a larger map scale than the valley bottoms, an effect that non-linearly distorts images. If left uncorrected, this distortion can create apparent pixel shifts when comparing two images.

Another apparent pixel offset due to topography stems from the calculation of geographic coordinates. The geographic coordinates of a pixel in an image are calculated from orbital and attitude data and lie on the Earths ellipsoid. Due to the fact the Earth has topography, however, the true latitude and longitude coordinates of a point on the Earths surface will be displaced from its ellipsoidal coordinates. Of the three apparent pixel offsets produced by topography, this last type can cause the greatest apparent shift [28].

All three topographic apparent pixel offsets can be minimized by *ortho-rectifying* the image with a DEM. The ortho-rectification process uses the DEM to remove topographic distortions and will re-project the ellipsoidal position of a given pixel to one that better approximates its true coordinates on the Earths surface. In our work, we have applied ortho-rectification techniques to pre-process the images before referencing them.

All images used in the Kunlun fault test cases are VNIR band 3n images from ASTER level 1B scenes, which has already been through pre-process for sensor artifacts by the ASTER corporation. This band was chosen for two reasons:

- because it has the highest resolution, and
- because a sensor model is available [31], which describes the interior and exterior orientations of the image ([17]).

The sensor model is required for the orthorectification process done using the Leica Photogrammetry Suite in 19 ERDAS IMAGINE 8.7 [31].

ASTER VNIR 3b bands were not considered in the test cases due to registration problems. Orthorectification of the 3b images in ERDAS IMAGINE was poor, and there were residual geometric pixel shifts of at least 20 pixels (210 m). The poor orthorectification can be due to either excessive pointing angle differences, or excessive variation in the viewing geometry.

The DEM used for orthorectification was the 90-m Shuttle Radar Topography Mission (SRTM; www.jpl.nasa.gov/srtm) DEM. ENVI 4.0 software was used to register the before and after images, and check the orthorectification process done in ERDAS IMAGINE.

4 How to Generalize the Existing Techniques to Images Known with Uncertainty

Problem

When we applied the above algorithm to detect the earthquake-caused shift between the pre- and post-earthquake images, we did not get any meaningful shift value. Specifically, the inverse Fourier transform $P(\mathbf{x})$ of the cross-correlation power spectrum looks random, and its maximum $\mathbf{a}$ was attained at values which are very different from the actual shift.

Analysis of the problem

In the above algorithm, for every frequency $\boldsymbol{\omega}$, we compute the complex-valued product $F(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})$ and then consider only the phase of this complex value – i.e., equivalently, the value $R(\boldsymbol{\omega})$ – in the further computations.

Due to the uncertainty with which we measure the images, the corresponding Fourier transforms $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$ are also only approximately known. So, for every $\boldsymbol{\omega}$, the product $F(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})$ is also only approximately known. How does this uncertainty translate into the uncertainty with which we know $R(\boldsymbol{\omega})$?

Let ε be the accuracy with which we know the value of the product. In general, if we multiply a value x known with uncertainty $\Delta x \approx \varepsilon$ by a number λ, the resulting new value $y = \lambda \cdot x$ is known with uncertainty $\Delta y = \lambda \cdot \Delta x = \lambda \cdot \varepsilon$. Similarly, if we divide the value x by a number λ, then the resulting new value $z = x/\lambda$ is known with uncertainty $\Delta z = \Delta x/\lambda \approx \varepsilon/\lambda$.

In our algorithm, the cross-correlation power spectrum $R(\boldsymbol{\omega})$ is obtained by dividing the product $F(\boldsymbol{\omega}) \cdot F'(\boldsymbol{\omega})$ by its absolute value $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$. Since we know the product with accuracy ε, we thus know the value $R(\boldsymbol{\omega})$ with accuracy $\varepsilon/(|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|)$.

As a result, for the frequencies $\boldsymbol{\omega}$ for which the absolute value is high, we know the corresponding value $R(\boldsymbol{\omega})$ with a reasonable accuracy. However, for the frequencies $\boldsymbol{\omega}$ for which the absolute value is low, the the corresponding value $R(\boldsymbol{\omega})$ is really inaccurate – all noise.

In the above algorithm, when we compute $P(\mathbf{x})$ as the Fourier transform of the function $R(\boldsymbol{\omega})$, we take all the values $R(\boldsymbol{\omega})$ with the same weight. In effect, we are taking the average of several values, some known with reasonable accuracy and some very inaccurate. Not surprisingly, the resulting average is very inaccurate.

For example, if we have two measurements of the same quantity whose actual value is 1.0,

- the first measurement is very accurate and results in 1.05, and
- the second measurement is way off and results in 5.61,

then when we take the average, we get $(1.05 + 5.61)/2 = 3.33$ in which the noisy values dominated the accurate ones.

How to make resulting estimate more accurate: an idea

In view of the above analysis, to make the measurements more accurate, we should:

- assign less weight to less accurate values $R(\boldsymbol{\omega})$, i.e., values for which the absolute value $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$ is small, and
- assign more weight to more accurate values $R(\boldsymbol{\omega})$, i.e., values for which the absolute value $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$ is large.

The simplest way is to assign weight 0 to less accurate measurements and weight 1 to more accurate measurements. In other words, the simplest way to implement this idea is:

- to fix some threshold, and
- for all the frequencies for which the absolute value $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$ is below this threshold, set $R(\boldsymbol{\omega})$ to 0.

A natural idea is to select, as a threshold, a certain portion of the largest (or mean) value of $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$. As a result, we arrive at the following algorithm.

Resulting algorithm: general description

To find the shift **a** *between the two images:*

- *first, we apply FFT to the original images $I(\mathbf{x})$ and $I'(\mathbf{x})$ and compute their Fourier transforms $F(\boldsymbol{\omega})$ and $F'(\boldsymbol{\omega})$;*
- *on the second step, we do the following:*
  - *we find the mean value m of the product $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})|$;*
  - *we set the threshold to be a certain portion of the mean, i.e., to $\alpha \cdot m$ for some $\alpha > 0$;*
  - *for those frequencies for which $|F(\boldsymbol{\omega})| \cdot |F'(\boldsymbol{\omega})| \geq \alpha \cdot m$, we compute the value $R(\boldsymbol{\omega})$ by using the formula (15);*

  - *for other frequencies, we set* $R(\boldsymbol{\omega}) = 0$*;*
- *on the third step, we apply the inverse FFT to the function* $R(\boldsymbol{\omega})$ *and compute its inverse Fourier transform* $P(\mathbf{x})$*;*
- *finally, on the fourth step, we determine the first approximation to the desired shift* $\mathbf{a}$ *as the point for which* $|P(\mathbf{x})|$ *takes the largest possible value, and then do the interpolation by using formulas (16)–(17).*

Empirically, the best value for the parameter α turned out to be 10^{-3}.

Resulting algorithm: details

We have mentioned that due to the inaccuracy, it is very difficult to detect the lateral shift. In general, when measurements are not very accurate, a natural way to increase the accuracy is to perform repeated measurements and then average the measurement results. With respect to images, this means that we have to consider more pixels, i.e., large parts of the compared image, corresponding to larger sliding window size.

However, the problem is that the lateral shift differs from location to location: its value decreases as we get farther away from the fault. So, when we increase the window size, instead of processing several pixels with the *same* shift (which would have helped), we instead bring together pixels corresponding to *different* values of lateral shift.

Good news is that while the *magnitude* of the lateral shift is different at different pixels, the *direction* of this shift remains largely the same. So, at the first stage of our analysis, we take a large sliding window (larger that 75×75 pixels, where 1 pixel is ≈ 15 m), and use the above algorithm to determine the *direction* of the lateral shift.

Once the direction at different locations is determined, we can now take smaller sliding windows (40×40 pixels), and determine the magnitudes of the lateral shift. The directions can also be obtained from these smaller windows, but these direction are determined from he analysis of fewer pixels and are, thus, much less accurate than the directions obtained form the analysis of a larger window. Thus, to get the best results, we combine the *direction* obtained form the analysis of a larger window with the *magnitude* obtained from the smaller window.

In other words, we need to apply the above algorithm twice:

- first, with a larger sliding window, to find the direction of the lateral shift;
- then, with a smaller sliding window, to find the shift's magnitude.

Finally, we combine the direction obtained from a larger window with the magnitude obtained from a smaller window.

5 Results

All test cases display good results in the near field of the faulted area; the accuracy with which we can determine the shift decreases as we move to distal areas.

Test Case 2 gives the best results, with a measured lateral displacement of 4.5 ± 0.4 m with left-lateral slip and an average slip direction of 270°. This magnitude is similar to the 4.6-4.8 m displacement of a gulley measured by Lin et al. [15] (site 2 in Table 2), and the sense and slip direction are consistent with the known trace and kinematics of the Kunlun fault.

Test Case 3 is fairly consistent in direction, with left lateral movement and an average slip direction of 265°. However, the magnitude obtained from this analysis, ≈ 8.4 m, is a much cruder approximation to the 5.7 m of displacement measured by Lin et al. [15] (site 3 in Table 2). This could be attributed to the long 13-month time separation window during which non-earthquake terrain change occurred, such as changes in the water level in Kusai Hu lake.

Test Case 1 results in left-lateral slip with an azimuth of 268° in the nearfield of the fault and a magnitude of ≈ 8.3 m. The sense of slip and azimuth are consistent with field observations, but assessing the accuracy of the resulting magnitude is less straightforward. The closest of the Lin et al. [15] field sites is site 7. Several offset features were measured here, with displacements ranging from 3.3 m on a road to 6.8 m on a stream channel. The latter is similar, as is the field measurement at another nearby locality, site 6, where Lin et al. [15] report 7.2 m of displacement on a stream channel.

6 Conclusions and Future Work

Main conclusion

Our results have shown that a new algorithm provide a good reconstruction of shift from the two images.

Using better DEM

Future work needed includes improving the pre-processing protocol. This improvement is needed in order to fully remove any residual apparent pixel offsets and to optimize the true pixel offsets. This can be accomplished by using a 30-m DEM instead of a 90-m DEM in the ortho-rectification process. A higher-resolution DEM can be obtained from aerial photographs or LIDAR, among other sources, but will require an investment of time and resources. By using a DEM with a higher resolution, the elevation uncertainty can be improved, thus lowering the apparent pixel offsets caused by parallax.

Comparing images of different type

Future work should also include applying the change detection procedures developed in this thesis to heterogenous input imagery, for instance, a combination of an ASTER "after" image with a Landsat TM scene or aerial photographs as "before" images. By using a heterogonous pair of input imagery, a greater number of possible image candidates can be used to do change detection. In addition, since Landsat images and aerial photographs are available for times prior to the beginning of ASTER image acquisition, using heterogeneous datasets can also lengthen the time separation windows that can be considered. This can be especially useful for monitoring terrain change due to slow processes such as glacier movement. It can also make possible the study of events that occurred before ASTER was launched.

Comparing images with gridded data sets

The algorithms desscribed in this paper should be able to detect lateral movements in any kind of registered imagery. Thus, the possibility exists to apply these methods to gridded gravity and other geophysical data.

Use redundancy of ASTER images

As an effort to improve our knowledge of ASTER attitude parameters and to minimize residual apparent pixel offsets during ortho-rectification, as well as improve the performance of the change detection techniques with ASTER data, it may be possible to exploit the redundancy in the ASTER VNIR imagery [28]. The redundancy provided by the ASTER is possible due to its stereo capability, a feature which, given two ASTER scenes, essentially provides four independent sets of images to process for terrain displacement. Given a single "before" scene and a single "after" scene, there are a total of four unique permutations of image pairs that can be used as input to a terrain change detection algorithm. All else being equal, each permutation should result in identical terrain change measurements. Differences in the estimates, however, can be reconciled by optimization of poorly-constrained parameters such as the satellite attitude (e.g., roll, pitch, and yaw).

We can also use the fact that the images are multi-spectra [3, 4] (see also Appendix).

Towards more accurate shift, rotation, and scaling

It is important to find the lateral shift between pre- and post-earthquake images. Once this relative shift is determined, it is desirable to shift one of the images to see what changes occurred.

The difficulty is that the shift is subpixel, so when we shift, we move away from the original rectangular grid: the brightness values of the first image were given on a grid, but the values of the shifted image are on the shifted points, which are in between the points of the original grid. Thus, to adequately compare the two images pixel-by-pixel, we must interpolate the brightness values from the shifted grid points (at which we know the brightnesses of the shifted first image) to the original grid point (at which we know the brightnesses of the second image).

In the above text, we used bilinear interpolation to actually perform the corresponding geometric transformation (shift, rotation, or scaling). This methods is efficient – it requires only a few computations per pixel – but because of its localized character, it is not always accurate. It is well known that the more points we use for interpolation, the better results we can achieve. Ideally, interpolation should use all the available points. Such methods have indeed been developed based on efficient FFT-based implementations of so-called chirp-z transform – a generalization of Fourier transform [26]. It is desirable to apply these methods to geosciences-related images.

Methods from [26] can perform shifts and scalings in an arbitrary rectangular grid, but efficient rotation techniques are only available for the case when we have a rectangular grid with exactly the same step sizes in two dimensions, i.e., when the grid is actually a square grid. For satellite images, it is often not the case. To handle such situations, we must thus:

- first, interpolate from the original rectangular grid to the square grid;
- then, perform the rotation in the square grid, and
- finally, interpolate the rotated image back into the original rectangular grid.

Towards explicit representation of interval and fuzzy uncertainty in images

In the current image processing, an image is represented as follows: for each pixel $\mathbf{x}$, we describe the approximate value $\widetilde{I}(\mathbf{x})$ of the brightness at this pixel. It is desirable to describe not only this approximate value, but also the accuracy with which we know this value.

For example, if for each pixel, we know the guaranteed upper bound $\Delta(\mathbf{x})$ for the inaccuracy of the corresponding brightness value, this means that at each pixel $\mathbf{x}$, the actual (unknown) value of the brightness $I(\mathbf{x})$ belongs to the interval

$$\mathbf{I}(\mathbf{x}) = [\underline{I}(\mathbf{x}), \overline{I}(\mathbf{x})] \stackrel{\text{def}}{=} [\widetilde{I}(\mathbf{x}) - \Delta(\mathbf{x}), \widetilde{I}(\mathbf{x}) + \Delta(\mathbf{x})].$$

In a more realistic situation, instead of the *guaranteed* bound, we may have different values which bound the difference $\Delta I(\mathbf{x}) \stackrel{\text{def}}{=} I(\mathbf{x}) - \widetilde{I}(\mathbf{x})$ with different *degrees of certainty.* In other words, for every pixel $\mathbf{x}$, we have nested intervals corresponding to different degrees of certainty – i.e., in effect, a fuzzy value $I(\mathbf{x})$. A fuzzy-valued image is thus simply a nested (layered) family of interval-valued images.

How can we process such interval and fuzzy images? To transform (shift, rotate, scale) an interval image $[\underline{I}(\mathbf{x}), \overline{I}(\mathbf{x})]$, it is sufficient to rotate the corresponding endpoint images $\underline{I}(\mathbf{x})$ and $\overline{I}(\mathbf{x})$. To transform a fuzzy image, it is sufficient to rotate the corresponding interval images layer-by-layer.

Acknowledgments

This work was supported in part by NASA under cooperative agreement NCC5-209, NSF grants EAR-0225670 and DMS-0532645, by Army Research Lab grant DATM-05-02-C-0046, by the Star Award from the University of Texas System, and by the Texas Department of Transportation grant No. 0-5453

References

1. Abrams M, Hook S, Ramachandran B (2002) Aster User Handbook, NASA Jet Propulsion Laboratory, Pasadena, California
2. Adidhela JE (2004) Using FFT-based Data Processing Techniques to Characterize Asphaltic Concrete Mixtures, Master Thesis, Department of Computer Science, University of Texas at El Paso
3. Araiza R (2003) Optimal FFT-Based Algorithms for Referencing Multi-Spectral Images, Master Thesis, Department of Computer Science, University of Texas at El Paso
4. Araiza R, Xie H, Starks SA, Kreinovich V (2002) Automatic Referencing of Multi-Spectral Images. In: Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, Santa Fe, New Mexico, USA, April 7–9, 2002, pp. 21–25
5. Brown LG (1992) A Survey of Image Registration Techniques. AMC Computing Surveys 24(4):325–376
6. Burgmann R, Rosen PA, Fielding EJ (2000) Synthetic Aperture Radar Interferometry to Measure Earth's Surface Topography and Its Deformation. Annu. Rev. Earth Planet. Sci. 28:169–209
7. Crippen RE, Blom RG (1991) Measurement of Subresolution Terrain Displacements using SPOT Panchromatic Imagery. In: Proceedings of the 1991 International Geoscience and Remote Sensing Symposium IGARSS'91 "Remote Sensing: Global Monitoring for Earth Management", June 3–6, 1991, Vol. 3, pp. 1667–1670
8. Dominguez S, Avouac JP, Michel R (2003) Horizontal coseismic deformation of the 1999 Chi-Chi earthquake measured from SPOT satellite images: Implications for the seismic cycle along the western foothills of central Taiwan. Journal of Geophysical Research 108(B2):2083
9. Feigl KL, Sarti F, Vadon H, McClusky S, Ergintav S, Durand P, Burgmann R, Rigo A, Massonnet D, Reilinger R (2002) Estimating Slip Distribution for the Izmit Mainshock from Coeseismic GPS, ERS-1, RADARSAT, and SPOT Measurements. Bulletin of the Seismological Society of America 92(1) 138–160
10. Garcia F, Araiza R, Rzycki B (2000) Towards optimal mosaicking of multi-spectral images. In: Proc. 2nd Nat'l NASA Student Conference, Nashville, Tennessee, April 7–10, 2000.
11. Gibson S (1999) An optimal FFT-based algorithm for mosaicking images, Master Thesis, Department of Computer Science, University of Texas at El Paso
12. Gibson S, Kreinovich V, Longpré L, Penn B, Starks SA (2001) Intelligent Mining in Image Databases, With Applications to Satellite Imaging and to Web

Search. In: Kandel A, Bunke H, Last M (eds.), Data Mining and Computational Intelligence, Springer-Verlag, Berlin, 309–336
13. King G, Klinger Y, Bowman D, Tapponnier P (2005) Slip-Partitioned Surface Breaks for the $M_w = 7.8$ 2001 Kokoxili Earthquake, China. Bulletin of the Seismological Society of America 95:731–738
14. Lang HR (1999) Algorithm Theoretical Basis Document for ASTER Digital Elevation Models (Standard Product AST14), Version 3.0: NASA Jet Propulsion Laboratory
15. Lin A, Fu B, Guo J, Qingli Z, Guangming D, He W, Zhao Y (2002) Coseismic strike-slip and rupture length produced by the 2001 $M_s = 8.1$ central Kunlun earthquake. Science 296(5575):2015–2017
16. Michel R, Avouac JP (2002) Deformation due to the 17 August 1999 Izmit, Turkey, earthquake measured from SPOT images. Journal of Geophysical Research 107(B4)
17. Mikhail EM, Bethel JS, McGlone JM (2001) Introduction to Modern Photogrammetry. New York, John Wiley & Sons, Inc.
18. Osegueda R, Keller GR, Starks SA, Araiza R, Bizyaev D, Kreinovich V (2004) Towards a General Methodology for Designing Sub-Noise Measurement Procedures. In: Proceedings of the 10th IMEKO TC7 International Symposium on Advances of Measurement Science, St. Petersburg, Russia, June 30–July 2, 2004, Vol. 1, pp. 59–64
19. Peltzer G, Crampe F, Rosen P (200) The $M_w = 7.1$, Hector Mine, California earthquake: surface rupture, surface displacement, and fault slip solution from ERS SAR data. Earth and Planetary Sciences 333:545–555
20. Reddy BS, Chatterji BN (1996) An FFT-based technique for translation, rotation, and scale-invariant image registration. IEEE Transactions on Image Processing 5(8):1266–1271
21. Riazanoff S (2002) SPOT Satellite Geometry Handbook. SPOT Image
22. Schiek CG (2004) Terrain change detection using aster optical satellite imagery along the Kunlun fault, Tibet, Master Thesis, Department of Geological Sciences, University of Texas at El Paso
23. Srikrishnan S (2002) Referencing noisy images: complexity of the problem, analysis and efficient implementation of the Fast Fourier approach, Master Thesis, Department of Computer Science, University of Texas at El Paso
24. Srikrishnan S, Araiza R, Xie H, Starks SA, Kreinovich V (2001) Automatic referencing of satellite and radar images. In: Proceedings of the 2001 IEEE Systems, Man, and Cybernetics Conference, Tucson, Arizona, October 7–10, 2001, pp. 2176–2181
25. Starks SA, Nazarian S, Kreinovich V, Adidhela J, Araiza R (2004) Using FFT-Based Data Processing Techniques to Characterize Asphaltic Concrete Mixtures. In: Proceedings of the 11th IEEE Digital Signal Processing Workshop DSP'04, Taos Ski Valley, New Mexico, August 1–4, 2004, pp. 241–245
26. Tong R, Cox RW (1999) Rotation of NMR images using the 2D chirp-z transform. Magnetic Resonance in Medicine 41:253–256
27. Van der Woerd J, Meriaux AS, Klinger Y, Ryerson FJ, Gaudemer Y, Tapponnier P (2002) The 14 November 2001, $M_w = 7.8$ Kokoxili Earthquake in Northern Tibet (Qinghai Province, China). Seismological Research Letters 73(2):125–135
28. Van Puymbroeck N, Michel R, Binet R, Avouac JP, Taboury J (2000) Measuring earthquakes from optical satellite images. Applied Optics 39(20):3486–3494

29. Xie H, Hicks N, Keller GR, Huang H, Kreinovich V (2000) Automatic image registration based on a FFT algorithm and IDL/ENVI. In: Proceedings of the ICORG-2000 International Conference on Remote Sensing and GIS/GPS, Hyderabad, India, December 1–4, 2000, Vol. 1, pp. 397–402
30. Xie H, Hicks N, Keller GR, Huang H, Kreinovich V (2003) Implementation, test, and analysis of an automatic image registration algorithm based on FFT and IDL/ENVI. Computers and Geosciences 29(8):1045–1055
31. Yu J, Cowgill E, Healy S (2003) Digital Stereomapping in Imagine Orthobase, California Institute of Technology, pp. 1–10

Appendix: Referencing Multi-Spectral Satellite Images

Formulation of the problem

With the new generation of multi-spectral satellites, for each area, we have several hundred images which correspond to different wavelengths. At present, when we reference two images, we only use one of the wavelengths and ignore the information from the other wavelengths. It is reasonable to decrease the referencing error by using images corresponding to all possible wavelengths in referencing.

Similarly, in detecting the known text in colored web images, we would like to take into consideration all color components.

In this appendix, we present an algorithm for such optimal referencing.

Derivation of the new algorithm

For multi-spectral imaging, instead of a single image $I(\boldsymbol{\omega})$, we get several images $I_i(\boldsymbol{\omega})$, $1 \le i \le n$, which correspond to different wavelengths. So, we have two groups of images:

- the images $I_i(\boldsymbol{\omega})$ which correspond to one area, and
- the images $I'_i(\boldsymbol{\omega})$ which correspond to an overlapping area.

Let us first consider the case when two images differ only by some (unknown) shift $\mathbf{a}$. For every wavelength i, the corresponding two images $I_i(\mathbf{x})$ and $I'_i(\mathbf{x})$ differ only by shift, i.e., $I'_i(\mathbf{x}) = I_i(\mathbf{x}+\mathbf{a})$. Therefore, for every wavelength i, their Fourier transforms

$$F_i(\boldsymbol{\omega}) = \frac{1}{2\pi} \cdot \int\int I_i(\mathbf{x}) \cdot e^{-2\pi \cdot \mathrm{i} \cdot (\mathbf{x} \cdot \boldsymbol{\omega})}\, \mathrm{d}x\mathrm{d}y,$$

$$F'_i(\boldsymbol{\omega}) = \frac{1}{2\pi} \cdot \int\int I'_i(\mathbf{x}) \cdot e^{-2\pi \cdot \mathrm{i} \cdot (\mathbf{x} \cdot \boldsymbol{\omega})}\, \mathrm{d}x\mathrm{d}y,$$

are related by the formula:

$$F'_i(\boldsymbol{\omega}) = e^{2\pi \cdot \mathrm{i} \cdot (\boldsymbol{\omega} \cdot \mathbf{a})} \cdot F_i(\boldsymbol{\omega}). \tag{18}$$

In the ideal no-noise situation, all these equations are true, and we can determine the value $r = e^{2\pi\cdot \mathrm{i}\cdot(\boldsymbol{\omega}\cdot\mathbf{a})}$ from any of these equations. In the real-life situations, where noise is present, these equations (18) are only approximately true, so we have the following problem instead: find r for which, for all i,

$$F_i'(\boldsymbol{\omega}) \approx r \cdot F_i(\boldsymbol{\omega}). \tag{19}$$

and which satisfies the condition (7).

We would like to get the best estimate for r among all estimates which satisfy the condition (7). To get the optimal estimate, we can use the Least Squares Method, according to which, for each estimate r and for each i, we define the error

$$E_i = F_i'(\boldsymbol{\omega}) - r \cdot F_i(\boldsymbol{\omega}) \tag{20}$$

with which the condition (19) is satisfied. Then, we find among all estimates which satisfy the additional condition (7), a value r for which the sum of the squares

$$|E_1|^2 + \ldots + |E_n|^2 = E_1 \cdot E_1^* + \ldots + E_n \cdot E_n^*$$

of these errors is the smallest possible.

The square $|E_i|^2$ of each error E_i can be reformulated as follows:

$$E_i \cdot E_i^* = (F_i'(\boldsymbol{\omega}) - r \cdot F_i(\boldsymbol{\omega})) \cdot \left({F'}_i^*(\boldsymbol{\omega}) - r^* \cdot F_i^*(\boldsymbol{\omega})\right) =$$

$$F_i'(\boldsymbol{\omega})\cdot {F'}_i^*(\boldsymbol{\omega}) - r^*\cdot F_i^*(\boldsymbol{\omega})\cdot F_i'(\boldsymbol{\omega}) - r\cdot F_i(\boldsymbol{\omega})\cdot {F'}_i^*(\boldsymbol{\omega}) + r\cdot r^*\cdot F_i(\boldsymbol{\omega})\cdot F_i^*(\boldsymbol{\omega}). \tag{21}$$

We need to minimize the sum of these expressions under the condition (7).

For this conditional minimization, we will use the Lagrange multipliers technique, which leads to the following unconditional minimization problem:

$$\sum_{i=1}^n \left(F_i'(\boldsymbol{\omega}) \cdot {F'}_i^*(\boldsymbol{\omega}) - r^* \cdot F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega}) - \right.$$

$$\left. r \cdot F_i(\boldsymbol{\omega}) \cdot {F'}_i^*(\boldsymbol{\omega}) + r \cdot r^* \cdot F_i(\boldsymbol{\omega}) \cdot F_i^*(\boldsymbol{\omega})\right) +$$

$$\lambda \cdot (r \cdot r^* - 1) \to \min. \tag{22}$$

Differentiating (22) relative to r^*, we get the following linear equation:

$$-\sum_{i=1}^n F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega}) + r \cdot \sum_{i=1}^n F_i(\boldsymbol{\omega}) \cdot F_i^*(\boldsymbol{\omega}) + \lambda \cdot r = 0. \tag{23}$$

From this equation, we conclude that

$$r = \frac{\sum_{i=1}^n F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega})}{\sum_{i=1}^n F_i(\boldsymbol{\omega}) \cdot F_i^*(\boldsymbol{\omega}) + \lambda}. \tag{24}$$

The coefficient λ can be now determined from the condition that the resulting value r should satisfy the equation (7). The denominator $\sum_{i=1}^{n} F_i(\boldsymbol{\omega}) \cdot F_i^*(\boldsymbol{\omega}) + \lambda$ of the equation (24) is a real number, so instead of finding λ, it is sufficient to find a value of this denominator for which $|r|^2 = 1$. One can easily see that to achieve this goal, we should take, as this denominator, the absolute value of the numerator, i.e., the value

$$\left| \sum_{i=1}^{n} F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega}) \right|. \tag{25}$$

For this choice of a denominator, the formula (23) takes the following final form:

$$r = R(\boldsymbol{\omega}) \stackrel{\text{def}}{=} \frac{\sum_{i=1}^{n} F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega})}{\left| \sum_{i=1}^{n} F_i^*(\boldsymbol{\omega}) \cdot F_i'(\boldsymbol{\omega}) \right|}. \tag{26}$$

So, for multi-spectral images, in the presence of noise, instead of using the exact ratio (4), we should compute, for every $\boldsymbol{\omega}$, the optimal approximation (26). Hence, we arrive at the following *algorithm*:

A new algorithm for determining the shift between two multi-spectral images

If we have images $I_i(\boldsymbol{\omega})$ and $I_i'(\boldsymbol{\omega})$ which correspond to different wavelengths, then, to determine the shift $\mathbf{a}$ between these two multi-spectral images, we do the following:

- *first, we apply FFT to the original images $I_i(\mathbf{x})$ and $I_i'(\mathbf{x})$ and compute their Fourier transforms $F_i(\boldsymbol{\omega})$ and $F_i'(\boldsymbol{\omega})$;*
- *on the second step, we compute the ratio (26) – setting the value to 0 if the denominator is below the threshold;*
- *on the third step, we apply the inverse FFT to the ratio $R(\boldsymbol{\omega})$ and compute its inverse Fourier transform $P(\mathbf{x})$;*
- *finally, on the fourth step, we determine the first approximation to the desired shift $\mathbf{a}$ as the point for which $|P(\mathbf{x})|$ takes the largest possible value, and perform the interpolation (16)–(17) to find the actual shift with subpixel accuracy.*

For rotation and scaling, we can use the same reduction to shift as for mono-spectral images. As a result, we get the desired values of shift, rotation, and scaling; hence, we get the desired referencing.