

A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration

Xunwei Xie, Yongjun Zhang, Xiao Ling & Xiang Wang

To cite this article: Xunwei Xie, Yongjun Zhang, Xiao Ling & Xiang Wang (2019): A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration, International Journal of Remote Sensing, DOI: [10.1080/01431161.2019.1579941](https://doi.org/10.1080/01431161.2019.1579941)

To link to this article: <https://doi.org/10.1080/01431161.2019.1579941>



Published online: 17 Feb 2019.



Submit your article to this journal [↗](#)



Article views: 39



View Crossmark data [↗](#)



A novel extended phase correlation algorithm based on Log-Gabor filtering for multimodal remote sensing image registration

Xunwei Xie^a, Yongjun Zhang^a, Xiao Ling^b and Xiang Wang^a

^aSchool of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei Province, China; ^bFuture Cities Laboratory, Singapore-ETH Centre, CREATE Tower, Singapore

ABSTRACT

Automatic registration of multimodal remote sensing images, which is a critical prerequisite in a range of applications (e.g. image fusion, image mosaic, and image analysis), continues to be a fundamental and challenging problem. In this paper, we propose a novel extended phase correlation algorithm based on Log-Gabor filtering (LGEPC) for the registration of images with nonlinear radiometric differences and geometric differences (e.g. rotation, scale, and translation). Our algorithm focuses on two problems that the traditional extended phase correlation algorithms cannot well handle: 1) significant nonlinear radiometric differences and 2) large-scale differences between image pairs. After an over-complete multi-scale atlas space of the original image is built based on the filtered magnitudes obtained by using Log-Gabor filters with different central frequencies, the phase correlation of the single scale images is extended by LGEPC to atlases phase correlation, which is conducive to solving the problem of large scale and rotation differences between the image pairs. Subsequently, LGEPC eliminates the interface of the significant nonlinear radiometric differences by superimposing multi-scale geometric structural spectra and carrying out the phase correlation module, so that the translation can be well determined. Our experiments on synthetic images demonstrated the rationality and effectiveness of LGEPC, and the experiments on a variety of multimodal images confirmed that LGEPC can ideally achieve pixel-wise registration accuracy for multimodal image pairs that conform to the similarity transformation model.

ARTICLE HISTORY

Received 12 August 2018
Accepted 22 December 2018

1. Introduction

The purpose of image registration is essentially to overlap two or more images of the same scene taken by different sensors, thereby geometrically aligning the reference and sensed images (Zitova and Flusser 2003; Szeliski 2010). It is a fundamental and crucial problem in remote sensing analysis tasks, such as change detection, image fusion, and image mosaic (Wong and Clausi 2007a). And the image registration results will seriously affect the performance of these follow-up processes.


CONTACT Xunwei Xie  xwxie@whu.edu.cn  School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei Province, China; Yongjun Zhang  zhangyj@whu.edu.cn  School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, Hubei Province, China

Image registration has been widely applied in the fields of remote sensing (Bentoutou et al. 2005; Dawn, Saxena, and Sharma 2010; C. Chen et al. 2014) mainly using these two types of methods: feature-based methods (FBMs) and area-based methods (ABMs).

FBMs first extract the salient structural information of the images and detect features like points, lines, and edges as the matching units. Then, the correspondences with similar local descriptors are found. In general, the common feature descriptors used in remote sensing images are Scale Invariant Feature Transform (SIFT) (Lowe 2004) and its various variants (Sedaghat, Mokhtarzade, and Ebadi 2011; Cai et al. 2013; Dai, Song, and Li 2014) Speed Up Robust Features (SURF) (Bouchiha and Besbes 2013) and wavelet feature (Wong and Clausi 2010b). For remote sensing image registration with linear radiometric differences, many related studies have been carried out and certain achievements have been made (Kim and Im 2003; Wan and Zhang 2017). The common practice is to combine a traditional SIFT algorithm with the geometric characteristics of the imaging; overcome the influence of big data and complex scenes on the accuracy of the SIFT matching algorithm; and obtain numerous corresponding points (Vural, Yardimci, and Temzel 2009; Xu, Zhang, and Li 2014). On basis of SIFT, a more uniform and robust SIFT method is proposed by (Sedaghat, Mokhtarzade, and Ebadi 2011), which can generate a sufficient number of reliable and uniformly distributed points. Also, an advanced self-similarity descriptor was proposed by Sedaghat and Ebadi (2015), which can exhibit high discrimination. For the registration tasks of the images with a large-scale difference, Sun et al. (2014) proposed an efficient SIFT matching method. However, these methods often suffer from the poor distribution and number of points in multimodal images (Ghassabi et al. 2013), further affecting the transformation model estimation. The reason for this is that it is difficult to extract highly repeatable shared features and describe their feature vectors for multimodal images.

On the whole, FBMs show robust geometric registration and low time complexity, but highly repeatable shared features extraction and description are not easy works for multimodal images.

ABMs can be subdivided into three categories: mutual information (MI) methods, correlation-like methods, and Fourier methods (Zitova and Flusser 2003).

The MI methods work directly with image intensities which are not sensitive to local differences such as inequality grey, geometric distortion, so they are suitable for multimodal remote sensing image registration (Inglada et al. 2007; Liang et al. 2014; Chen, Arora, and Varshney 2003; S. Chen et al. 2018; Zhang et al. 2018; Gong et al. 2014). And they can achieve excellent performance and robustness with only slightly more time expended (Gharbia, Ahmed, and Hassanien 2015). The basic principle of most of these methods is to solve the transformation parameters that maximizes the similarity metric of corresponding voxels. However, some researches show that MI as a similarity measure may not obtain accurate registration results, while the normalized MI (NMI) may be more robust. In addition, in order to integrate spatial information, Regional Mutual Information (RMI) calculates joint probability distribution by eight-adjacent information instead of original grey information (Zhao et al. 2015), which can be considered a replacement of MI metric in some special cases. Inspired by RMI, a rotationally invariant regional and mutual information (RIRMI) method is proposed by (S. Chen et al. 2018), which considers not only spatial information but also the effect of the local grey variations and rotation changes on computing probability density function. Their experiments verified that RIRMI was more robust and accurate than the original MI-based registration methods.

Correlation-like methods, such as Normalization Cross Correlation (NCC), are early classical representatives of the ABMs (Gonzalez and Wintz 1987), which strive to compute the similarity metrics of the image windows and thus consider the one with the largest similarity as the correspondence. However, the traditional correlation-like methods based on grey information are not robust to multimodal images and the situations where the regions lack texture.

One common technique of the Fourier methods is phase correlation method (PC) based on the Fourier shift theorem (Reddy and Chatterji 1996; Chen, Defrise, and Deconinck 1994). They have advantages in efficiency and are robust to frequency-dependent noise. The original PC was proposed to match two images with translation differences (Kuglin and Hines 1975; Foroosh, Zerubia, and Berthod 2002), and some extended phase correlation (EPC) methods were proposed to solve rotation and scaling factor (Decastro and Morandi 1987; Reddy and Chatterji 1996; Ge, Lan, and Wang 2014). As an increase in the radiometric differences between two images can cause a gradual decrease in the salience of phase correlation, the EPC only can tolerate the inhomogeneous illumination difference between the two images. Furthermore, in theory, EPC can deal with images with any scale differences in a continuous frequency domain. However, because of the discretization selection of frequency domain in Fourier transform, EPC also cannot handle the large scale differences between images. Therefore, EPC cannot realize the theoretical effects in many multimodal image registration tasks. Some research showed that the results of EPC will not be reliable when the scaling ratio between two images is more than 1.8 due to the great difficulty in finding the peak in the phase correlation module (Reddy and Chatterji 1996). In (Tong et al. 2015), a new EPC registration method in frequency domain was proposed so as to achieve affine variance using singular value decomposition and unified random sample consensus. However, they focused not on significant radiometric differences.

Strictly speaking, EPC and its variants belong to the transformation model methods (TMMs), because they directly yielded the parameters of transformation model other than the discretely distributed correspondences like Histogram of Orientated Phase Congruency (HOPC) (Ye et al. 2017). Although some researchers utilize the idea of Fourier methods, their methods actually belong to NCC-like methods. These methods often use a similarity metric based on phase congruency other than NCC in the template window, among which Automatic Registration of Remote-Sensing Images (ARRSI) and HOPC are the representatives (Wong and Clausi 2007a; Ye and Shen 2016; Ye et al. 2017). These methods and the correlation-like methods can be collectively called spatial domain methods (SDMs). Both ARRSI and HOPC used a matching progress based on the phase congruency model, and their difference is that HOPC constructs a more robust descriptor for multimodal images. ARRSI employed a phase congruency model-based method in the control point detection to address global and local contrast and illumination conditions that may affect the accuracy of the detected control points. And ARRSI also used a phase-congruency moment-based patches as local feature descriptors that are invariant to intensity-mapping conditions during the matching process. HOPC extracted the highly consistent structural information of the same features in the images of interest (Anuta 1970) using Log-Gabor (Field 1987; Gabor 1946), utilized these structural information to calculate phase congruency orientation and further constructed HOPC descriptor. As HOPC is a feature descriptor that captures the internal structures of images, it can be used to match two images with significant nonlinear radiometric differences using NCC framework (Ye et al. 2017). The reason for using Log-Gabor is that Gabor filter can

seize multi-direction and multi-scale local structures with excellent spatial locality and directional selectivity (Feichtinger and Strohmer 2012), so it is widely used in image edge detection and visual information comprehension (Kokila and Thangavel 2014; Lee 1996; Serrano et al. 2010). In addition, extracting geometric structures with Log-Gabor showed a better performance than other methods, such as contour gradient-based methods (Elder and Zucker 1998; Arbelaez et al. 2011) and wavelet transform-based methods (Kovesi 1999). However, the prerequisite of using HOPC is that the rotation and scale differences should be first eliminated. Compared to TMMs, these SDMs need extract interest points as candidate points so as to utilize the similarity metric like HOPC, and most of their authors do not tell the readers that what to do if we extract few or no candidate points.

Specifically, ABMs will obtain high registration accuracy, while they also require more preprocessing, such as eliminating most of rotation or scale differences.

As the traditional EPC cannot well deal with the problems caused by nonlinear radiometric and large scale differences, we combined the advantages of EPC and Log-Gabor and proposed a novel extended phase correlation algorithm based on Log-Gabor filtering (LGEPC) in this paper. The idea is based on the fact that the multi-scale geometric structural information of image without being affected by the radiometric differences will be conducive to the phase correlation between images. The proposed method has the following two contributions:

- (1) we use the filtered magnitude images obtained by Log-Gabor filtering with different central frequencies to construct an over-completed multi-scale atlas space. Thereby a process of atlases phase correlation is conducted to obtain the rotation and scale. Specifically, the phase correlation of every magnitude image pair in the multi-scale atlas spaces between the reference and sensed images is constructed using the phase correlation method (Kuglin and Hines 1975; Decastro and Morandi 1987; Reddy and Chatterji 1996). And the scale factor and rotation angle can be determined by calculating the maximum peak of the atlases phase correlation. These processes can help solve the large scale difference problems in the step of solving rotation-scale in EPC module.
- (2) we superimpose the filtered structural spectra obtained by Log-Gabor filtering with different central frequencies together to enhance the overall structural information, so as to be conducive to eliminate the influence of radiometric differences as accurately as possible in the step of solving translation in EPC module.

The above two points can effectively improve the applicability and stability of EPC to a certain extent. And it must be emphasized that LGEPC is an extension of EPC and essentially also belongs to a TMM method within the scope of ABMs.

2. Methodology

This section describes the process of LGEPC for multimodal remote sensing image registration. We start by briefly introducing an important module in our method, i.e. EPC, and then demonstrate how to and why we extract the multi-scale structural information with a Log-Gabor filter. Some details of LGEPC are thereafter described, followed by the parameter settings to conclude this section.

2.1. Extended phase correlation algorithm

The phase correlation image registration is possible to obtain good robustness against correlated and frequency dependent noise, and it depends on the transformation property of Fourier transform known as the shift theorem (Sarvaiya, Patnaik, and Kothari 2012). Given a reference image $\mathbf{m}(x, y)$ and a sensed image $\mathbf{n}(x, y)$, the translation (x_0, y_0) differs them as

$$\mathbf{n}(x, y) = \mathbf{m}(x - x_0, y - y_0) \quad (1)$$

According to Fourier theorem (Reddy and Chatterji 1996), we can get the following relationship, as in

$$\mathbf{F}_n(\xi, \eta) = \mathbf{F}_m(\xi, \eta)e^{-i2\pi(\xi x_0 + \eta y_0)} \quad (2)$$

where $\mathbf{F}_m(\xi, \eta)$ and $\mathbf{F}_n(\xi, \eta)$ represent the Fourier transform of $\mathbf{m}(x, y)$ and $\mathbf{n}(x, y)$; (ξ, η) represents the frequency variables in the frequency domain; $i^2 = -1$.

From Equation (2), the translation will not affect the Fourier magnitude but differ the phase. The cross-power spectrum of $\mathbf{m}(x, y)$ and $\mathbf{n}(x, y)$ is calculated by

$$\mathbf{C}(\xi, \eta) = \frac{\mathbf{F}_m(\xi, \eta)\mathbf{F}_m^*(\xi, \eta)}{\mathbf{F}_n(\xi, \eta)\mathbf{F}_n^*(\xi, \eta)}e^{-i2\pi(\xi x_0 + \eta y_0)} \quad (3)$$

Taking the inverse Fourier transform of Equation (3), we can get a delta function that is an impulse, in which the location of non-zero specifies the translation (Reddy and Chatterji 1996).

The above theorem defines how to compute the translation, while the rotation-scale can be obtained by Fourier-Mellin transform theorem (Reddy and Chatterji 1996). Further, if $\mathbf{m}(x, y)$ and $\mathbf{n}(x, y)$ are differs by rotation θ , isotropic scale s and transform $t = (x_0, y_0)$, i.e.

$$\mathbf{n}(x, y) = \mathbf{T}(\theta, s, t)\mathbf{m}(x, y) \quad (4)$$

where $\mathbf{T}(\theta, s, t)$ represents a rotation-scale-translation transform. Their Fourier transform $\mathbf{F}_m(\xi, \eta)$ and $\mathbf{F}_n(\xi, \eta)$ have the following relationship, i.e.

$$|\mathbf{F}_n(\xi, \eta)| = s^{-2}|\mathbf{F}_m(s^{-1}\xi\cos(\theta) + s^{-1}\eta\sin(\theta), -s^{-1}\xi\sin(\theta) + s^{-1}\eta\cos(\theta))| \quad (5)$$

where $|\cdot|$ represents magnitude. From Equation (5), their magnitudes are independent of translation, so the rotation and scale can be calculated by the magnitudes of the reference and sensed images.

As the scaling and rotation in the Cartesian domain correspond to pure translation in log-polar domain, we can utilize the magnitude images of reference and sensed images in the log-polar space to calculate scale and rotation. The magnitudes \mathbf{F}_{mp} and \mathbf{F}_{np} of the reference and sensed images in the log-polar domain, are related by

$$\mathbf{F}_{np}(\alpha, \ln \rho) = \mathbf{F}_{mp}(\alpha - \theta, \ln \rho - \ln s) \quad (6)$$

where α, ρ represent the radial distance and angle from the centre in log-polar domain, respectively.

Note that Equation (6) and Equation (1) are the same in form, so that θ and s (actually, $\ln s$) can be calculated by Equations (1)-(3) in log-polar domain using the Fourier transform theorem.

In accordance with the recovered scale factor and rotation, the sensed image is resampled to eliminate its rotation and scale difference to the reference image, followed by removing the translations if necessary. However, the difference between our method and previous ones is that we do not directly use the magnitude spectra of original reference and sensed images but the magnitude spectra of their multi-scale structural information. Next section we will elaborate how our method combines the multi-scale structural information generated by multi-scale Log-Gabor and phase correlation algorithm in detail.

2.2. Multi-scale structural information extraction via Log-Gabor filtering

The Gabor filter is defined as a plane wave constrained by the Gauss envelope function in the two-dimensional (2D) space domain (Daugman 1985; Lades et al. 1993), as in

$$g_{\vec{k}}(\vec{x}) = \frac{\vec{k}^2}{\delta^2} e^{-0.5\delta^{-2}\vec{k}^2\vec{x}^2} (e^{i\vec{k}\vec{x}} - e^{-\frac{\delta^2}{2}}) \quad (7)$$

where the first term is the oscillating part of the Gabor kernel function; the second term is the direct-current (DC) compensation component, which eliminates the influence of the luminance difference on the kernel function; \vec{x} is the 2D plane coordinate; \vec{k} controls the wavelength and direction of the oscillating part; and δ determines the proportion between the width of the Gaussian and the wavelength. However, when the width of the Gabor filter reaches a certain value, its response values will be affected by the image brightness, which is not conducive to the structural information extraction of the same feature under different radiometric conditions.

On this basis, when compared to the Gabor filter, the Log-Gabor filter introduced by Field (Field 1987) does not contain the DC compensation component and is more suitable for dealing with natural images. As the edge structural information independent of the direction is what we need, the transfer function of the 2D Gabor filter, which is unaffected by direction in the frequency domain, is what we focus on Sarvaiya, Patnaik, and Kothari (2012), as in

$$\mathbf{G}(f) = e^{-\frac{\ln^2(f/f_0)}{2\ln^2(\delta/f_0)}} \text{ or } \mathbf{G}(\xi, \eta) = e^{-\frac{\ln^2(\sqrt{\xi^2+\eta^2}/f_0)}{2\ln^2(\delta/f_0)}} \quad (8)$$

where f_0 is the centre frequency of the filter; δ is the bandwidth of the filter; (ξ, η) is the coordinates in the frequency domain. If δ/f_0 is constant, the form of the Gabor filter will remain the same.

Figure 1 shows the 2D images of Log-Gabor filtering with a different f_0 in the frequency domain and the results obtained by using the Gabor filters to filter the image shown in Figure 3(a). It is apparent that the lower f_0 is, the smaller the range of the Log-Gabor and the closer it is to the origin in the frequency domain, which is to say that a lower f_0 corresponds to the large scale structural information. On the other hand, the higher f_0 is, the more it corresponds to the medium or small scale structural information. The enhancement of multi-scale structural information is conducive to the phase correlation of the images with large scale differences.

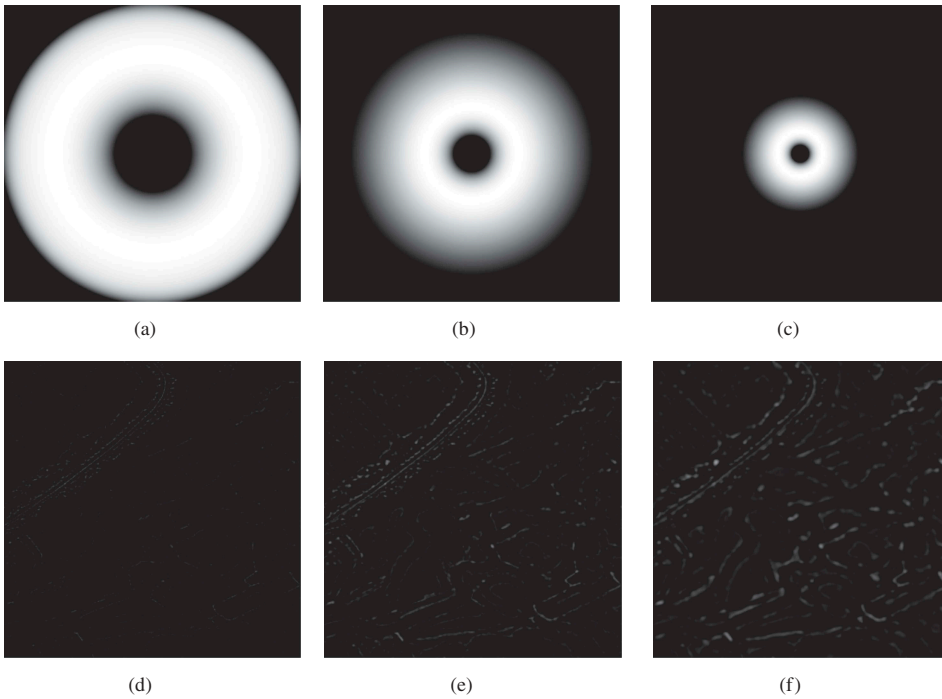


Figure 1. 2D images of Log-Gabor with different f_0 in the frequency domain and the filtering results of Figure 3(a). (a)–(c) are the 2D images of Log-Gabor with $f_0 = 0.333$, $f_0 = 0.158$ and $f_0 = 0.076$ in the frequency domain; (d)–(f) are the corresponding filtered structural information of Figure 3(a).

Take the log-polar coordinates (α, ρ) back to the Cartesian domain of Equation (8), i.e. $(\xi, \eta) = (e^\rho \cos \alpha, e^\rho \sin \alpha)$, the Log-Gabor has the corresponding transfer function, as in

$$\mathbf{G}_{\log}(\alpha, \rho) = e^{-\frac{(\rho - \ln f_0)^2}{2\ln^2(\sigma/f_0)}} \quad (9)$$

Equation (9) demonstrates that the transfer function of Log-Gabor filter without considering the direction information in log-polar space is very simple in form, and it is the extension of 1D Gaussian function in a certain direction of 2D space.

To sum up, the proposed LGEPC method utilizes the Log-Gabor filter to extract the geometric structural information for the following three reasons.

- (1) The Log-Gabor filter does not contain the DC component and is insensitive to the local radiometric differences, which is more suitable for the identical structural information extraction in the multimodal images than the algorithms based on contour gradients.
- (2) The Log-Gabor filter is able to extract multi-scale structural information, which can make a big difference in phase correlation of images with scale differences. We will discuss it in more detail later in this paper.
- (3) In addition, it is easy to achieve the Log-Gabor filter in the log-polar domain without overburdening the process.

2.3. The LGEPC algorithm

There are two main components to estimate the similarity transformation model in this paper, which is discussed below: 1) solving the rotation and scale differences and 2) solving the translation differences between the reference image and the sensed image. Also, the workflow of LGEPC is shown in Figure 5, in which only some of the key steps are described because some steps of LGEPC are similar to those of the traditional EPC.

2.3.1. Solving rotation and scale differences

Similar to the traditional EPC, LGEPC obtains the magnitude spectra of the reference image and sensed image in the log-polar domain via a series of transformations (i.e. the Fourier transform and the Log-Polar transform), which are shown in part one of Figure 2. Given a reference image $I_r(x, y)$ and a sensed image $I_s(x, y)$, they are both transformed by Fourier transform and Log-Polar transform:

$$F_i^p(\xi, \eta) = T_{\log}(\text{FFT}(I_i(x, y))), i = r, s \quad (10)$$

where $\text{FFT}(\cdot)$ represents Fourier transform; $T_{\log}(\cdot)$ represents Log-Polar transform; $(\cdot)^p$ represents the component in the log-polar domain; $i = r$ and $i = s$ represent reference image and sensed image, respectively. $|F_i^p|$ represents the magnitude of reference image or sensed image in the log-polar domain.

(1) Log-Gabor filtering with multi-scale

This step aims to build the multi-scale filtered atlases of the reference and sensed images with their magnitudes filtered by the Log-Gabor filter with different central

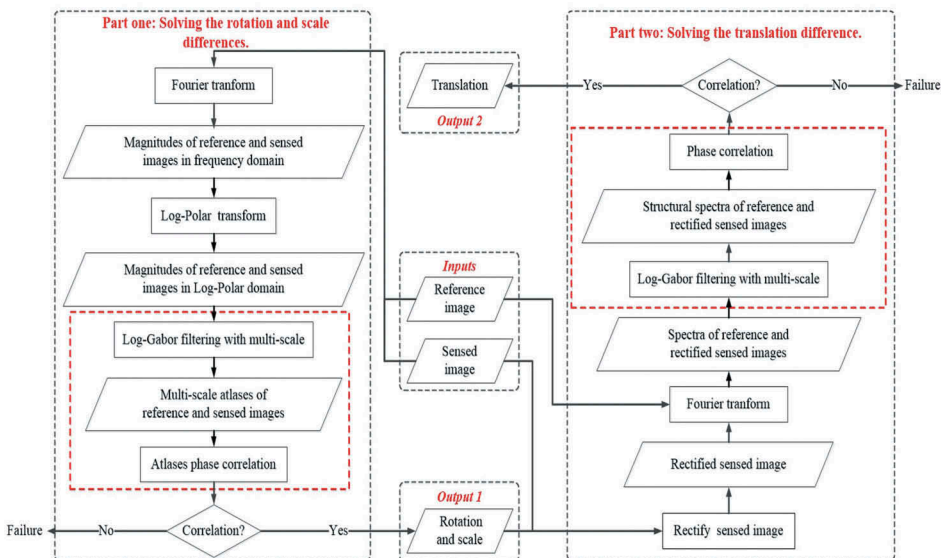


Figure 2. Flowchart of LGEPC. Inputs: reference image and sensed image. Outputs: similarity transformation model parameters (e.g. rotation, scale, translation). The red dotted rectangle areas represent the specific parts of LGEPC.

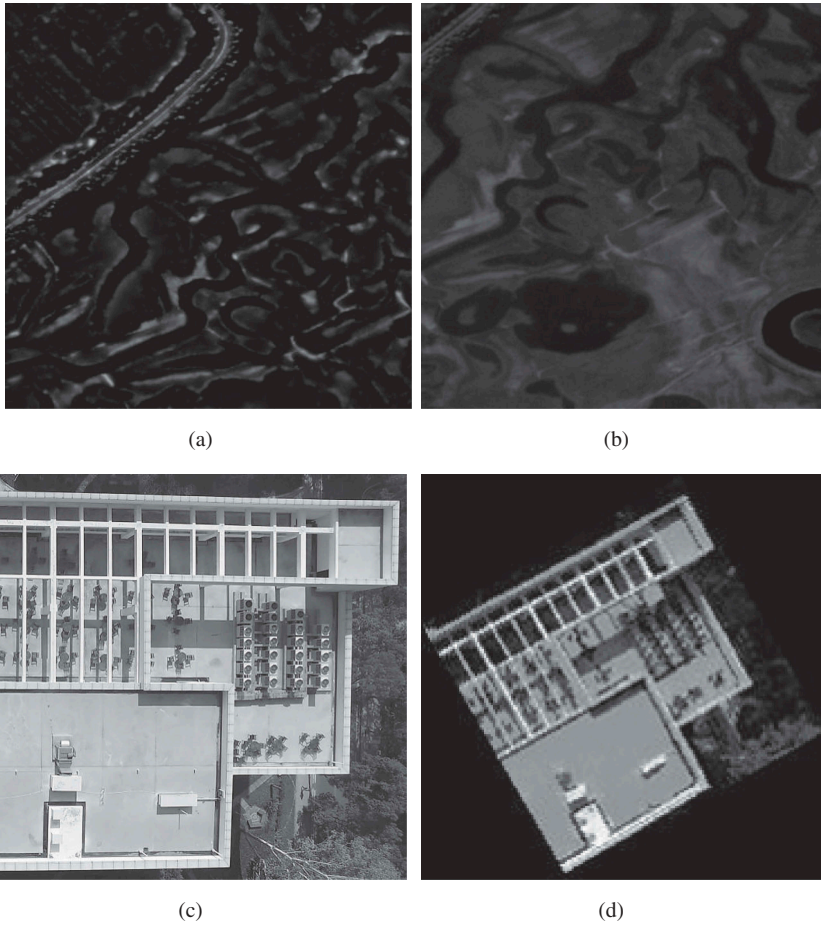


Figure 3. Data sets of test 1 and test 2. (a) test1-reference image (Chinese ZY-3 satellite image). (b) test1-sensed image. (c) test2-reference image (low-altitude visible spectral). (d) test2-sensed image. The original images are scaled for better visualization.

frequencies. Due to the scale difference between the two images reaches a certain value, the phase correlation will fail to solve the scaling ratio. Therefore, we use the atlases phase correlation algorithm to eliminate the scale differences, especially large scale differences, so as to overcome this problem.

Three kinds of Log-Gabor filters with different central frequencies (i.e. f_{\min} , f_{mid} and f_{\max}) are used to filter the magnitudes of the reference and sensed images in the log-polar domain using Equation (11),

$$\mathbf{A}_i^{(j)}(\xi, \eta) = |\mathbf{F}_i^p(\xi, \eta)| * \mathbf{G}_j(\xi, \eta), j = 1, 2, 3; i = r, s \quad (11)$$

where $\mathbf{A}_i^{(j)}(\xi, \eta)$ represents the Log-Gabor filtered result of the magnitude of the reference image or sensed image; $\mathbf{G}_j(\xi, \eta)$ represents Log-Gabor filter, whose form is as Equation (9); $j = 1, 2, 3$ represents the Log-Gabor filters with f_{\min} , f_{mid} and f_{\max} .

From (11), we obtain three filtered magnitude images corresponding to the reference image and sensed image, respectively, whose properties are summarized in Table 1.

Table 1. The properties of the multi-scale filtered magnitude images.

| Filtered magnitude image | Log-Gabor filter | structural information |
|--|---|------------------------|
| $\mathbf{A}_r^{(1)}, \mathbf{A}_s^{(1)}$ | low central frequency f_{\min} | large scale |
| $\mathbf{A}_r^{(2)}, \mathbf{A}_s^{(2)}$ | medium central frequency f_{mid} | medium scale |
| $\mathbf{A}_r^{(3)}, \mathbf{A}_s^{(3)}$ | high central frequency f_{\max} | small scale |

Note that the filtered magnitude images $\mathbf{A}_i^{(j)}$ can reflect the structural information at different levels and there is nearly no overlapping information between them.

Subsequently, we respectively utilize the filtered magnitude images of the reference image and sensed image to construct their multi-scale atlas spaces, so as to facilitate phase correlation between every image pair. The basic principle of multi-scale atlas space is that the upper structural information in the multi-scale atlas space will be completely contained in the next layer, which is much similar to an image pyramid. That is to say, the multi-scale atlas space belongs to an over-complete scale space. Here, a mathematical formula is used to express the relationship between the multi-scale atlas space $\mathbf{S}_i = \{\mathbf{S}_i^t, \mathbf{S}_i^m, \mathbf{S}_i^b\}$ and the corresponding magnitude images $\mathbf{A}_i^{(j)}$ as follows:

$$\begin{cases} \mathbf{S}_i^t = \mathbf{A}_i^{(1)} \\ \mathbf{S}_i^m = \sqrt{(\mathbf{A}_i^{(1)})^2 + (\mathbf{A}_i^{(2)})^2} \\ \mathbf{S}_i^b = \sqrt{(\mathbf{A}_i^{(1)})^2 + (\mathbf{A}_i^{(2)})^2 + (\mathbf{A}_i^{(3)})^2}, i = r, s \end{cases} \quad (12)$$

where \mathbf{S}_i^t contains only the large scale structural information and \mathbf{S}_i^m contains the large and medium scale structural information, and yet \mathbf{S}_i^b contains all the structural information. The multi-scale atlas space constructed according to Equation (12) is only formally similar to the image pyramid.

Because the multi-scale atlas space is based on the magnitude of the signal, it is constructed by using the root of the sum of squared magnitudes in Equation (12).

(2) Atlases phase correlation

After the Log-Gabor filtering with multiple central frequencies is conducted, we build the multi-scale atlas space \mathbf{S}_r of the reference image and the multi-scale atlas space \mathbf{S}_s of the sensed image. This step aims to conduct phase correlation between each image pair in their multi-scale atlas spaces, and then obtain the image pair corresponding to the maximum response peak of atlases phase correlation so as to obtain the rotation and scaling parameters. The special process to realize that is as follows.

First, the phase correlation between each image pair in their multi-scale atlas spaces is carried out to solve rotation and scale using EPC module as shown in Equations (1–6), and this process can be simplified as:

$$\mathbf{T}_{so}(i, j) \leftarrow \mathbf{S}_r(i) \oplus \mathbf{S}_s(j), i, j = N \quad (13)$$

where \oplus represents the phase correlation operation using the EPC; N represents the layer number of the multi-scale atlas space, and it is fixed as three in this paper; and $\mathbf{T}_{so}(i, j)$ represents the rotation-scale transformation model of image pair (i, j) .

In this process, the reliability of phase correlation results must be tested (Reddy and Chatterji 1996). There are two key thresholds in this step: response intensity t_{mag} , and the ratio t_{ratio} of the sub-maximum response peak and the maximum response peak. The maximum response peak calculated in the EPC module directly reflects the strength of phase correlation between each image pair. If the maximum response peak is less than some given threshold t_{mag} , we consider that phase correlation failure may occur. Besides, if the ratio of the sub-maximum response peak and the maximum response peak is greater than t_{ratio} , we believe that the correlation between the two images is not significant enough and also determine that phase correlation failure occurs. We record every rotation-scale transformation model $\mathbf{T}_{so}(i,j)$ and the maximum response peak $Peak(i,j)$ after the EPC module is carried out.

Second, the maximum among $IEOP_A_1574745$ can be obtained, and its corresponding rotation-scale transformation model \mathbf{T}_{so} is what we want.

2.3.2. Solving translation differences

First, we rectify the sensed image via bilinear interpolation to eliminate its rotation and scale differences to the reference image according to rotation-scale transformation model \mathbf{T}_{so} obtained in part one of Figure 2:

$$\mathbf{I}_s^{rec}(x,y) = \mathbf{T}_{so} * \mathbf{I}_s(x,y) \quad (14)$$

where $\mathbf{I}_s^{rec}(x,y)$ and $\mathbf{I}_s(x,y)$ represent the rectified sensed image and the sensed image, respectively.

Then, we obtain the spectra of the reference image and the rectified sensed image via the Fourier transform, followed by filtering their spectra to obtain the structural spectra using the Log-Gabor filters with different central frequencies as Equation (15):

$$\mathbf{F}_i^{(j)\dagger}(\xi,\eta) = \mathbf{G}_j(\mathbf{F}_i(\xi,\eta)), j = 1, 2, 3; i = r, rec \quad (15)$$

where $\mathbf{F}_i(\xi,\eta)$ represents the Fourier transform result of the image; $\mathbf{F}_i^{(j)\dagger}(\xi,\eta)$ represents the filtered result in the frequency domain; $\mathbf{G}_j(\bullet)$ represents the Log-Gabor filtering operation in frequency domain; $j = 1, 2, 3$ represents the Log-Gabor filters with f_{min} , f_{mid} and f_{max} ; $i = r, rec$ represents the reference image and the rectified sensed image. It is noted that the purpose of Log-Gabor filtering with multiple central frequencies in this step is to extract as much of the structural information as possible so as to eliminate the interference of radiometric differences. Therefore, unlike part one in Figure 2, the multi-scale atlas space is not needed to deal with the scale differences.

Thereafter, the three filtered structural spectrum images obtained by the Log-Gabor filters with different central frequencies are simply superimposed together to obtain the final structural spectrum as follows:

$$\mathbf{F}'_i(\xi,\eta) = \sqrt{[\mathbf{F}_i^{(1)\dagger}(\xi,\eta)]^2 + [\mathbf{F}_i^{(2)\dagger}(\xi,\eta)]^2 + [\mathbf{F}_i^{(3)\dagger}(\xi,\eta)]^2}, i = r, rec \quad (16)$$

where $\mathbf{F}'_i(\xi,\eta)$ is the overall superimposed structural spectrum.

In addition, in order to directly implement EPC module, we apply the inverse Fourier transform to get the corresponding filtered result of the Log-Gabor wavelet in the spatial domain:

$$\mathbf{lout}'_i = \text{IFFT}(\mathbf{F}'_i(\xi, \eta)), i = r, \text{rec} \quad (17)$$

where \mathbf{lout}'_i represents the corresponding filtered result of the Log-Gabor in the spatial domain, and IFFT represents the inverse Fourier transform.

Finally, we implement the phase correlation to solve translation as shown in Equations (1–3), and obtain the maximum response peak in an impulse, whose coordinates specify the translation between the reference image and the rectified sensed image. In accordance with the above process, the rotation, scale, and translation between the reference image and the sensed image can be obtained.

In order to facilitate the display in [Figure 2](#), here we rectify the sensed image to eliminate its geometric difference to the reference image. However, it is actually necessary to decide whether to down-sample the large-scale image or up-sample the small-scale image according to the actual scale parameters.

2.4. Parameter settings

In order to build the multi-scale atlas space, there are three parameters in the LGEPC: δ/f_0 , N and $f_0(i)$ in the Log-Gabor filter. Parameter δ/f_0 is the ratio of the bandwidth and the central frequency of the filter. Parameter N is the number of layers of the multi-scale atlas space. Parameter $f_0(i)$ is the central frequency of the i^{th} layer. The $f_0(i)$ of the i^{th} layer is not set directly, but rather is calculated by

$$f_0(i) = \frac{1}{\lambda_{\min} s^{i-1}} \quad (18)$$

where λ_{\min} is the minimum wavelength and s is the ratio of the wavelength between the adjacent layers.

According to the recommendation of Kovessi, the parameters of the Log-Gabor filter are set as follows: $\delta/f_0 = 0.55$, $\lambda_{\min} = 3$, $s = 2.1$ and $N = 3$ (Kovessi 1999). That is to say, there are three layers in the multi-scale atlas space, and the scale ratio between the adjacent layers is 2.1. There are also two parameters in the phase correlation module: t_{mag} and t_{ratio} . Likewise, according to Reddy, the two parameters are selected for $t_{\text{mag}} = 0.03$ and $t_{\text{ratio}} = 0.75$ (Reddy and Chatterji 1996).

3. Experimental results and analysis

In this section, the rationality and effectiveness of LGEPC will be tested with the synthetic images, and the performance of LGEPC also is evaluated with three different types of multimodal remote sensing images. In following, the data sets, implementation details and experimental results are presented as well as a failure case.

3.1. Data sets

The data sets are divided into two classes: 1) synthetic images which are used to evaluate the resistance of LGEPC on images with the nonlinear radiometric differences and the large-scale differences and 2) multimodal remote sensing images which are used to evaluate the performance of LGEPC.

The first data set class mainly tested two properties of LGEPC: 1) whether LGEPC can successfully degenerate the similarity transform model into a translation model when there are only nonlinear radiometric differences and translation differences between the two images; and 2) whether LGEPC can successfully estimate a transformation model when there are nonlinear radiometric differences, rotation and large-scale differences between the two images. We designed two sets (i.e. test 1 and test 2) of synthetic images to verify the two properties, as shown in Figure 3.

Figure 3(a) and Figure 3(c) are a Chinese ZY-3 satellite image and a low-altitude visible spectral image, respectively. Figure 3(b) is obtained by the brightness and contrast enhancement of Figure 3(a), followed by translating (100, 100). Figure 3(d) is obtained by the brightness enhancement and nonlinear stretching of Figure 3(c), followed by rotating 30° and down-scaling 4 times. Figure 3(a) and Figure 3(b) are both 400×400 , while Figure 3(c) and Figure 3(d) are 524×532 , 181×181 in size, respectively. The nonlinear radiometric differences are common between each image pair.

In addition, to evaluate the performance of LGEPC, we also selected three multimodal remote sensing image pairs, which were divided into three categories: Visible-to-Infrared (test 3), LiDAR-to-Visible (test 4), and Panchromatic-to-Multispectral (test 5), as shown in Figure 4. Significant radiometric differences are common between these image pairs

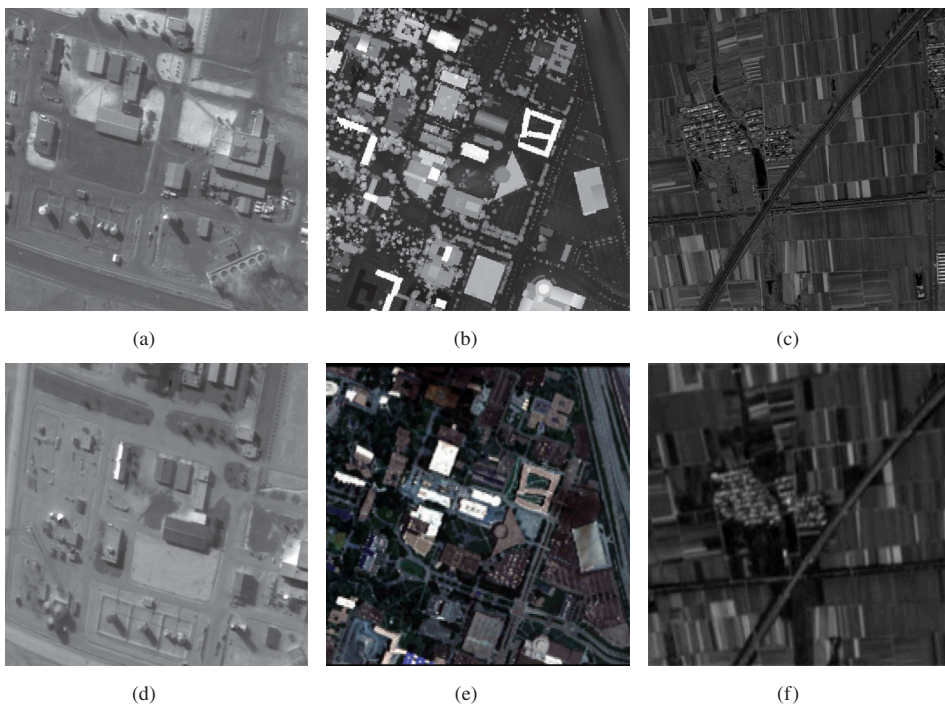


Figure 4. Multimodal remote sensing images. (a) test 3-reference image (visible spectral). (b) test 4-reference image (LiDAR depth). (c) test 5-reference image (panchromatic). (d) test 3-sensed image (infrared). (e) test 4-sensed image (visible spectral). (f) test 5-sensed image (multi-spectral: red band). The original images are scaled for better visualization.

Note that the original images shown in Figure 3 and Figure 4 are scaled for better visualization, therefore, the real scale differences are not reflected.

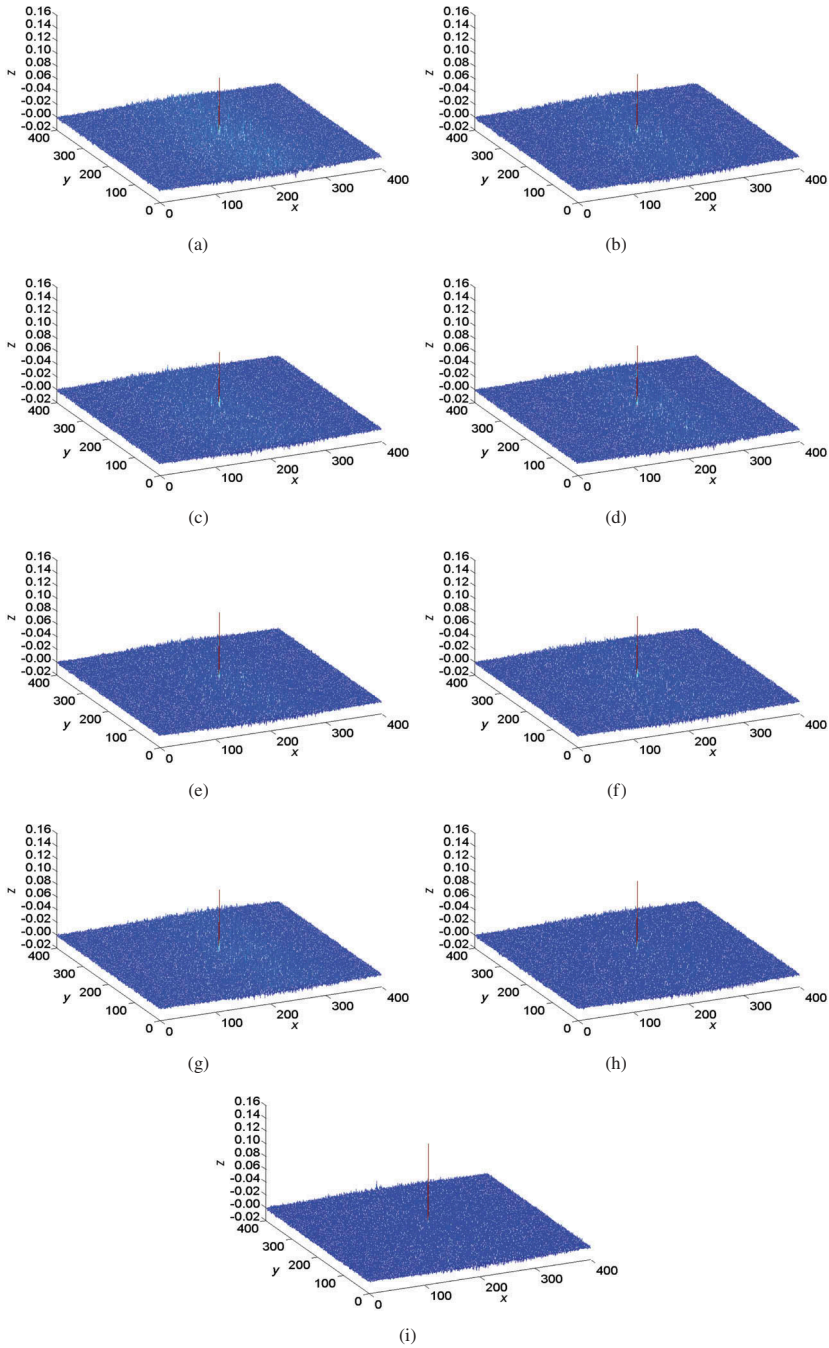


Figure 5. Results of atlases phase correlation on test 1. (a)-(i): phase correlation results between S_r^t, S_r^m, S_r^b of the atlas S_r and S_s^t, S_s^m, S_s^b of the atlas S_s , respectively. (x, y) are the coordinates in the pulse image obtained by LGEPC in solving the rotation and scale, and z is the peak corresponding to each (x, y) . The maximum peaks of the z -axis in each sub-figure, which reflect the strength of the correlation, are 0.093, 0.099, 0.090, 0.100, 0.109, 0.103, 0.102, 0.116, and 0.131 in (a)-(i), respectively, for comparing their value. The z -coordinates are taken from -0.02 to 0.16 . All the above results both gave the correct scale $s_0 = 1.0$, and rotation $\theta_0 = 0.0^\circ$. However, the peak in (i) was the most significant, so its results were the most reliable according to (Reddy and Chatterji 1996).

because they were acquired by different imaging modalities. Among them, the test 3 and test 4 were provided by (Ye and Shen 2016). And the test data are described below.

The first group was visible spectral (Daedalus) and infrared (Daedalus), respectively, and both were 400×400 in size. The image pair contained bare land and very few buildings. There were large translation and nonlinear radiometric differences between them; however, the geometric structural information of the same feature was nearly identical. The second group was the LiDAR depth image and visible spectral image (Airborne), 524×524 and 220×174 in size, respectively. The overlapped areas were the urban setting areas. Although there were obvious radiometric differences and geometric differences, the geometric structural information of the buildings was very similar.

The third group was panchromatic image and multi-spectral image (red band), 1000×800 and 250×200 in size, respectively. They both come from Chinese GF1 satellite imagery. The resolution of panchromatic image is 2 m and that of multi-spectral image was 8 m, so there are a big scale difference between this image pair. The image pair was covered by lots of croplands, roads and buildings. However, many details cannot be manually identified due to the resolution difference. There were also significant nonlinear radiometric differences near the centre between images. In addition, there was also a big translation between them.

3.2. Implementation details

We chose the root mean squared error (RMSE) of the residuals of checkpoints (CPs) as the criterion to evaluate the quality of estimated transformation model (or registration accuracy). According to our prior knowledge of various multimodal matching methods mentioned in the introduction, we found that HOPC can generate many correct correspondences uniformly. In order to avoid the hassle of manually selecting CPs, we implemented HOPC on these multimodal image pairs and converted the correct matching points of HOPC to the original image as CPs. We must emphasize that the correct matching points of HOPC had passed manual re-examination, and those whose residuals are two pixels are considered mismatches. And then the residuals of CPs are calculated using the similarity transformation model estimated by LGEPC.

3.3. Performance of LGEPC

(1) Results of test 1

The original image pair of test 1 is shown in Figure 3(a) and Figure 3(b). Since there was no scale difference between the two images, in theory, the atlases phase correlation results of multi-scale atlas $\mathbf{S}_r = (\mathbf{S}_r^t, \mathbf{S}_r^m, \mathbf{S}_r^b)$ of the reference image and the multi-scale atlas $\mathbf{S}_s = (\mathbf{S}_s^t, \mathbf{S}_s^m, \mathbf{S}_s^b)$ of the sensed image had an important characteristic. That was that the more similar the structural information of each image pair, the more significant their phase correlation results (e.g. the correlation of $(\mathbf{S}_r^b, \mathbf{S}_s^b) >$ that of $(\mathbf{S}_r^b, \mathbf{S}_s^m) >$ that of $(\mathbf{S}_r^b, \mathbf{S}_s^t)$).

The atlases phase correlation results in solving the rotation and scale differences are shown in Figure 5, and the phase correlation result in solving the translation difference

is shown in Figure 6. As shown in Figure 5(g), Figure 5(h) and Figure 5(i), the results are basically in accordance with the above theoretical analysis. Besides, the peak in Figure 5(i) > that in Figure 5(e) > that in Figure 5(a) meant that the enhancement of the multi-scale structural information helps to improve the significance of the phase correlation. This experiment demonstrated that LGEPCC can successfully solve the rotation and scale differences without being affected by nonlinear radiometric differences. In addition, it also was confirmed that LGEPCC can successfully obtain a reliable translation because Figure 6 obtained by the phase correlation module is a standard pulse image with a significant peak whose coordinates specify the translation. In general, LGEPCC exhibited good robustness to image pair with nonlinear radiometric differences.

(2) Results of test 2

The original image pair of test 2 is shown in Figure 3(c) and Figure 3(d). The atlases phase correlation results of LGEPCC in solving the rotation and scale differences, as well as Reddy's (Reddy and Chatterji 1996), are shown in Figure 7(a-c), respectively. Both the results in Figure 7(a-b) successfully obtained the approximate rotation angle and scale factor. However, according to our method, the result in Figure 7(a) was deemed the final acceptable one because the peak in Figure 7(a) was larger than that in Figure 7(b). Besides, the ratio of the sub-maximum peak and the maximum peak in Figure 7(a) and Figure 7(b) is 0.74 and 0.88, respectively, and the former was within the given threshold t_{ratio} , while the latter was beyond t_{ratio} , so the result in Figure 7(b) was not reliable according (Reddy and Chatterji 1996). The phase correlation failure occurred in solving the correct rotation angle and scale factor as shown in Figure 7(c), which demonstrated that the correct results may not be extracted by directly conducting the phase correlation module when there were large scale differences between the image pair. This experiment verified that it was effective to build the multi-scale atlas space to enhance the phase correlation for resistance to the large scale differences between the image pairs.

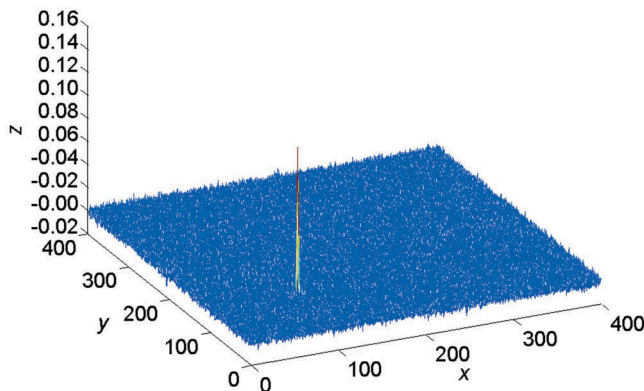


Figure 6. Result of phase correlation in solving the translation on test 1. (x, y) are the coordinates in the pulse image obtained by LGEPCC in solving the translation, and z is the peak corresponding to each (x, y) . The coordinates of the maximum peak (0.128) on the z -axis is (100, 100), which is the actual offset of the test data.

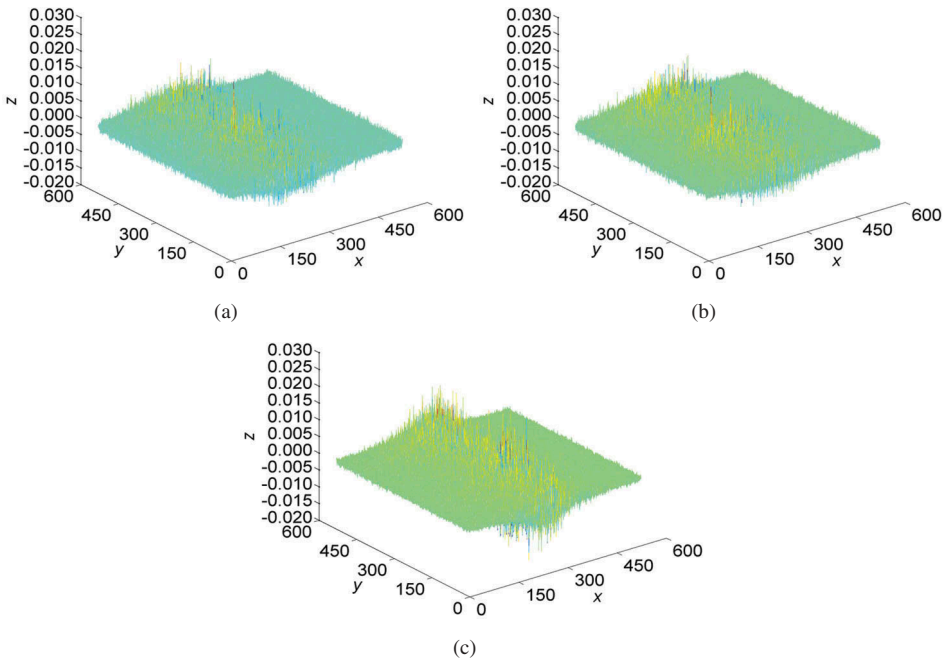


Figure 7. Results of atlases phase correlation on test 2. (a) phase correlation result of S_r^t and S_s^b . (b) phase correlation result of S_r^m and S_s^b . (c) phase correlation result of the original image pair using the method of (Reddy and Chatterji 1996). The meanings of x , y , z are the same as those in Figure 5. The transformation parameters are $(s_0 = 0.255, r_0 = 30.45^\circ)$, $(s_0 = 0.255, r_0 = 30.45^\circ)$, and $(s_0 = 0.844, r_0 = 57.18^\circ)$ in (a)-(c), respectively, for comparison. Obviously, the parameters are not correct in (c).

The experimental results of the two sets of synthetic images show that LGEPCC is an effective applicability extension of EPC. Furthermore, LGEPCC can deal with image pairs with significant nonlinear radiometric differences and large scale differences to a certain extent, and it can correctly calculate the similarity transformation model between the images.

3.4 Tests on multimodal remote sensing image pairs

3.4.1. Visualization and accuracy

Figure 8(a-c) show the registration results of LGEPCC, the details of which are displayed using staggered grids for visualization. For the Visible-to-Infrared, LGEPCC performed very well in terms of the displayed details shown in Figure 8(a). It demonstrated that the part two of LGEPCC that utilized the enhanced structural spectra to solve translation was as robust as the original EPC. For the LiDAR-to-Visible, there were a twice scale difference and a small rotation angle between them. The registration details shown in Figure 8(b) demonstrated that LGEPCC also achieved good registration accuracy in this case. For the Panchromatic-to-Multispectral, a big scale difference and translation existed between the image pair, and the original EPC cannot handle this case as the scale factor beyond its tolerance. However, LGEPCC can well solve their scale factor, which indirectly confirmed that the necessity and effectiveness of atlases phase correlation in part one of LGEPCC. In addition, it can be seen that the significant nonlinear radiometric differences are ubiquitous between these three image pairs, while

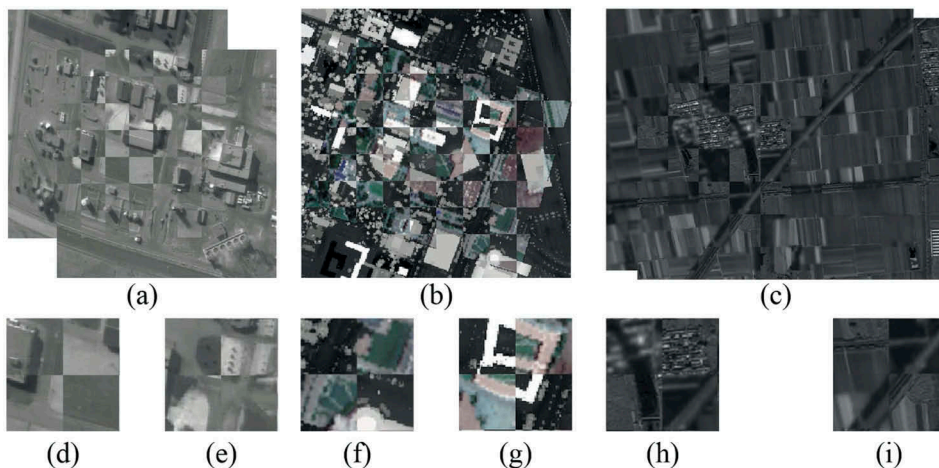


Figure 8. Registration results of LGEPC. (a) test 3. (b) test 4. (c) test 5. (d)-(e), (f)-(g) and (h)-(i) are the amplifications of the two local overlapped areas of test 3, test 4 and test 5, respectively. In (a)-(c), the local registration details of the two images in the overlapped areas are displayed using staggered grids. The sensed image is resampled to the coordinate system of the reference image and superposed with the later for visualization.

LGEPC was able to resist these differences very well. It is also worth noting that the re-sampling errors will inevitably occur in Test 4 and Test 5 when rectifying the sensed images due to the rotation-scale differences, which may affect the effect of visualization.

On the whole, LGEPC can resist nonlinear radiometric and geometric differences (including scale, rotation and translation) to a certain extent, and achieve a good level of registration accuracy for multimodal remote sensing images in terms of visualization.

As LGEPC is a kind of transformation model method in essence, it does not provide discrete matching points as the methods such as NCC or HOPC. Therefore, in order to quantitatively evaluate the registration accuracy of LGEPC, we converted the correct matching points of HOPC to the original images as CPs. And then the RMSE of the residuals of CPs was calculated as the evaluation criterion. The registration accuracy of the above experiments on multimodal remote sensing images is presented in Table 2. As shown, the registration accuracy of LGEPC can achieve one pixel for image pairs which well conform to the similarity transformation model. It must be emphasized that the correct matching points of HOPC were manually re-examined, and the mismatches whose residuals beyond 2 pixels were eliminated. As present, the max error of the residuals of CPs was within 2 pixels, which suggested that there was no obvious error in the transformation model obtained by LGEPC. The RMSEs of Test 4 and Test 5 were slightly larger than that of Test 3, and part of the reason may be attributed to re-sampling errors to some extent because rectifying the sensed images both occurred in the previous two cases.

Table 2. Registration accuracy of LGEPC.

| Image pair | Number of checkpoints | RMSE (pix.) | Max error (pix.) |
|--------------------------------------|-----------------------|-------------|------------------|
| Test 3-Visible-to-Infrared | 99 | 0.701 | 1.502 |
| Test 4-LiDAR-to-Visible | 142 | 0.961 | 1.670 |
| Test 5-Panchromatic-to-Multispectral | 141 | 0.945 | 1.414 |

3.4.2. Comparisons

The traditional FBMs such as SIFT suffer to extract highly repeatable shared features and describe their features, so that few correspondences can be to obtain and then transformation model estimation may fail. Therefore, without consideration of FBMs, we compared our method with some ABMs: NCC, MI, RMI and HOPC, utilizing once again test 3 in terms of registration accuracy. The reason for choosing test 3 is that the ABMs often force a similar scale between image pairs.

Our aim in this part was to illustrate our method as a good registration method as well, but not to demonstrate it as a replacement for the existing advanced methods. The experimental results are presented in Table 3. As can be seen, NCC achieved the worst registration accuracy among these methods. The reason for this is as follows. The area-based method using NCC metrics to detect the correspondences cannot effectively handle nonlinear radiometric differences between images, so that it affects the distribution and quality of the correspondences and then affects the transformation model estimation. MI obtained the best registration result in this case, while its variant RMI yielded almost identical transformation model with the exception of one-pixel difference in the x-direction. On the basis of the structural properties of multimodal images, HOPC adopted a histogram of oriented phase congruency to construct feature descriptors. Furthermore, the NCC of HOPC descriptors, i.e. $HOPC_{ncc}$ is utilized to register multimodal images based on a fast template matching scheme. In this experiment, HOPC achieved slightly lower registration accuracy than MI. Unlike HOPC which uses Log-Gabor to obtain phase congruency orientation as so to generate the feature descriptors, LGEPC uses the filtered results of multi-scale Log-Gabor to enhance the structural information at different scales. The registration accuracy of LGEPC is basically the same as that of HOPC, and the difference in translation is within 0.7 pixels. It is worth noting that the location accuracy of CPs obtained by HOPC is pixel-wise. Therefore, the registration accuracy must also be the pixel-wise in theory as long as the transformation model is strictly correct. That is to say that the registration results within one pixel can be both considered reliable. Therefore, both MI, RMI, HOPC and LGEPC obtained the correct transformation model in this test. In addition, the statistically weak advantage of LGEPC is not sufficient evidence that it is better or worse than the compared methods.

3.5. Failure case and limitations

The LGEPC algorithm is still essentially an extended phase correlation algorithm, which forces a similarity transformation model between image pairs. However, the complex transformation models between image pairs such as non-similarity transformation may lead to its failure. We will give an example under this case.

The test data was visible spectral (Google Earth) and SAR imaging (TerraSAR-X), 528×520 and 320×320 in size, respectively, as shown in Figure 9(a-b). The images contained

Table 3. Transformation model and registration accuracy of NCC, MI, RMI, HOPC and LGEPC on test 3.

| Method | Scale | Rotation ($^{\circ}$) | Translation (pix.) | RMSE (pix.) | Max error (pix.) |
|--------|--------|-------------------------|--------------------|-------------|------------------|
| NCC | 1.0082 | -0.120 | (72.188, 91.240) | 1.197 | 2.116 |
| MI | 1.0000 | 0.000 | (71.000, 89.000) | 0.284 | 1.000 |
| RMI | 1.0000 | 0.000 | (70.000, 89.000) | 1.000 | 1.414 |
| HOPC | 1.0002 | 0.054 | (70.748, 89.258) | 0.691 | 0.871 |
| LGEPC | 1.0000 | 0.000 | (70.512, 88.579) | 0.701 | 1.502 |

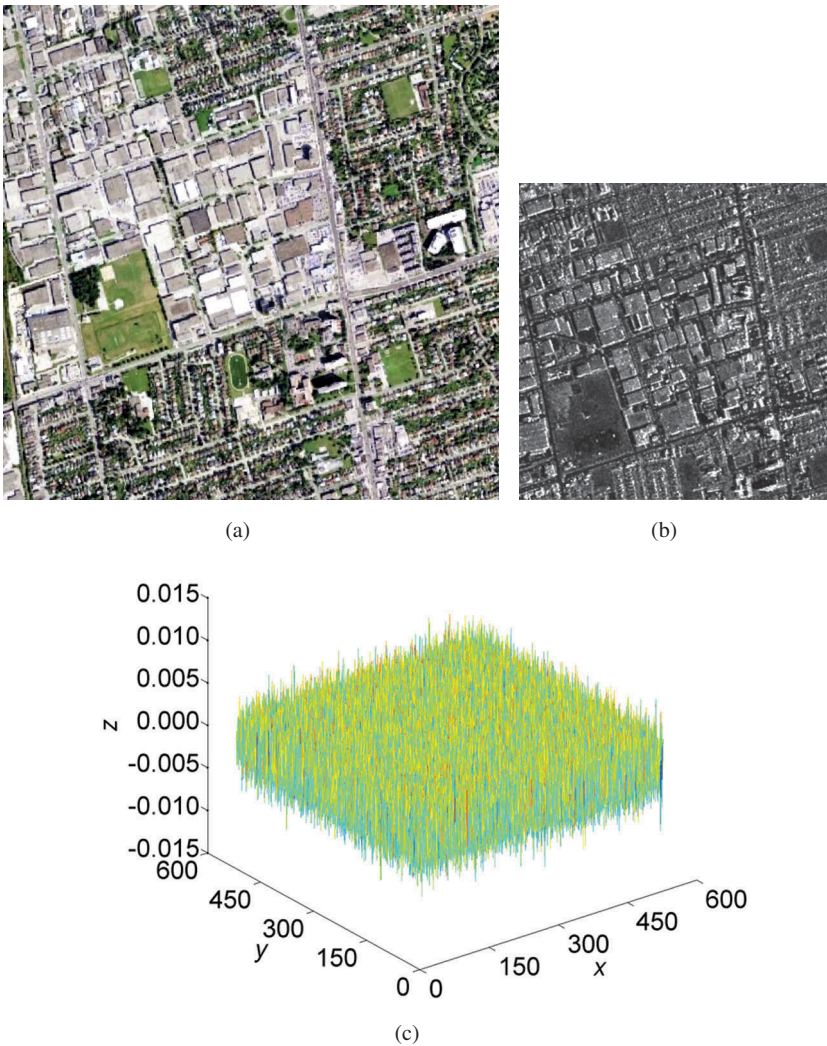


Figure 9. A failure case. (a) reference image (visible). (b) sensed image (SAR). (c) phase correlation result in solving translation via LGEPc. The meanings of x , y , z are the same as those in Figure 6.

mainly urban areas covered by dense artificial buildings. Compared to the visible spectral image, there was a significant noise on the SAR image that made the building roofs on the SAR image appear less smooth than those on the visible spectral image. In addition, there have significant local deformation caused by relief displacement of buildings.

The phase correlation result in solving translation via LGEPc is shown in Figure 9(c). As shown, the translation cannot be explicitly determined because the maximum response peak was not significant. This may be mainly caused by the following two reasons. On one hand, there are more noise on the SAR image than the visible spectral image, leading to a difference in the geometric structural information extracted by LGEPc using Log-Gabor filtering as shown in Figure 10. On the other hand, although the SAR image and the visible spectral image conformed to a similar transformation model on a global scale, there was still a scale difference in the local region because of

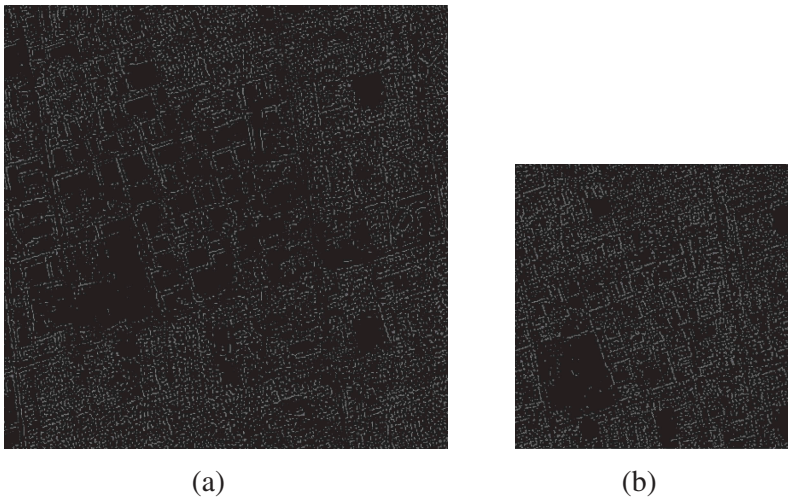


Figure 10. Results of Log-Gabor filtering on the image pair shown in Figure 9(a-b). (a) reference image. (b) sensed image. The results show the results of Log-Gabor filtering of the reference image and the sensed image corresponding to the shortest wavelength $\lambda_{\min} = 3$. Compared with the results of Log-Gabor filtering of the reference image, the filtering results of the sensed image have many noises, causing a difference in the structural information between them.

the inconsistent scale caused by the radar imaging mode. Meanwhile, the tilted imaging viewpoint of SAR image also leads to the relief displacement of buildings. And the second argument demonstrated that we cannot utilize a simple translation model to express their relationship under the situation with local distortions. And this situation cannot be solved by using an image-to-image registration method, such as EPC or LGEPC, until an orthorectification is applied to the image with local distortions.

In addition, the layer number of our multi-scale atlases space is three, which does not guarantee that our algorithm is suitable for the image pair with any scale difference. Also, the optimal multi-scale atlas building will be most consideration in our future work.

4. Conclusions and discussions

In this paper, we proposed a novel extended phase correlation algorithm based on Log-Gabor filtering, called LGEPC, for multimodal remote sensing image registration. LGEPC mainly focuses on the issue of significant nonlinear radiometric and large scale differences that traditional extended phase correlation cannot well handle. First, the multi-scale atlas spaces of the reference and sensed images, which are both composed of different geometric structural information, are respectively built based on their magnitudes obtained by using Log-Gabor filtering with different central frequencies. And a process of atlases phase correlation is conducted using phase correlation module, among which we obtain the optimal result that determines the scale factor and rotation angle between image pairs. Then, we rectify the sensed image using the rotation and scale parameters. Thereafter, the multi-scale filtered structural spectra of the reference and rectified sensed images, which are obtained by using the Log-Gabor filtering with

different central frequencies, were superimposed to obtain the final structural spectra. This can eliminate the interference of radiometric differences in solving translation when using the phase correlation module as much as possible. Our qualitative and quantitative experiments demonstrated that LGEPC was very resistant to significant nonlinear radiometric differences and geometric differences and can meet some needs of multimodal remote sensing image registration tasks. For image pairs that conform to the similarity transformation model, the registration accuracy of LGEPC can ideally reach within one pixel. However, the registration accuracy of LGEPC may be unsatisfactory and even failure may occur for image pairs with more complex non-similarity transformation model.

Since LGEPC depends on geometric structural information, the performance of LGEPC may decline if the image pairs contain few structures. In this situation, an image enhancement method would be conducive to enhance the edge or contour features of the images in order to be beneficial to their registration via LGEPC. However, the situation that lacks structures may also lead its failure. In addition, more quantitative comparison experiments and evaluations between LGEPC and other state-of-the-art algorithms should be addressed in the future research using more multimodal remote sensing images.

Acknowledgments

We would like to thank Ye's public MATLAB code for implementing HOPC.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was supported in part by the National Key R&D Program of China under Grant No. 2017YFB0503004, in part by the National Natural Science Foundation of China with project number 41571434 and 41322010.

References

- Anuta, P. E. 1970. "Spatial Registration of Multispectral and Multitemporal Digital Imagery Using Fast Fourier Transform Techniques." *IEEE Transactions on Geoscience Electronics* 8 (4): 353–368. doi:10.1109/TGE.1970.271435.
- Arbelaez, P., M. Maire, C. Fowlkes, C. Fowlkes, and J. Malik. 2011. "Contour Detection and Hierarchical Image Segmentation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 33 (5): 898–916. doi:10.1109/TPAMI.2010.161.
- Bentoutou, Y., N. Taleb, K. Kpalma, and J. Ronsin. 2005. "An Automatic Image Registration for Applications in Remote Sensing." *IEEE Transactions on Geoscience and Remote Sensing* 43 (9): 2127–2137. doi:10.1109/TGRS.2005.853187.
- Bouchiha, R., and K. Besbes. 2013. "Automatic Remote-Sensing Image Registration Using Surf." *International Journal of Computer Theory and Engineering* 5 (1): 88–92. doi:10.7763/IJCTE.2013.V5.653.

- Cai, G., P. M. Jodion, S. Li, Y. Wu, S. Su, and Z. Huang. 2013. "Perspective-Sift: An Efficient Tool for Low-Altitude Remote Sensing Image Registration." *Signal Process* 93 (11): 3088–3110. doi:10.1016/j.sigpro.2013.04.008.
- Chen, C., Y. Li, W. Liu, and J. Huang. 2014. "Image Fusion with Local Spectral Consistency and Dynamic Gradient Sparsity." *IEEE Conference on Computer Vision and Pattern Recognition*: 2760–2765. doi:10.1109/CVPR.2014.347.
- Chen, H.-M., M. K. Arora, and P. K. Varshney. 2003. "Mutual Information-Based Image Registration for Remote Sensing Data." *International Journal of Remote Sensing* 24 (18): 3701–3706. doi:10.1080/0143116031000117047.
- Chen, Q. S., M. Defrise, and F. Deconinck. 1994. "Symmetric Phase-Only Matched Filtering of Fourier-Mellin Transforms for Image Registration and Recognition." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 16 (12): 1156–1168. doi:10.1109/34.387491.
- Chen, S., X. Li, L. Zhao, and H. Yang. 2018. "Medium-Low Resolution Multisource Remote Sensing Image Registration Based on SIFT and Robust Regional Mutual Information." *International Journal of Remote Sensing* 39 (10): 3215–3242. doi:10.1080/01431161.2018.1437295.
- Dai, J., W. Song, and Y. Li. 2014. "Progressive SIFT Matching Algorithm for Multi-Source Optical Satellite Images." *Acta Geodaetica Et Cartographica Sinica* 43 (7): 746–752. doi:10.13485/j.cnki.11-2089.2014.00.
- Daugman, J. G. 1985. "Uncertainty Relation for Resolution in Space, Spatial Frequency, and Orientation Optimized by Two-Dimensional Visual Cortical Filters." *Journal of the Optical Society of America A-Optics Image Science and Vision* 2 (7): 1160–1169. doi:10.1364/JOSAA.2.001160.
- Dawn, S., V. Saxena, and B. Sharma. 2010. "Remote Sensing Image Registration Techniques: A Survey." In *Paper presented at 4th International Conference on Image and Signal Processing*, Trois Rivieres, Canada, Jun 30-Jul 02.
- Decastro, E., and C. Morandi. 1987. "Registration of Translated and Rotated Images Using Finite Fourier Transforms." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 5: 700–703. doi:10.1109/TPAMI.1987.4767966.
- Elder, J. H., and S. W. Zucker. 1998. "Local Scale Control for Edge Detection and Blur Estimation." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 20 (7): 699–716. doi:10.1109/34.689301.
- Feichtinger, H. G., and T. Strohmer. 2012. *Gabor Analysis and Algorithms: Theory and Applications*. Springer Science and Business Media. Birkhäuser, Boston, Inc. Secaucus, NJ, USA.
- Field, D. J. 1987. "Relations between the Statistics of Natural Images and the Response Properties of Cortical Cells." *Journal of the Optical Society of America A-Optics Image Science and Vision* 4 (12): 2379–2394. doi:10.1364/JOSAA.4.002379.
- Foroosh, H., J. B. Zerubia, and M. Berthod. 2002. "Extension of Phase Correlation to Subpixel Registration." *IEEE Transactions on Image Processing* 11 (3): 188–200. doi:10.1109/83.988953.
- Gabor, D. 1946. "Theory of Communication. Part 1: The Analysis of Information." *Journal of the Institution of Electrical Engineers-Part III, Radio and Communication Engineering* 93 (26): 429–441. doi:10.1049/ji-3-2.1946.0074.
- Ge, P., C. Y. Lan, and H. Wang. 2014. "An Improvement of Image Registration Based on Phase Correlation." *Optik - International Journal for Light and Electron Optics* 125 (22): 6709–6712. doi:10.1016/j.ijleo.2014.07.086.
- Gharbia, R., S. A. Ahmed, and A. E. Hassanien. 2015. "Remote Sensing Image Registration Based on Particle Swarm Optimization and Mutual Information." *Springer India* 127 (1): 399–408.
- Ghassabi, Z., J. Shanbehzadeh, A. Sedaghat, and E. Fatemizadeh. 2013. "An Efficient Approach for Robust Multimodal Retinal Image Registration Based on UR-SIFT Features and PIIFD Descriptors." *Eurasip Journal on Image and Video Processing* 2013 (1): 1–16. doi:10.1186/1687-5281-2013-25.
- Gong, M. G., S. M. Zhao, L. C. Jiao, D. Y. Tian, and S. Wang. 2014. "A Novel Coarse-to-Fine Scheme for Automatic Image Registration Based on SIFT and Mutual Information." *IEEE Transactions on Geoscience and Remote Sensing* 52 (7): 4328–4338. doi:10.1109/TGRS.2013.2281391.
- Gonzalez, R. C., and P. Wintz. 1987. *Digital Image Processing*. New York, NY, USA: Addison-Wesley.

- Inglada, J., V. Muron, D. Pichard, and T. Feuvrier. 2007. "Analysis of Artifacts in Subpixel Remote Sensing Image Registration." *IEEE Transactions on Geoscience and Remote Sensing* 45 (1): 254–264. doi:[10.1109/TGRS.2006.882262](https://doi.org/10.1109/TGRS.2006.882262).
- Kim, T., and Y. J. Im. 2003. "Automatic Satellite Image Registration by Combination of Matching and Random Sample Consensus." *IEEE Transactions on Geoscience and Remote Sensing* 41 (5): 1111–1117. doi:[10.1109/TGRS.2003.811994](https://doi.org/10.1109/TGRS.2003.811994).
- Kokila, R., and P. Thangavel. 2014. "Image Registration Based on Fast Fourier Transform Using Gabor Filter." *International Journal of Computer Science and Electronics Engineering* 2 (1): 31–37. ISSN: 2320-401X; EISSN: 2320-4028.
- Kovesi, P. 1999. "Image Features from Phase Congruency." *Videre* 1 (3): 1–26.
- Kuglin, C. D., and D. C. Hines. 1975. "The Phase Correlation Image Alignment Method." In *In Proceeding of IEEE International Conference on Cybernetics and Society*, 163–165. New York, USA.
- Lades, M., J. Vorbruggen, J. Buhmann, J. Lange, C. Vndermalsburg, R. Wurtz, and W. Konen. 1993. "Distortion Invariant Object Recognition in the Dynamic Link Architecture." *IEEE Transactions on Computers* 42 (3): 300–311. doi:[10.1109/12.210173](https://doi.org/10.1109/12.210173).
- Lee, T. S. 1996. "Image Representation Using 2d Gabor Wavelets." *IEEE Transactions on Pattern Analysis and Machine Intelligence* 18 (10): 959–971. doi:[10.1109/34.541406](https://doi.org/10.1109/34.541406).
- Liang, J., X. Liu, K. Huang, X. Li, D. Wang, and X. Wang. 2014. "Automatic Registration of Multisensor Images Using an Integrated Spatial and Mutual Information (SMI) Metric." *IEEE Transactions on Geoscience and Remote Sensing* 52 (1): 603–615. doi:[10.1109/TGRS.2013.2242895](https://doi.org/10.1109/TGRS.2013.2242895).
- Lowe, D. G. 2004. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journal of Computer Vision* 60 (2): 91–110. doi:[10.1023/B:VISI.0000029664.99615.94](https://doi.org/10.1023/B:VISI.0000029664.99615.94).
- Reddy, B. S., and B. N. Chatterji. 1996. "An FFT-based Technique for Translation, Rotation, and Scale-Invariant Image Registration." *IEEE Transactions on Image Processing* 5 (8): 1266–1271. doi:[10.1109/83.506761](https://doi.org/10.1109/83.506761).
- Sarvaiya, J. N., S. Patnaik, and K. Kothari. 2012. "Image Registration Using Log Polar Transform and Phase Correlation to Recover Higher Scale." *Journal of Pattern Recognition Research* 7: 90–105. doi:[10.13176/11.355](https://doi.org/10.13176/11.355).
- Sedaghat, A., and H. Ebadi. 2015. "Distinctive Order Based Self-Similarity Descriptor for Multi-Sensor Remote Sensing Image Matching." *ISPRS Journal of Photogrammetry and Remote Sensing* 108: 60–71. doi:[10.1016/j.isprsjprs.2015.06.003](https://doi.org/10.1016/j.isprsjprs.2015.06.003).
- Sedaghat, A., M. Mokhtarzade, and H. Ebadi. 2011. "Uniform Robust Scale Invariant Feature Matching for Optical Remote Sensing Images." *IEEE Transactions on Geoscience and Remote Sensing* 49 (11): 4516–4527. doi:[10.1109/TGRS.2011.2144607](https://doi.org/10.1109/TGRS.2011.2144607).
- Serrano, I. M. de Diego, C. Conde, and E. Cabello. 2010. "Recent Advances in Face Biometrics with Gabor Wavelets: A Review." *Pattern Recognition Letters* 31 (5): 372–381. doi:[10.1016/j.patrec.2009.11.002](https://doi.org/10.1016/j.patrec.2009.11.002).
- Sun, Y. B., L. Zhao, S. D. Huang, L. Yan, and G. Dissanayake. 2014. "L2-SIFT: SIFT Feature Extraction and Matching for Large Images in Large-Scale Aerial Photogrammetry." *ISPRS Journal of Photogrammetry and Remote Sensing* 91: 1–16. doi:[10.1016/j.isprsjprs.2014.02.001](https://doi.org/10.1016/j.isprsjprs.2014.02.001).
- Szeliski, R. 2010. *Computer Vision: Algorithms and Applications*. Part of the Texts in Computer Science book series (TCS). London: Springer. doi: [10.1007/978-1-84882-935-0](https://doi.org/10.1007/978-1-84882-935-0).
- Tong, X. H., Z. Ye, Y. S. Xu, S. J. Liu, L. Li, H. Xie, and T. Li. 2015. "A Novel Subpixel Phase Correlation Method Using Singular Value Decomposition and Unified Random Sample Consensus." *IEEE Transactions on Geoscience and Remote Sensing* 53 (8): 4143–4156. doi:[10.1109/TGRS.2015.2391999](https://doi.org/10.1109/TGRS.2015.2391999).
- Vural, M. F., Y. Yardimci, and A. Temzel. 2009. "Registration of Multispectral Satellite Images with Orientation-Restricted Sift." Paper presented at IEEE International Geoscience and Remote Sensing Symposium, Cape Town, South Africa, July 12-17.
- Wan, Y., and Y. Zhang. 2017. "The P2L Method of Mismatch Detection for Push Broom High-Resolution Satellite Images." *ISPRS Journal of Photogrammetry and Remote Sensing* 130: 317–328. doi:[10.1016/j.isprsjprs.2017.06.009](https://doi.org/10.1016/j.isprsjprs.2017.06.009).

- Wong, A., and D. A. Clausi. 2007a. "ARRSI: Automatic Registration of Remote Sensing Images." *IEEE Transactions on Geoscience and Remote Sensing* 45 (5): 1483–1493. doi:[10.1109/TGRS.2007.892601](https://doi.org/10.1109/TGRS.2007.892601).
- Wong, A., and D. A. Clausi. 2010b. "AISIR: Automated Inter-Sensor/Inter-Band Satellite Image Registration Using Robust Complex Wavelet Feature Representations." *Pattern Recognition Letters* 31 (10): 1160–1167. doi:[10.1016/j.patrec.2009.05.016](https://doi.org/10.1016/j.patrec.2009.05.016).
- Xu, Q., Y. Zhang, and B. Li. 2014. "Improved Sift Match for Optical Satellite Images Registration by Size Classification of Blob-Like Structures." *Remote Sensing Letters* 5 (5): 451–460. doi:[10.1080/2150704X.2014.917774](https://doi.org/10.1080/2150704X.2014.917774).
- Ye, Y., J. Shan, L. Bruzzone, and L. Shen. 2017. "Robust Registration of Multimodal Remote Sensing Images Based on Structural Similarity." *IEEE Transactions on Remote Sensing* 55 (5): 2941–2958. doi:[10.1109/TGRS.2017.2656380](https://doi.org/10.1109/TGRS.2017.2656380).
- Ye, Y., and L. Shen. 2016. "Hopc: A Novel Similarity Metric Based on Geometric Structural Properties for Multi-Modal Remote Sensing Image Matching." *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences* 3 (1): 9–16. doi:[10.5194/isprsannals-III-1-9-2016](https://doi.org/10.5194/isprsannals-III-1-9-2016).
- Zhang, J., M. Zareapoor, X. He, D. Shen, D. Feng, and J. Yang. 2018. "Mutual Information Based Multi-Modal Remote Sensing Image Registration Using Adaptive Feature Weight." *Remote Sensing Letters* 9 (7): 646–655. doi:[10.1080/2150704X.2018.1458343](https://doi.org/10.1080/2150704X.2018.1458343).
- Zhao, L. Y., B. Y. Lv, X. R. Li, and S. H. Chen. 2015. "Multi-Source Remote Sensing Image Registration Based on Scale-Invariant Feature Transform and Optimization of Regional Mutual Information." *Acta Physica Sinica* 64 (12): 190–200. doi:[10.7498/aps.64.124204](https://doi.org/10.7498/aps.64.124204).
- Zitova, B., and J. Flusser. 2003. "Image Registration Methods: A Survey." *Image and Vision Computing* 21 (4): 977–1000. doi:[10.1016/S0262-8856\(03\)00137-9](https://doi.org/10.1016/S0262-8856(03)00137-9).