Advances and Opportunities in Remote Sensing Image Geometric Registration

A systematic review of state-of-the-art approaches and future research directions

RUITAO FENG, HUANFENG SHEN, JIANJUN BAI, AND XINGHUA LI

Geometric registration is often an accuracy assurance for most remote sensing image processing and analysis, such as image mosaicking, image fusion, and time-series analysis. In recent decades, geometric registration has attracted considerable attention in the remote sensing community, leading to a large amount of research on the subject. However, few studies have systematically reviewed its current status and deeply investigated its development trends. Moreover, new approaches are constantly emerging, and some issues still need to be solved. Thus, this article presents a survey of state-of-the-art approaches for remote sensing image registration in terms of intensity-based, feature-based, and combination techniques. Optical flow estimation and deep learning-based methods are summarized, and software-operated registration and registration evaluation are introduced. Building on recent advances, promising opportunities are explored.

OVERVIEW

Remote sensing images from various sensors, periods, and viewpoints can provide complementary information about regions of interest (ROIs) and Earth surface observation. Owing to various factors, such as Earth's rotation and curvature and variations in platform altitudes, remote sensing images contain systematic geometric distortions that cannot be thoroughly corrected without high-precision elevation data [through the digital elevation model (DEM) or the digital surface model (DSM)] and control points on the ground. Although the true digital orthophoto map (TDOM) promises accurate spatial positions, it has high production costs and is difficult for general users to obtain. Therefore, most available remote sensing images retain small geometrical distortions after systematic correction, resulting in objects in one image not spatially corresponding to those in another image, as in Figure 1.

Furthermore, topographical fluctuations in mountainous regions, differences in imaging viewpoints (shown in Figure 2), and spatial resolutions cause dislocation in two
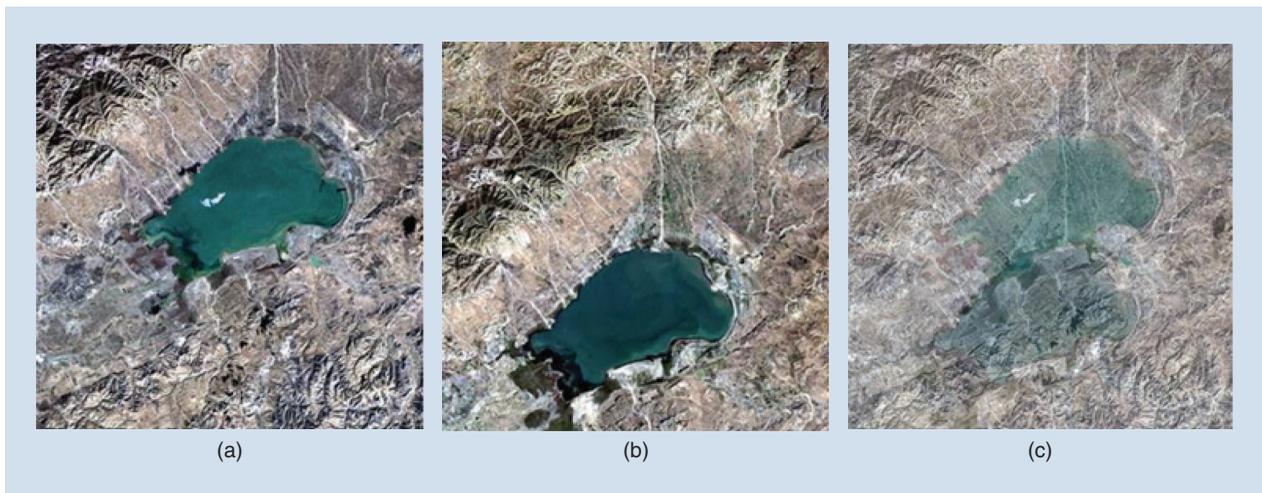
FIGURE 1. Multitemporal optical image geometrical dislocation. (a) A reference image taken by *Landsat 5* on 15 October 1990. (b) A sensed image taken by *Landsat 5* on 15 September 1993. (c) The overlapping images of (a) and (b).

images covering the same scene. Thus, geometrical registration techniques are implemented to align two or more images from the image-to-image perspective rather than the imaging mechanism. Consequently, geometrical registration is an image-processing technique that aligns different images of the same scene acquired at various times and viewing angles and with multiple sensors [1]. As a fundamental task in remote sensing information processing, it is a prerequisite for many practical applications, such as image mosaicking [2], image fusion [3], land cover change detection [4], [5], and disaster evaluation [6], [7].

It worth noting that there is a technical term, *coregistration*, that is similar but not exactly the same as image registration. It is now commonly used in aerial and unmanned aerial vehicle image registration, generally including multimode registration and alignment through the aid of auxiliary data. When the registration is conducted with a GPS/inertial measurement unit, it usually establishes a connection between an image and the simulated or real ground [8]. Certainly, the registration technology works on tie points generation for the construction of relationships. With real ground control points (GCPs), the tie points between the reference and sensed bands are produced to register different bands of hyperspectral images [9]. Additionally, when the orientation of the reference image is determined, without GCPs, the coregistration of multitemporal high-resolution image blocks is automatically achieved [10]. Although there are time-increasing papers focused on coregistration techniques doing some auxiliary work with the positioning data, the core of the process is image registration, as far as we are concerned. Therefore, the emphasis is put on the opportunities and challenges of geometrical registration in remote sensing fields.

Geometrical registration can be traced to the 1970s, when the United States proposed image registration to analyze target objects in aircraft-aided navigation and weapons systems. Since then, it has rapidly developed, particularly in the domains of remote sensing, computer vision,

and medical image processing. Some conclusive studies of computer vision and medical image processing have been published [11]–[16]. Building on a widespread survey of image registration, published in 1992 by Brown [15], a 2003 review [16] comprehensively summarized the subsequent research. In recent years, several overviews of image registration have focused on newly developed approaches inspired by extant versions [17]–[19]. However, these surveys are limited to analyzing and drawing conclusions based on conventional approaches [20]–[22]. Since the first study of multispectral and multitemporal digital imagery registration in 1970 [23], an increasing number of papers have contributed to the field. A total of 140,983 related studies with the keywords *image registration* or *image matching* were retrieved, from 1979 to January 2021, from Web of Science (WoS). When screening again using the keyword *remote sensing*, 46,141 articles were found, as plotted, based on their publication year, in Figure 3. The respective proportions of the total number of papers on WoS per year are also presented. It can be seen that a small number of papers
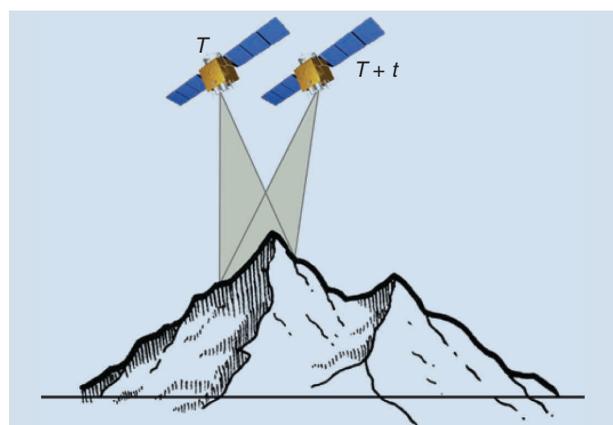

FIGURE 2. The angle difference from multitemporal images in a mountainous region.

were presented early in the field's development, with remote sensing image registration accounting for a minimal percentage of annual WoS publications. More recently, a considerable number of studies have been published, peaking in 2019. Thus, comprehensive analysis is necessary to identify unsolved problems for the rapid development of this field.

In this article, we summarize various classical approaches to remote sensing image registration as well as recent methods based on deep learning, optical flow estimation, and image registration software. We also point out interesting aspects and analyze development trends from our perspective, without describing specific approaches in detail. Concretely, the registration approaches can be classified into three categories, namely, intensity-based,

feature-based, and combination registration, as detailed in Figure 4. The intensity-based technique directly uses pixel intensity information to register images, including the conventional area-based approach and optical flow estimation. The geometrical and advanced features used to register images instead of intensity information are defined as feature-based approaches. Combination registration mainly consists of the integration of feature- and area-based methods as well as two geometric feature-based techniques. Many detailed classifications are presented in each category.

All registration approaches must undergo coordinate transformation and resampling to ultimately acquire the aligned image, as demonstrated in Figure 5. Before this step, transformation models for coordinate recalculation other than optical flow estimation should be constructed. In general, transformation models, such as the affine, projective, piecewise linear, and thin spline models, are derived from global or local parametric models. To calculate these models, images are preprocessed to extract representative features through techniques including geometrical- and advanced-feature extraction and matching. Given that intensity information is directly utilized in area-based registration, feature extraction is omitted, and the transformation model is constructed when matching the intensity information. Since most approaches prefer to contribute to the preliminary steps (e.g., feature extraction, feature matching, and mismatched feature elimination) rather
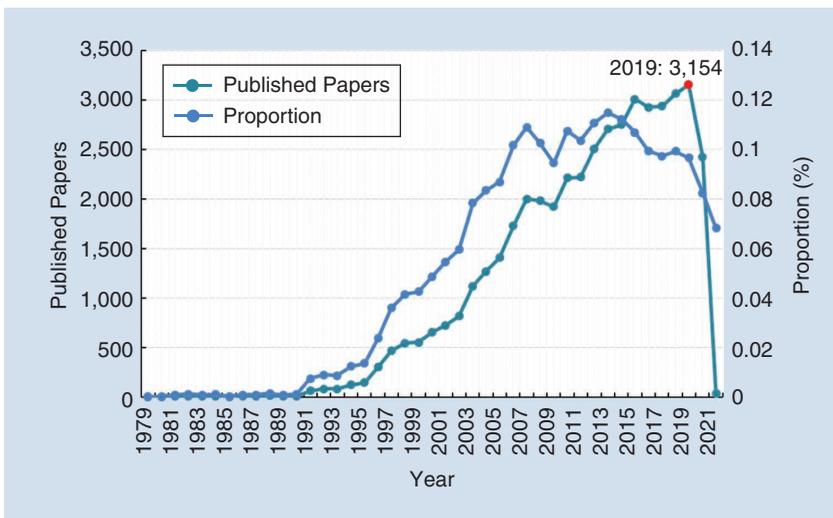

FIGURE 3. The number of papers about remote sensing image registration on WoS, per year.
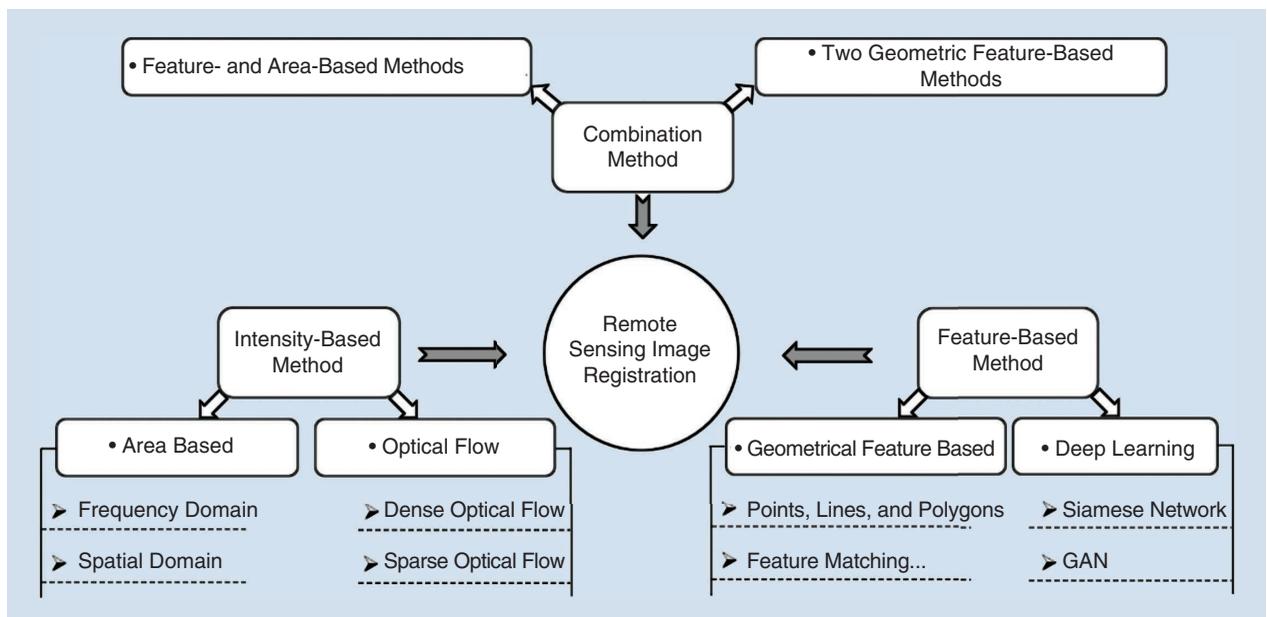

FIGURE 4. The remote sensing image registration algorithms. GAN: generative adversarial network.

than designing new transformation models and presenting novel resampling techniques, this article emphasizes the previous steps, as well, comprehensively summarizing studies and further predicting development trends.

INTENSITY-BASED REGISTRATION

Intensity-based registration directly employs original or extended intensity information, such as gradients, for registering remote sensing images. In addition to the traditional area-based approach, we classify optical flow estimation, a direct calculation of the increased displacement of corresponding pixels with intensity information, as intensity-based registration.

AREA-BASED METHOD

In general, area-based registration accords with a similarity criterion established in advance and adopts the optimal search strategy to iteratively find the parameters of the transformation model that yield the maximum or minimum similarity measurement to achieve the spatial registration of images, as illustrated in Figure 6. With the transformation model constantly being optimized, the aligned image changes gradually, which is mainly reflected in the growing black area in the lower- and upper-left-hand corners of the aligned image. This approach differs from image matching, which is generally understood as template matching. Although both methods directly employ intensity information, template matching aims to extract the centroids of matched windows as a feature point. This process is not true geometric registration, but it constitutes an important step. Here, we introduce area-based registration. The well-known core of this technique is the similarity metric, which has been researched in terms of spatial- and frequency-domain approaches [16], [24], [25].

SPATIAL-DOMAIN APPROACH

Spatial-domain techniques directly employ intensity difference and statistical information of all pixels, without any image transformation. These methods generally come at the problem from one of two perspectives, namely, the correlation-like technique or the mutual information (MI) algorithm.

CORRELATION-LIKE SIMILARITY METRIC

This technique determines the spatial alignment of images by directly comparing the similarity of corresponding pixels. It is vulnerable to intensity changes, which may be introduced, for instance, by noise, thick or thin clouds, and differences in the photosensitive components of various sensors. As a fundamental similarity metric, the cross-correlation (CC) algorithm directly calculates the difference between corresponding pixels to iteratively register images until they have the largest CC, which is useful for small rigid-body and affine transformation [26], [27]. Many other correlation-like similarity metrics are available, including the sequential similarity detection algorithm [28], correlation coefficient [29], [30], normalized CC (NCC) [31]–[33], sum of squared differences [34], Hausdorff distance [35], and other minimum distance criteria. NCC, in particular, is very popular and widely applied due to its invariance to linear intensity variations [31], [36], [37]. Recently, the centers of windows well-matched by NCC have been used as feature points to solve transformation model parameters [38], namely, image matching. Supposing $\rho(R, S)$ to be the NCC coefficient of matched windows, we calculate NCC as follows:

$$\rho(R, S) = \frac{\sum\limits_{i=1}^{m \times n} (R(i) - \mu_R)(S(i) - \mu_S)}{\sqrt{\sum\limits_{i=1}^{m \times n} (R(i) - \mu_R)^2} \sqrt{\sum\limits_{i=1}^{m \times n} (S(i) - \mu_S)^2}}, \quad (1)$$

where the predefined window consists of $m \times n$ pixels, $R(i)$ and $S(i)$ denote specified positions in the windows of the reference and sensed images, and μ_R and μ_S are the average intensity values of a specified window. The algorithm was developed to generate tie points that resist complicated geometric deformation [31], [38], [39]; it has recently been integrated with a novel feature descriptor [e.g., the local self-similarity (LSS) descriptor] for robust feature extraction in multimodal remote sensing image registration [36]. Although NCC is superior to the traditional correlation-like similarity metric, it is unable to handle the nonlinear radiometric difference, which is a common problem for correlation-like similarity metrics.

MI APPROACH

MI has appeared recently compared with correlation-like techniques; it has been successfully applied to multispectral and multisensor image registration due to its robustness against nonlinear radiation differences [40]–[43], which are usually calculated by (2). The normalized MI (NMI) method is a measure that is independent of changes in the marginal entropies of two images in their region of overlap [44], [45]. MI and NMI are the same type of statistical similarity measurement, and both are prone to registration errors. Inspired by these approaches, the region–MI approach was developed [46] with consideration of structural information.
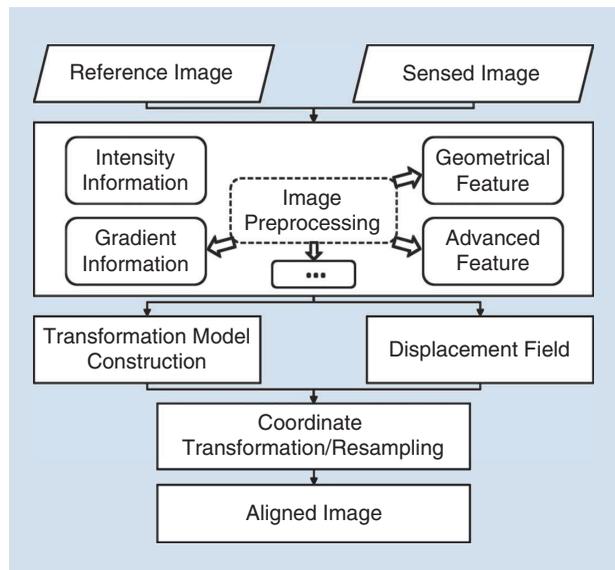

FIGURE 5. General geometrical registration.

Furthermore, rotationally invariant regional MI considers not only the spatial information but also the influence that local gray variations and rotation changes have on the computation of the probability density function [45]:

$$MI(R,S) = H(R) + H(S) - H(R,S),$$
$$H(R) = -\sum_{r \in R} P(r) \log_2 P(r),$$
$$H(S) = -\sum_{s \in S} P(s) \log_2 P(s),$$
$$H(R,S) = -\sum_{r \in R, s \in S} P(r,s) \log_2 P(r,s), \qquad (2)$$

where $H(R)$ and $H(S)$ are the Shannon entropies of the reference and sensed images, respectively; $H(R,S)$ represents the mutual entropy; $P(r)$ and $P(s)$ are the marginal probability distributions of R and S; and $P(r,s)$ is the joint probability distribution that is calculated, in practice, by 2D histogram binning as the discrete random variables. Additionally, there is an MI registration based on displacement maps, which is similar to optical flow estimation. In this variational framework, MI is employed as the similarity metric for displacement calculation [47]. Overall, the MI-like algorithms originating from information theory are a measure of the statistical

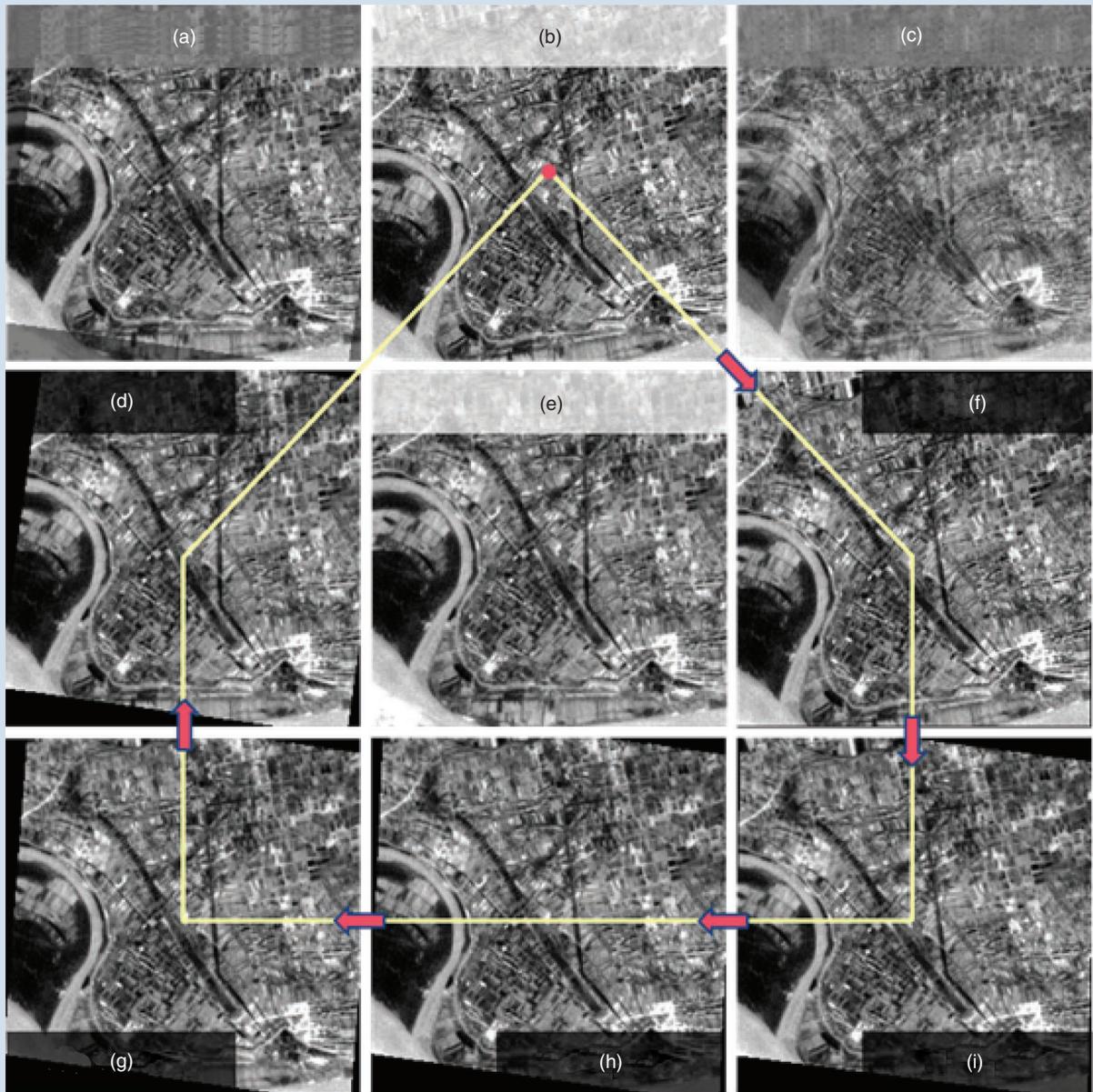**FIGURE 6.** Conventional area-based registration. Pay attention to how the black-edge region changes in the lower- and upper-left corners of the aligned image. (a) The aligned images overlapping. (b) The sensed image. (c) The original images overlapping. (d) The fifth iteration. (e) The reference image. (f) The first iteration. (g) The fourth iteration. (h) The third iteration. (i) The second iteration.

dependence between two data sets and particularly suitable for registration with different imaging mechanisms. However, they are computationally expensive, which may be restrictive, as remote sensing images are always relatively large.

FREQUENCY-DOMAIN APPROACHES

Frequency-domain approaches indirectly utilize intensity information, transforming an image and exploiting its frequency-domain features for registration. By so doing, they accelerate the computational speed of relatively small geometric dislocations. Fourier techniques are typical representations of frequency-domain registration, which were first used to register images with translational changes [48]. Phase-based correlation approaches [23], [49]–[51] exploit the Fourier transform to register images by searching for global optimal matching [53]; they compute the cross-power spectra of the sensed and reference images and seek the location of the peak. The translational and rotational properties of the Fourier transform are employed to calculate the transformation parameters [53]. Frequency domain approaches are robust against frequency-dependent noise and illumination changes. They also contribute to the acceleration of computational efficiency [54] since they neither involve feature extraction, as feature-based approaches do, nor require an optimization approach in the spatial domain, which would increase their computational complexity [53]. However, given that the Fourier transform offers poor spatial localization, the operation can be replaced by a wavelet transform with strong spatial and frequency localization [55], which can be applied to remote sensing image registration [56]. Recently, phase congruency (PC) has been used to represent structural information in remote sensing images; it is similar to the image gradient but is invariant in terms of image contrast and brightness variations [57], [58].

In short, most correlation-like approaches are statistical similarity metrics that do not facilitate structural information or high computational complexity. Owing to their easy hardware implementations, they remain in frequent use for registration evaluation [59]. Fourier techniques have some advantages in terms of computational sufficiency, and they are robust against frequency-dependent noise. However, they have limitations in the case of image pairs with significantly different spectral content. Although MI methods offer outstanding performance compared with the two aforementioned algorithms, they do not always provide a global maximum of the entire search space for the transformation, as images containing insufficient information

or the overlap between two scenes inevitably reduces their robustness [16], [25]. Overall, intensity-based approaches directly use the pixel value of an image, without error accumulation, offering high-precision registration. However, these algorithms have limitations in terms of large rotations, translations, scale differences, and so on and are quite time-consuming.

OPTICAL FLOW ESTIMATION

Similar to the area-based approaches, optical flow estimation calculates object motions with direct and indirect consistency constraints based on pixel intensity. This technique is popular in computer vision for motion estimation. Owing to the similarity between the displacements of corresponding pixels under the same coordinate system and the optical flow of an object, some studies have utilized optical flow estimation to register remote sensing images [60], [61]. Unlike area-based approaches, optical flow estimation calculates pixel displacement based on intensity and gradient consistency constraints for coordinate recalculation. After resampling, the intensity value is assigned to the new noninteger position, and the aligned image is acquired [62], as summarized in Figure 7.

Optical flow is a 2D displacement field that describes the apparent motion of brightness patterns between two successive images [63], and its concept was proposed by Gibson [64]. Horn and Schunck (HS) [63] and Lucas and Kanada (LK) [65] proposed a differential approach for optical flow calculation in 1981. Since then, many extensions and modifications have been proposed for video image processing [66]–[68]. Given that the process is at the initial stage of development in the remote sensing field and that many studies have focused on differential techniques, the following
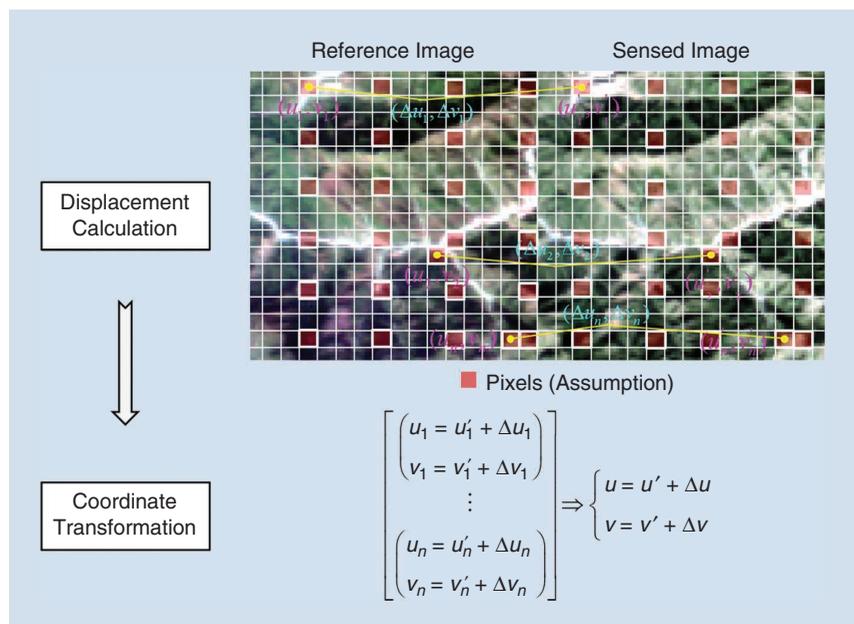

FIGURE 7. Optical flow estimation for remote sensing image registration. [(u_i, v_j) indicates the pixel coordinates in the reference image, and (u'_i, v'_j) indicates the coordinates of the corresponding pixel in the sensed image. The coordinate difference, which we called the *displacement*, is depicted as $(\Delta u_i, \Delta v_j)$.]

aspects are generally emphasized in research on remote sensing image registration.

DENSE OPTICAL FLOW ESTIMATION

The differential method for dense optical flow calculation proposed by HS is generally called the *typical global approach* [63]. Dense optical flow calculates each pixel's motion in a scene, as in Figure 8. The regular grid represents image pixels, and the displacement is displayed at equal intervals, where only the displacement directions and magnitudes of the green pixels are marked, for brevity. The HS optical flow integrates the brightness constancy assumption and the global smoothness constraint to separately estimate the pixel motion in the x and y directions. The intensity value constancy assumption is markedly susceptible to slight brightness changes [69], which are inevitable for remote sensing images. Applying the spatial gradient constancy assumption to the HS equation [as in (3)] is popular in research on multitemporal remote sensing image registration [62], [69]:

$$E(u,v) = \int_{\Omega} \psi \left(|I(x+w) - I(x)|^2 + \gamma |\nabla I(x+w) - \nabla I(x)|^2 \right) dx$$
$$+ \alpha \int_{\Omega} \psi \left(|\nabla_3 u|^2 + |\nabla_3 v|^2 \right) dx, \tag{3}$$

where $w = (u,v,1)^T$ is the pixel displacement to be solved, $X = (x,y,t)^T$ is a pixel coordinate, $\psi(s^2) = \sqrt{s^2 + \varepsilon^2}$ is an increasing concave function, and ε is a fixed value. Here, α and γ are the weights for the gradient and smoothness terms, respectively, and $\nabla_3 = (\partial_x, \partial_y, \partial_t)^T$ indicates a spatiotemporal smoothness assumption and is often replaced by the spatial gradient when used for remote sensing image registration.

Owing to the advantages of the per-pixel computation of optical flow estimation, very local deformation due to terrain elevations can be eliminated. Occlusion remains a challenge for accurate dense optical flow calculation [66], which is similar to land use (LU) and land cover (LC) changes in remote
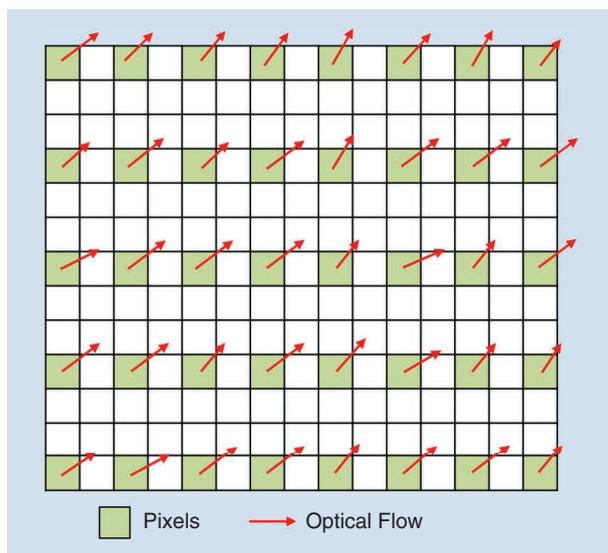
FIGURE 8. Dense optical flow.

sensing images [62]. Under this circumstance, an object in the reference (sensed) image cannot be sought in the sensed (reference) image. For example, in the yellow, rounded rectangles in Figure 9(a) and (b), a road disappears in the sensed image. This leads to further abnormal pixel displacement, in Figure 9(c), where the magnitudes and directions of the displacements are inconsistent with the neighborhood. The successive abnormal displacements further change the content of the aligned image, although it is highly geometrically aligned with the reference image in Figure 9(d). This change opposes the principle of image registration in that it does not alter the image content but spatially aligns the sensed and reference images. After the abnormal displacement correction, the recalculated displacement is similar to that of the surrounding region, as in Figure 9(e). Furthermore, the aligned image is similar to the corresponding region in the sensed image in Figure 9(b), and the two are spatially aligned with the reference image, as in Figure 9(f).

For large-scale movements, which are another concern when applying optical flow for remote sensing image registration, an improved approach was proposed in [70]. The pixel displacement calculated by the extended phase correlation technique is determined as the initial motion estimator for the global optical flow to achieve general remote sensing image registration, especially for large-scale movement deformation [70]. However, given that dense optical flow estimation calculates the displacement for each pixel, it is unavailable for the real-time registration of large images, although it provides a high-precision result.

SPARSE OPTICAL FLOW ESTIMATION

Sparse optical flow estimation is more popular for remote sensing image registration than its dense counterpart is. The sparse optical flow represented by the local difference may be supported in a specified local region, such as the position of the feature points extracted by popular extractors, including the scale-invariant feature transform (SIFT), as shown in Figure 10. This approach assumes that pixel motions are identical within a local neighborhood and estimates the optical flow by performing least-squares regression with a set of similar equations [66]. The LK gradient-based approach [65], as the origin, is widely used to estimate the motion of video images, on an equal footing with the HS model. The GeFOLKI algorithm was developed from LK and implemented on a graphics processing unit to achieve real-time and robust optical flow estimation [60], [71]. Furthermore, the GeFOLKI algorithm is adopted for the coregistration of heterogeneous data, such as synthetic aperture radar (SAR) lidar images and SAR optical images [61]. Subsequently, given the different imaging mechanisms of SAR and high-resolution optical images, which benefit from the high registration precision of optical flow estimation, two dense feature descriptors replace raw intensities when aligning images by an optical-to-SAR flow; this combines the global and local optical flow estimation approaches [72]. Sparse optical flow based on specified and distinct pixels is computationally time saving, whereas
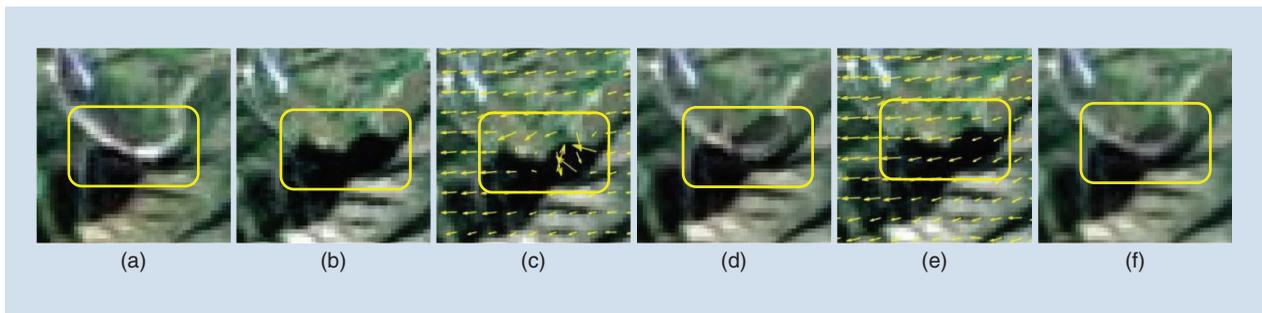
FIGURE 9. Abnormal displacement detection and correction. (a) The reference image. (b) The sensed image. (c) The displacement field estimated by (3). (d) The aligned image overlapping (a). (e) The corrected displacement field. (f) The aligned image formed by overlapping the corrected optical flow with (a) The highlighted road in (a) disappears in (b), leading to similar occlusion.

its accuracy for remote sensing image registration is relatively low compared with the dense optical flow approach. In addition, it is not vulnerable to LU–LC changes because it does not have similar features for sparse optical flow estimation in the changed region.

In summary, optical flow estimation has been developed in computer vision for motion estimation in superresolution reconstruction for several decades, whereas it is in the initial stage of use in remote sensing image registration. Optical flow estimation is a superior pixel displacement calculation approach that is particularly interesting in the case of very local deformation due to, for example, terrain elevation, which has considerable influence on high-resolution image registration [61]. The efficiency of optical flow estimation should be considered when applying it to remote sensing because a wide field of view (WFV) is a characteristic of remote sensing image. Therefore, due to social development and seasonal changes, LU–LC changes are frequent phenomena for multitemporal remote sensing images. The dense optical flow approach is sensitive to such changes, leading to abnormal displacement and the alteration of the content of an aligned image. Therefore, efficient and accurate correction should be integrated into the initial optical flow estimation when used for registration.

FEATURE-BASED REGISTRATION

The feature-based approach directly exploits the abstract features of an image, rather than the pixel intensity, for registration. *Feature* refers to a distinct geometrical or advanced characteristic extracted by a specified approach. Geometrical features are distinct points, line segments, and closed boundary regions in a remote sensing image that can be detected or extracted by extant or novel approaches. Advanced features are abstract descriptions of local regions, which are extracted by a neural network (NN) (especially in the deep learning approach) to represent the original image. Geometric features are understood as being conventional for feature-based registration, and the use of advanced features is defined as novel feature-based registration.

CONVENTIONAL FEATURE-BASED METHOD

In general, salient and distinctive features, such as points, line segments, and closed boundary regions, are manually

and automatically detected to represent the original remote sensing image. The feature correspondence is then established between the reference and sensed images by a similarity comparison of the feature descriptors. The geometric relationship is calculated, guiding a sensed image that is spatially aligned with the reference. Ultimately, coordinates in the sensed image are transformed. The transformed coordinates are usually noninteger, and they are calculated by interpolation to acquire their intensity values, as demonstrated in Figure 11. In the following, we summarize geometrical feature extraction and matching because research into this subject has been at the core of the traditional feature-based approach.

FEATURE EXTRACTION

The feature extraction mentioned here is a representation of feature detection and extraction. Detection aims to locate distinctive features in an image and determine their positions. In the feature-extraction stage, the recognizable descriptor is uniquely constructed, identifying the detected feature. Formerly, features were manually selected. This approach is still in use today, as in the "image-to-image registration" module in Environment for Visualizing Images (ENVI) software. Experts require a considerable amount of time for this approach,
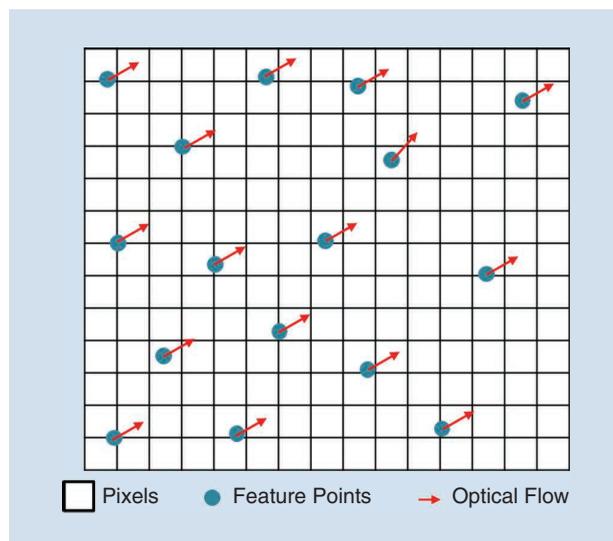

FIGURE 10. Sparse optical flow.

especially for large remote sensing images. At present, many methods have been proposed to automatically acquire representative features. Common geometrical features, including salient points (line intersections, corners, points on curves with high curvature, and road crossings) [73], [74], polylines (roads, contours, and edges) [41], [75], and polygons (closed boundary regions and lakes) [76], are selected by the specified approach. As shown in Figure 11, the yellow points, line segments, and regions are detected to abstractly describe the original image.

FEATURE POINTS

The local points at which the gray value varies dramatically in all directions are feature points, including corner points, inflection points, and T-intersection points. Many attempts have been made to extract them in computer vision, inspiring the development of feature point extraction in remote sensing. The first corner detection approach was proposed by Moravec in 1977 [77]. This algorithm has fast computation but is sensitive to noise and vulnerable to image rotation, leading to its rare use in the remote sensing field. The Harris corner detector was proposed in 1988 [78]. This algorithm is invariant under grayscale and rotational changes. It and improved Harris algorithms are applied to remote sensing image processing [38], [74], [79], [80], mainly with respect to multiscale corner detection.

Smith and Brady presented the smallest unvalued segment assimilating nucleus operator [81], which is insensitive to local noise and has high anti-interference ability [82]. However, it is not widely used in remote sensing image registration [83], whereas the SIFT algorithm is [45], [58], [74], [84]–[90]. The SIFT was developed by Lowe [92] and is invariant under rotation, scale, and translational changes [93]. It has been followed by many improved versions, such as principal component analysis SIFT [94], scale-restriction (SR) SIFT [36], [95], affine

SIFT [96], and uniform robust SIFT [97], [98]. Moreover, the speeded-up robust features (SURF) [99] algorithm was proposed, by Bay et al. to overcome the time-consuming nature of the SIFT for large-scale remote sensing images [100]–[102]. SURF applies an integral image to compute image derivations and quantifies the gradient orientations in a small number of histogram bins [103]. Additionally, the features from accelerated segment test (FAST) [104]; binary, robust, independent elementary features (BRIEF) [105]; oriented FAST and rotated BRIEF [106], [107]; Kaze [108]; and accelerated Kaze [109] algorithms are fast tools for descriptor construction but are less widely utilized in remote sensing.

In addition, a novel key point detector combining corners and blobs for remote sensing image registration is under development to increase the number of correctly matched features [110]. Recently, looking at intensity differences in multimodal remote sensing images, robust and novel feature descriptors have been adopted to depict detected feature points; these include the LSS descriptor, which accommodates effects such as nonlinear intensity differences [36]; the histogram of oriented PC, based on structural similarity measures [57]; and maximally stable PC, representing a novel affine and contrast-invariant descriptor [111]. All these coincidentally absorb PC information. PC is similar to the image gradient, presenting structural information with resistance to variations in illumination [112]. Therefore, the use of phase consistency information is a trend in the construction of robust feature descriptors for multimodal remote sensing image.

FEATURE LINES

A feature line is also known as a *line feature*; it is the generalization of feature points, such as general line segments [113], object contours [75], roads, coast lines [114], and rivers [115]. Given that feature lines have more attributes than feature points as control features [116], they have been gradually developed for use in image registration [117] as well as remote sensing image registration [116], [118], [119]. Standard edge detection, as with the Canny detector [120], [121] and detectors based on the Laplacian of Gaussian [122] are conventional feature line detection approaches [16]. Recently, some excellent detectors generating precise and robust line segments have been proposed [123], [124], and they are suitable for line detection in remote sensing images. Feature lines are comparatively less utilized in the remote sensing field than are feature points because matching them is an obstacle. They are often abstracted from corners, midpoints, and endpoints as final features [16], thereby losing their geometric value.
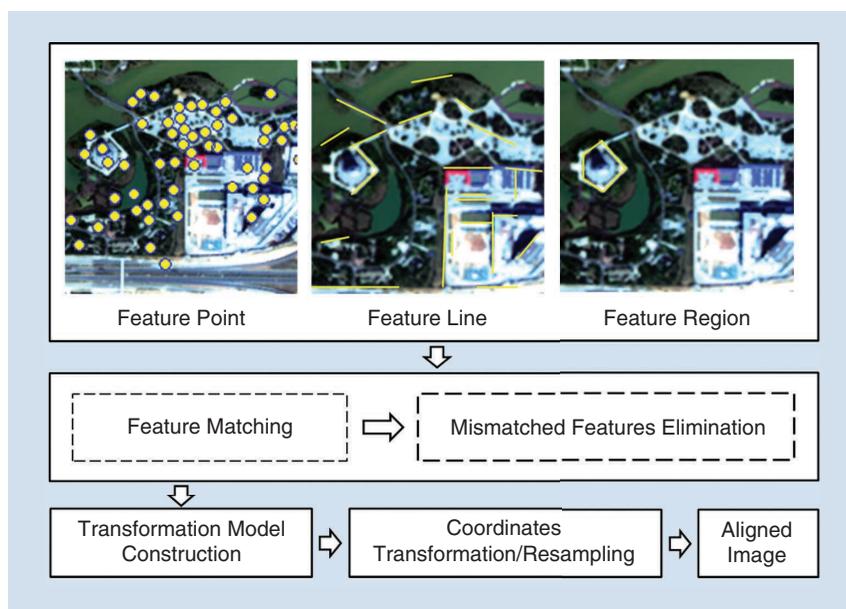

FIGURE 11. The geometrical feature-based registration algorithm.

FEATURE REGION

Feature region is a general term for all closed boundary regions of appropriate size, e.g., lakes [125], forests [126], buildings [113], urban areas [127], and so on. Before the robust feature point extraction approach was developed, the feature region was used to indirectly extract feature points. Regions with high contrast were extracted by filtering [128] and image segmentation [129] and described with moment-invariant descriptors [130], [131]. They are often abstracted by their centers of gravity [128], [132]–[135], which are invariant with respect to rotation, scaling, and skewing and are stable under random noise and gray-level variation [16]. Compared with feature points and lines, the extraction and description of feature regions were relatively early foci of research, and they have been used less for recent feature-based registration.

FEATURE MATCHING AND MISMATCHED FEATURE ELIMINATION

The correspondence relationship between reference and sensed images can be established based on detected feature points, lines, and regions, exploiting various descriptors of features [16], [136], [137]. Mismatched features are an inevitable byproduct of general feature matching, the elimination of which purifies correspondences for generating transformation models that are as accurate as possible. A pair of features with similar attributes is considered a selectable matching despite radiometric differences, noise, image distortion, and so forth. Under the circumstances, a robust matching measurement is essential. Feature matching approaches can be generally classified into two categories, namely, feature similarity and spatial relations.

FEATURE SIMILARITY

The constructed feature descriptors are used to establish the correspondence between extracted features in the reference and sensed images through feature similarity comparison. Feature similarity is conducted in the feature space by using the Euclidean distance ratio between the first and second nearest neighbors [92]. For efficiency, the k-dimensional tree and the best-bin-first algorithms are employed for feature similarity determination [93], [138]. The clustering technique [140], chamfer matching [141], and PC models are frequently used matching approaches, and they are invariant under intensity changes during matching [1].

SPATIAL RELATIONS

Aimed at tie point matching in poor textural regions, approaches based on spatial relations have been developed. Representative of these, graph-based feature points matching considers feature points as graph nodes. Feature matching is then transformed into a node-correspondence problem and solved by graph matching [125], [142]. Graph matching is applied to image feature correspondences, although it is not affine invariant [143]. By finding a consensus nearest-neighbor graph from candidate matches, a graph-transformation matching approach is developed [144]. Targeting the problem

in [143], a similar graph matching for tie point matching in poor textural images is proposed [101]. Furthermore, Xiong and Zhang introduced a novel interest point matching for high-resolution satellite images [145]. For this, the relative position and angle are used to reduce ambiguity and to avoid false matching, as the approach is suitable for image shifting and rotation. Affine and large-scale transformations are not considered [144].

MISMATCHED FEATURE ELIMINATION

Although the extracted features in a reference image have been matched with the corresponding ones in the sensed image via the aforementioned approach, some mismatched feature points are inevitable, further affecting the transformation model estimation [32], [76]. Therefore, eliminating mismatched features with a specified approach is necessary [146], [147]. Generally, based on the initial matching result, random sample consensus (RANSAC) is used to remove a mismatched point. This method randomly selects a sample from the consensus set in each iteration and finds the largest consensus set to calculate the final model parameters [33], [148]. RANSAC performs well and robustly when there are no more than 50% outliers [144], [149], [150]. Combining the local structure with global information, a restricted spatial order constraints algorithm is developed to find exact matched feature points in reference and sensed images [144].

Based on the affine-invariance property of the triangle-area representation (TAR), a robust sample consensus judging algorithm is proposed to efficiently identify bad samples and ensure accuracy with a light computational load [151]. For images with simple patterns, large affine transformations, and low overlapping areas, a mismatch- removal principle based on the TAR value of the k-nearest neighbors is proposed and referred to as *k-nearest neighbors–TAR* [149]. Furthermore, an improved RANSAC approach called *fast sample consensus* is developed to obtain correct matching in a few iterations [150], [152]. Thus, most of the reserved feature points in the reference image accurately correspond to the specified feature points in the sensed image, as the feature points connected by the yellow lines in Figure 12 will add precision to the transformation model estimation in the following step. The geometrical feature-based approach abstracts an original remote sensing image with distinct features instead of its intensity information, which is efficient and can easily process large rotations, translations, and scale differences between reference and sensed images. However, position errors in the automatically extracted features are inevitable, and a few mismatched features cannot be eliminated. This leads to a relatively low registration precision compared with the intensity-based approach.

NOVEL FEATURE-BASED REGISTRATION BY DEEP LEARNING

Deep learning provides a new concept for remote sensing image registration. It essentially refers to image registration based on advanced feature extraction [153]. Deep learning

originated in computer vision and has a long history [154]. In recent years, it has gradually entered use in remote sensing image applications, such as image fusion [155], [156], LC classification [157], [158], and segmentation [159]. The framework is data driven and can generate image features by learning from many training data sets with a specified principle [158]. Therefore, it is suitable for remote sensing image registration.

Some studies have focused on feature matching for this purpose [158], [160]. Most utilize a Siamese network consisting of two parts to train a deep NN (DNN) [161]–[164]. One part extracts features from image patch pairs by training a Siamese, pseudo-Siamese, or improved Siamese network [165]; the other part measures the similarity between these features for image matching. In [164], the DNN inspired the construction of a deep learning framework for remote sensing image registration. In addition, generative adversarial networks (GANs) are applied to image matching and registration [166], [167]. These approaches first translate an image into another one by training the GANs, enabling two images to have similar intensities and feature information [166], [168]. Feature extraction and matching are subsequently performed between two artificially generated images, effectively improving the performance of image matching. For the deficiencies of specified-scale NNs, multitask learning is introduced to improve the registration precision [169]. Wang et al. break through the limitations of the traditional deep learning approach, which extracts image features in one network and matches them with the other NN. They design an end-to-end network using forward propagation and backward feedback to learn the mapping functions of the patches and their matching labels for remote sensing image registration [164]. Recently, Li et al. paired image blocks from sensed and reference images and directly learned the displacement parameters of four corners of the sensed block relative to the reference image on a deep learning regression network, which differs from the traditional deep learning method [170].

Deep learning has advantages over the traditional registration approach. It is completely data driven and has strong flexibility, enabling it to theoretically fit any complex mapping function, whereas the traditional registration method can deal only with fixed pattern registration. Moreover, deep learning extracts abstract and high-level semantic information. Compared with low-level gray and gradient data, deep semantic information is more consistent with the way humans understand images. Therefore, deep learning methods can extract robust features. However, deep learning has challenges. It highly depends on image samples; when there is a lack of data or the data quality is poor, deep learning methods have difficulty ensuring the effectiveness of the registration results. Although remote sensing images are now easy to acquire, the lack of manual annotation and standard data is still very serious. Deep learning, in essence, learns the statistical characteristics of a large number of similar images, but its input–output process is a complex, nonlinear mapping without clear physical significance. Additionally, deep learning requires high computing power and has major hardware requirements, limiting its applicability.

In short, remote sensing image registration based on deep learning is still in its infancy, and its registration framework is not mature. However, many studies have demonstrated that deep learning methods can achieve or even surpass the optimal level of traditional registration approaches in terms of accuracy and efficiency. We predict that deep learning-based methods will become important solutions to the problem of real-time and high-precision remote sensing image registration.

REGISTRATION BASED ON THE COMBINATION METHOD

As mentioned, feature- and intensity-based approaches have their own advantages. Different feature extractors also have various precisions. To integrate these strengths as fully as possible, combination techniques have been developed. Typically, popular combinations consist of two aspects, namely, feature- and area-based approaches; however, some integrate two geometric feature-based approaches, such as the SIFT and Harris detectors.

COMBINATIONS OF FEATURE- AND AREA-BASED ALGORITHMS

Feature-based approaches are typically suitable for images with more significant structural data than intensity information. However, they are restricted by the distribution and accuracy of the features. On the other hand, area-based approaches are appropriate for images with more distinctive intensity information; however, they require the intensity information of the reference and sensed images to be correlated. Thus, the two methods have complementary pros and cons. To further improve registration accuracy and robustness, some studies focus on a combination of geometric feature- and area-based techniques [171]. Huang et al. [172] proposed a hybrid approach to aligning images by intensities within a scale-invariant feature region. Elsewhere, a wavelet-based feature extraction technique and an area-based method with NCC were combined to reduce the local distortion caused by terrain relief [173]. In a wavelet-based hierarchical pyramid framework, Mekky et al. [174] proposed a hybrid approach using MI and the SIFT; employing the rough registration parameters of the area-based approach for MI, the number of false alarms obtained by the SIFT was reduced. In addition,
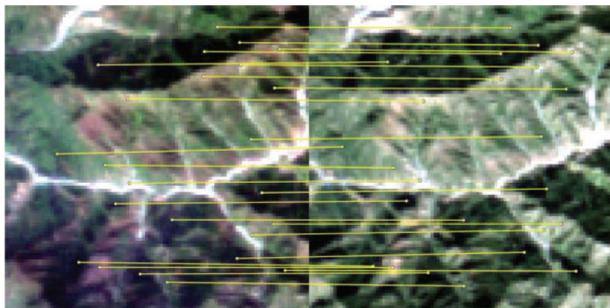

FIGURE 12. Feature matching examples.

Gong et al. employed the robustness of the SIFT and the accuracy of MI, proposing a novel coarse-to-fine registration framework aimed at registering optical and SAR remote sensing images [90]. For multisensor SAR image registration, Suri et al. proposed a multistage registration strategy. The rough parameters of the transformation model are estimated by MI, and this model is introduced during the SIFT matching phase to increase the number of tie points [175]. Under the SIFT and MI combinations, Heo et al. introduced a stereo matching method that produces accurate depth maps [176]. All these approaches can be considered coarse-to-fine-processing chains. The basic idea is to improve the result of the feature-based approach by adopting an optimization process from an area-based technique [90], [171].

The combined methods integrate the robustness of the feature-based algorithm with the accuracy of the area-based approach. They are relatively few compared with individual methods, but their combination will be the focus in the near future, from our point of view. To deal with the possible accumulation of errors, bundle block adjustment is usually needed [178], [179] to register sequential images. Moreover, the integration of different geometric feature-based approaches is being developed, as well, for ever-increasing transformation model estimation accuracy, generating precise registration results to the greatest extent possible.

INTEGRATION OF TWO GEOMETRIC FEATURE-BASED APPROACHES
In addition to combinations of feature- and area-based techniques, the integration of two geometric feature-based approaches is a developing trend for high-precision registration. In particular, the feature points extracted by different methods are used to register images in two stages. Yu et al. proposed to extract feature points using the SIFT for the preregistration of *Satellite Pour l'Observation de la Terre-5*/Thematic Mapper/*Quickbird* images from different sensors [74]. In the fine registration stage, the Harris algorithm for corner point detection is enforced to detect the distinct corner, and the extracted point is matched by the NCC algorithm. Similarly, Lee used SURF to extract the feature point of a low-resolution image after Harr wavelet transformation, which is defined as rough registration [180]. Fine registration is the same as the approach proposed by Yu et al. Recently, Ye et al. utilized SR–SIFT to extract the feature point in the preregistration stage for distinct translation, rotation, and scale difference elimination.

To further optimize registration, the Harris algorithm was employed to detect feature points in the reference and prealigned images and describe them by LSS for matching [36]. To register large, high-resolution remote sensing images, a coarse-to-fine strategy combining the Harris–Laplace detector with the SIFT descriptor has been proposed. After rough registration, a large image is divided into small, processable blocks for fine alignment [181]. Additionally, in a new two-step registration, the approximate spatial relationship is calculated with the deep features using a convolutional NN in the first step. Then, the previous result is adjusted based on the extracted local features [182]. Another technique combines feature point and feature line methods for the registration of images covering low-texture scenes in the computer vision field [183]. Since low- and repeated-texture regions are common in remote sensing images, feature lines can be employed to supplement the number of feature points. Therefore, beside the combination of two geometric feature-based methods, the integration of different geometrical features has great potential for the high-precision alignment of remote sensing images [22].

Since combination schemes integrate the advantages of two or more registration approaches, they offer remarkable precision. Moreover, in general, preregistration provides a rough result that approximates the final alignment. With fine-tuning in the optimized registration stage, a high-precision registration result is finally acquired. This algorithm is suitable for remote sensing image registration with large spatial position differences. It is as time-consuming as two or more alignment strategies.

SOFTWARE-BASED REGISTRATION
Most reviews emphasize the ever-increasing number of image registration approaches that are improved on the basis of existing methods for registering larger and more complicated images [16], [184]. Few studies have evaluated the performance of software-embedded image registration modules and the packages/tools for image geometric registration [185]. Thus, in this section, we present some examples.

The Earth Resources Data Analysis System (ERDAS), ENVI, PCI Geomatica, ER Mapper, and Arc Geographic Information System (GIS) are well-known software packages for remote sensing image processing that include registration modules. ER Mapper was acquired by ERDAS a few years ago. They integrate conventional manual and automatic registration programs. Concretely, ENVI could register two remote sensing images or align one image with a map covering the same scene. A user can extract tie points by observing similar objects lying on two images, such as corners of buildings, road intersections, inflection points of rivers, and so on. With a uniform point distribution, the parameters of a specified transformation model can be estimated. There are some general geometric mapping functions, including affine, polynomial, and triangulation transformation models. Geometric mapping is generally conducted by an expert and is time-consuming and tedious. It is difficult to avoid subjective factors while extracting tie points, especially when registering WFV images that require more time than general image registration. To liberate the productive forces and improve the registration efficiency, the automatic alignment technique is also put into ENVI.

We should point out the reference and sensed images, respectively. After setting the area-based matching parameters, the tie point for transformation model construction is automatically extracted; soon, the aligned image is obtained. Neither the manually extracted tie point nor the automatically acquired point in ENVI is sufficiently accurate. For example, the coordinates of the extracted feature point are $(157.05, 171)$,

which may suggest the neighborhood of the real corner. Under this circumstance, the calculated geometric spatial relationship is not as precise as it could be. The obtained registration result is usually worse than expected, especially for high-resolution remote sensing images with inconsistent local deformation.

ERDAS was developed by the ERDAS Corporation, in the United States. Compared with ENVI, it can produce tie points with higher location accuracies [for instance, the coordinate of the extracted feature point is $(385.776, 75.161)$, which has more decimal places] to generate precise mapping functions between reference and sensed images that approximate real geometric relations. Additionally, there are abundant transformation models, such as linear rubber sheeting, nonlinear rubber sheeting, and the direct linear transform. Elevation data are introduced into the registration to generate the high-precision alignment of mountainous remote sensing images, even using the digital terrain model (DTM). Furthermore, the region and interval of the selected tie point can be set manually in the "AutoSync" module. To acquire a high-precision registration result, the elevation data (DEM or DTM) should be input at the same time as the image to be registered. If higher-spatial-resolution elevation data were included in ERDAS, the corresponding information would be automatically extracted when an image's geographic information was identified to register the input image.

Image registration can also be conducted in ArcGIS, although most researchers would probably utilize the software to solve problems with the GIS, such as spatial analysis. PCI Geomatica prefers to produce orthophoto and fusion images, rather than registering remote sensing images. However, both ArcGIS and PCI Geomatica contain an image registration module. The steps for alignment processing are similar to those for the aforementioned software, including manual registration and automatic operation. Some different transformation models, such as spline, similarity polynomial, and projective transformations, are used to achieve the high-precision registration of complicated remote sensing images. However, sometimes the result is unsatisfactory for further applications, as the tie points are not uniformly distributed and their number is small. Pixel Information Expert is a new generation of remote sensing image processing software that was developed by Beijing Aerospace Hongtu Information Technology. It can handle the dislocation of multisource, heterogeneous remote sensing images since it integrates a novel algorithm with a focus on multimodal remote sensing image registration. It can be tested free for 30 days. In addition, copyrighted geometric registration software, such as the Hyperspectral Image Processing and Analysis System, GeoImager, Titan Image, and so forth, were generated by the Institute of Remote Sensing, Chinese Academy of Sciences.

Because high-resolution image registration is an important task in remote sensing image processing, much emphasis has been placed on it. To extract dense tie points representing local geometric relationships, SURF and an adaptive binning SIFT descriptor have been combined [186]. With the guidance of the local transformation model, an accurate registration result is obtained. The MATLAB code for the algorithm is provided, with experimental data, at https://www.researchgate .net/publication/320354469_HRImReg. The code is encrypted, and the parameters cannot be adjusted. It can be used only for comparative experiments to evaluate a proposed approach. When doing simulation experiments to assess a feature point detector or to evaluate a mismatched elimination approach with real data, the progressive sparse spatial consensus algorithm can be employed [187]. The code, with experimental data, is publicly available at https://github.com/ jiayi-ma?tab=repositories. It has been tested on photographs from the computer vision field. To apply it to remote sensing images, some improvements are needed. Beyond these, there are many commercial and open-source software packages/ tools for geometric registration. There are also different points of view, which should be discussed in depth in the future as more resources become available. However, an evaluation of registration approaches should be conducted, as well, whenever an aligned image is generated from software or a proposed method.

EVALUATION OF IMAGE REGISTRATION ACCURACY

For the spatial alignment of remote sensing images, it is highly desirable to provide users with an estimate of how accurate the registration actually is. Accuracy evaluation is a nontrivial problem that is present in all literature on remote sensing image registration. We have identified three aspects to measuring the registration accuracy on the basis of different considerations, including tie point identification, the transformation model performance, and the alignment error. In this section, we review basic approaches for alignment assessment.

ACCURACY OF TIE POINTS

The quality and quantity of tie points are important to guarantee high-precision image registration. The number of redundant tie points, in addition to the elementary computation of the specified transformation model, is essential information since we generally use as many tie points as possible to calculate the parameters of the mapping function for alignment. Furthermore, we must allow for a residual $(\Delta x_i, \Delta y_i)$ for the ith extracted feature point compared with the origin of the image [188]. If there are N tie points, the root-mean-square error (RMSE) can be estimated as follows:

$$RMS_{tp} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left((\Delta x_i)^2 + (\Delta y_i)^2\right)}. \qquad (4)$$

To enable general comparison, the RMSE should be computed across the normalized (to the pixel size) residuals. Additionally, the bad point proportion should be calculated to evaluate the extracted feature point. This is the number of residuals that lie above a certain threshold multiplied by the ellipse formed by the pixel size. Besides the mentioned criteria, the distribution of tie points is attracting increased attention. To design a uniform distribution of tie points, some papers have proposed to extract feature points within a specified

subregion [30]. A detection approach is employed to extract the specified number of feature points. Tie points affect the registration accuracy but are not the sole influencer.

TRANSFORMATION MODEL PERFORMANCE

The transformation model abstractly represents the geometric mapping function from a sensed image to a reference image. The actual between-image geometric distortion is difficult to obtain without prior information, and the estimated transformation approximates the real geometric relationship between images. One part of the N pairs of tie points is taken for mapping function estimation through the least-squares method, assuming N matched feature points. The left part in the sensed image is employed as the test point to be transformed into the reference image system [188]. The distance between the transformed coordinate and the corresponding point in the reference image is calculated as the residual, the mean of which is a representation of the estimated transformation model:

$$RMS_{N-te} = \sqrt{\frac{1}{N-T}\sum_{j=1}^{N-T}((x-Hx')^2 + (\gamma-H\gamma')^2)}, \quad (5)$$

where H denotes the estimated transformation model by T pairs of tie points, (x, γ) and (x', γ'), which represent the corresponding points in the reference and sensed images, respectively. Furthermore, a χ^2 goodness-of-fit test may be applied [188] to analyze whether the residuals are equally distributed across all quadrants. However, "overfitting" may yield zero error for a mapping model with sufficient degrees of freedom; this is a well-known phenomenon in numerical analysis. Under this circumstance, the registration results may not be optimal.

ALIGNMENT ERROR

The oldest method for estimating registration accuracy is visual assessment by a domain expert, which is still in use and remains the most effective technique, although it cannot be quantified [16], [188]. At present, this is performed using professional software, such as ENVI and ArcGIS, with shutter tools. Similarity metrics in area-based registration, such as MI, NMI, CC, and so on, are frequently employed to evaluate alignment accuracy [59]. The indicators are easily influenced changes in the information with development and differences in radiation. To quantitatively present the alignment error, the RMSE is calculated using feature points manually extracted by a specialist employing (4) [85]. Since image registration aims to achieve the relative spatial alignment of two different images, there is no gold standard reference image with which to evaluate the registration accuracy. When evaluating outcomes according to at least three criteria, the most indicative results point to the best registration, as different assessments have their own advantages and disadvantages.

FUTURE TRENDS

There has been a large number of independent studies on remote sensing image registration, and much effort has been put into constructing robust feature descriptors and eliminating mismatched features. With the development of sensor technology and application requirements, some novel opportunities and challenges must be addressed for remote sensing image registration. To us, it seems likely that the future of this field will include accelerated, combined, heterogenous, cross-scale, and smart remote sensing image registration techniques, which are introduced in detail in the following.

ACCELERATED REMOTE SENSING IMAGE REGISTRATION

With the ongoing development of sensor technology, the spatial resolution of remote sensing images increases, resulting in a growing number of features with distinctive details. The huge number of features lengthens the distance to the real-time registration of remote sensing images, causing inefficiency when aligning large-scale images. Thus, constructing descriptors and matching the detected features is time-consuming for general images, especially WFV ones. As proposed in [52], to achieve real-time registration to the greatest extent possible, remote sensing image registration can be operated on a cloud platform based on finite-state chaotic compressed sensing theory. Similarly, cloud computing [91] and some hardware systems may also be effective for accelerating image registration. At present, parallel computing [139] is the easiest path to implementation. Here, an image is divided into several subregions, and the image features in each one are simultaneously extracted, based on the same principles, on different parallel processors, as is the transformation model construction. The parallel commands are easy to implement on MATLAB and other platforms.

COMBINED APPROACHES FOR IMAGE REGISTRATION

With the development of imaging sensors, the resolution of remote sensing images has increased, and local deformation has become obvious. For example, the geometric distortion caused by terrain relief and high-rise buildings leads to inaccurate registration [36], introducing difficulties for remote sensing image applications. The reference and sensed images cover the plain and mountainous regions simultaneously in Figure 13(c). Calculating the displacements of corresponding pixels for spatial registration, the enlarged displacements in the specified rectangular regions are shown in Figure 13(d) and (e). The magnitude and direction of the displacements in the plain region are similar, but they differ in the mountainous region. Here, multistage registration with a global mapping function cannot exactly describe the spatial relationship between the reference and sensed images, and neither can the local transformation model.

Given that displacements vary in different terrain regions, dividing images into a series of regions and registering with a specified approach may yield a high-precision alignment, indicating a combination of different techniques. Concretely,
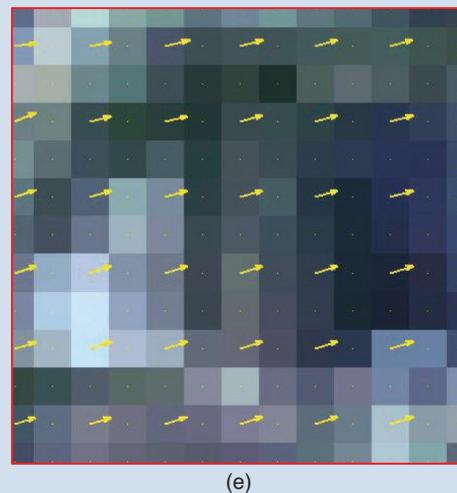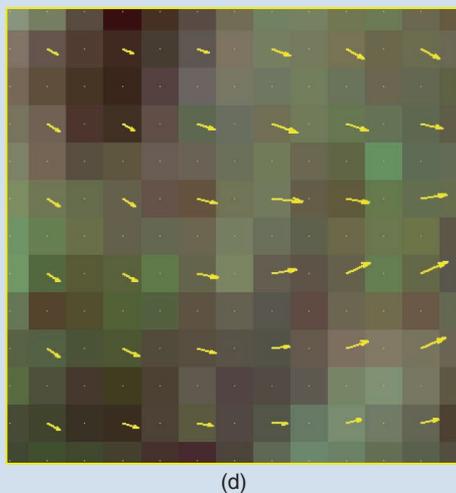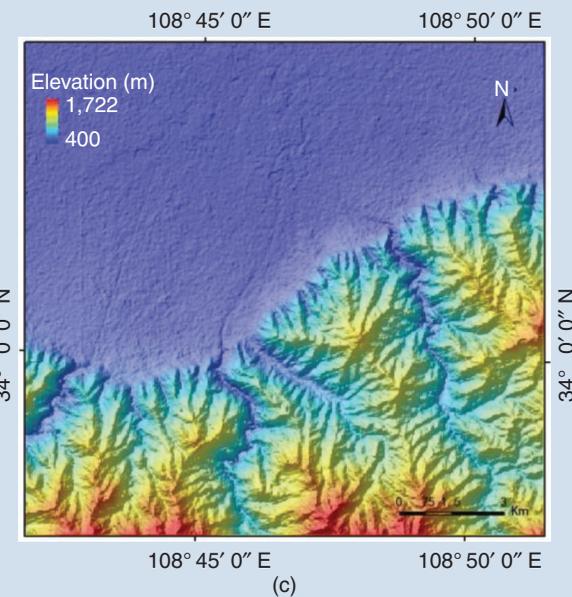
FIGURE 13. The spatial position of corresponding pixels in a remote sensing image of complex terrain. (a) The reference image. (b) The sensed image. (c) The topographic image. (d) The displacements in the mountainous region marked with a yellow rectangle in (a) and (b). (e) The displacements in the plain region marked with a red rectangle in (a) and (b).

this transformation model is calculated with distinct tie features in the plain region. With the transformation model, rather than directly obtaining the aligned plain region, the displacement guiding pixels to alignment is estimated. In mountainous regions, the dense optical flow estimation borrowed from computer vision is utilized to acquire the displacement of each corresponding pixel. Then, the displacement fields from different terrain regions are mosaicked (e.g., using the inverse distance weighted function for uniform transitions in image stitching) to obtain a seamless displacement field of the entire image [177]. This is a creative combination of different registration approaches in a coordinated way, differing from the combined approaches mentioned in the "Registration Based on the Combination Method" section with the serial mode. Therefore, regional registration accommodating complex geometric relationships that vary with terrain differences may become a significant trend in remote sensing image registration, giving full play to the registration advantages of different approaches in various terrain regions.

HETEROGENOUS AND CROSS-SCALE IMAGE REGISTRATION

Heterogenous and cross-scale images collected all at once and at different times provide complementary information to improve our understanding of an entire scene during Earth observation or even during disaster rescues. However, such data usually have dramatically different spatial resolutions, intensities, noise, geometries, and so on, owing to different imaging principles. Some studies have focused on spatial registration, including optical image and SAR registration, optical image and infrared image registration, and satellite image and map registration [36], [57]. These works emphasized the robust construction of descriptors to resist intensity and noise differences and other influential factors. Large-scale differences between cross-scale images (which are much greater than four times the resolution difference between the panchromatic and multispectral images) introduce difficulties for extracting geometrical features from low-resolution images that are similar to those from high-resolution images. Thus, generating the tie features of cross-scale images for transformation model construction, even during high-precision registration, is difficult. Additionally, high-efficiency heterogenous and cross-scale image registration remains an open problem that is worth researching in the near future. For a concrete example, the approximately real-time registration of optical and SAR images may offer an approach for analyzing disaster regions as quickly as possible for rescue purposes by means of registering and comparing images before and after an event. These applications are vital for rescue operations. Precise and efficient heterogeneous and cross-scale image registration is a mandatory prerequisite for high-precision, real-time applications.

SMART REMOTE SENSING IMAGE REGISTRATION

To register multiple remote sensing images, one simple and conventional idea is to align them frame by frame, namely, by converting multiple image registration into pair-to-pair alignment. This process, learning from the simultaneous mosaicking of multiframe images, specifies a reference image connected to others and stitches other images to the reference one. Therefore, when images to be registered are read into the program, the coordinates of the four corners in each image are extracted. The reference image is determined by comparing these coordinates. As presented in Figure 14, images A, B, C, and D are simultaneously aligned with the reference image (marked in green) according to a general registration strategy, as there is overlap between two images. Unlike frame-to-frame approaches, this technique needs to specify only the reference image, and the intermediate results do not output and input many times, which saves memory and improves computational efficiency. From our point of view, this is smart registration, which is particularly useful for WFV-image generation. However, when images overlap, a more intelligent approach needs to be developed.

Moreover, images to be registered may have small overlapping areas. This overlap presents a challenge for high-accuracy alignment because a small number of geometric and intensity features is available for constructing the transformation model. This problem should be intelligently solved to register images with a low ratio of overlapping regions. Typically, these images are used to produce WFV images by means of stitching. Further solutions should be provided in the future.

Therefore, the large-scale, complex distortion of high-resolution, heterogenous, and cross-scale remote sensing images must be a focus of future research. In this situation, the traditional single-registration approach may not meet requirements. For real time, high-precision registration, a combination of alignment approaches and high-performance computing is considered very promising.

CONCLUSIONS

In this article, we presented a comprehensive and quantitative summary of intensity-based, feature-based, and combined approaches to remote sensing image registration. Conventional methods and new applications of deep learning and optical
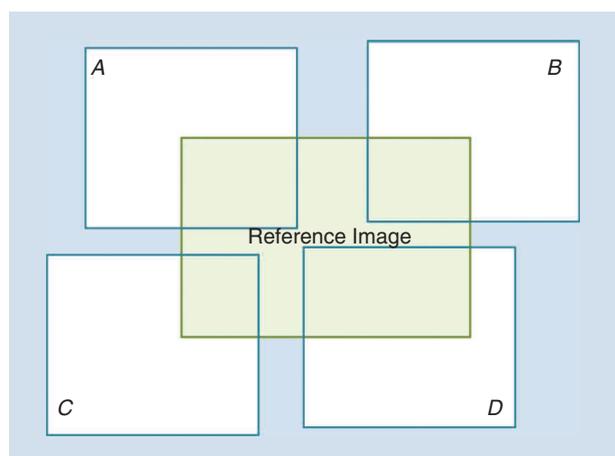
FIGURE 14. The spatial position of multiple images to be registered.

flow techniques were included. The performance of registration software packages and tools was analyzed. Additionally, novel registration evaluations were presented to support an effective assessment. The development of any approach aims to improve registration accuracy as much as possible because registration is an important step for preprocessing remote sensing images. Several such techniques have been developed, as recounted in this article.

However, as resolutions increase, the problem of inconsistent local distortion caused by high-rise buildings and topographic relief has become apparent; this cannot be exactly described by the transformation model. Moreover, WFV images are an emerging trend in satellite image production, enabling a whole ROI to be contained within one image. This poses a challenge for real-time registration and memory for registration processing. Therefore, we believe that future research on remote sensing image registration will use accelerated registration, combined approaches for remote sensing image registration, heterogeneous and cross-scale image registration, and smart registration. Challenges remain, and considerable additional research is required. We perform this research with the advantage of lower entrance barriers than the TDOM generation.

AUTHOR INFORMATION
Ruitao Feng (feng-rt@snnu.edu.cn) is with the School of Geography and Tourism, Shaanxi Normal University, Xi'an, 710062, China.

Huanfeng Shen (shenhf@whu.edu.cn) is with the School of Resource and Environment Science and the Collaborative Innovation Center for Geospatial Technology, Wuhan University, Wuhan, 430072, China. He is a Senior Member of IEEE.

Jianjun Bai (bjj@snnu.edu.cn) is with the School of Geography and Tourism, Shaanxi Normal University, Xi'an, 710062, China.

Xinghua Li (lixinghua5540@whu.edu.cn) is with the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, 430072, China. He is a Senior Member of IEEE.

REFERENCES
[1] A. Wong and D. A. Clausi, "ARRSI: Automatic registration of remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1483–1493, 2007. doi: 10.1109/TGRS.2007.892601.

[2] X. Li, N. Hui, H. Shen, Y. Fu, and L. Zhang, "A robust mosaicking procedure for high spatial resolution remote sensing images," *ISPRS J. Photogram. Remote Sens.*, vol. 109, pp. 108–125, Nov. 2015. doi: 10.1016/j.isprsjprs.2015.09.009.

[3] H. Shen, X. Meng, and L. Zhang, "An Integrated Framework for the Spatio-Temporal-Spectral Fusion of Remote Sensing Images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 12, pp. 7135–7148, 2016. doi: 10.1109/TGRS.2016.2596290.

[4] Y. Lu, P. Wu, X. Ma, and X. Li, "Detection and prediction of land use/land cover change using spatiotemporal data fusion and the Cellular Automata–Markov model," *Environ. Monitoring Assessment*, vol. 191, no. 2, p. 68, 2019. doi: 10.1007/s10661-019-7200-2.

[5] Z. Lv, T. Liu, C. Shi, J. A. Benediktsson, and H. Du, "Novel land cover change detection method based on k-means clustering and adaptive majority voting using bitemporal remote sensing images," *IEEE Access*, vol. 7, pp. 34,425–34,437, Jan. 2019. doi: 10.1109/ACCESS.2019.2892648.

[6] C. Yuan, F. Wang, S. Wang, and Y. Zhou, "Accuracy evaluation of flood monitoring based on multiscale remote sensing for different landscapes," *Geomatics, Natural Hazards Risk*, vol. 10, no. 1, pp. 1389–1411, 2019. doi: 10.1080/19475705.2019.1580224.

[7] L. Yang and G. Cervone, "Analysis of remote sensing imagery for disaster assessment using deep learning: A case study of flooding event," *Soft Comput.*, vol. 23, no. 24, pp. 13,393–13,408, 2019. doi: 10.1007/s00500-019-03878-8.

[8] K. Barbieux, "Pushbroom hyperspectral data orientation by combining feature-based and area-based co-registration techniques," *Remote Sens.*, vol. 10, no. 4, p. 645, 2018. doi: 10.3390/rs10040645.

[9] Y. Jiang, J. Wang, L. Zhang, G. Zhang, X. Li, and J. Wu, "Geometric processing and accuracy verification of zhuhai-1 hyperspectral satellites," *Remote Sens.*, vol. 11, no. 9, p. 996, 2019. doi: 10.3390/rs11090996.

[10] I. Aicardi, F. Nex, M. Gerke, and A. M. Lingua, "An image-based approach for the co-registration of multi-temporal UAV image datasets," *Remote Sens.*, vol. 8, no. 9, p. 779, 2016. doi: 10.3390/rs8090779.

[11] F. P. M. Oliveira and J. M. R. S. Tavares, "Medical image registration: A review," *Comput. Methods Biomech. Biomed. Eng.*, vol. 17, no. 2, pp. 73–93, 2014. doi: 10.1080/10255842.2012.670855.

[12] A. Sotiras, C. Davatzikos, and N. Paragios, "Deformable medical image registration: A survey," *IEEE Trans. Med. Imag.*, vol. 32, no. 7, pp. 1153–1190, 2013. doi: 10.1109/TMI.2013.2265603.

[13] M. A. Viergever, J. B. A. Maintz, S. Klein, K. Murphy, M. Staring, and J. P. W. Pluim, "A survey of medical image registration," *Med. Image Anal.*, vol. 33, pp. 140–144, Oct. 2016. doi: 10.1016/j.media.2016.06.030.

[14] G. Haskins, U. Kruger, and P. Yan, "Deep learning in medical image registration: A survey," *Mach. Vis. Appl.*, vol. 31, nos. 1–2, p. 8, 2020. doi: 10.1007/s00138-020-01060-x.

[15] L. G. Brown, "A survey of image registration techniques," *ACM Comput. Surv.*, vol. 24, no. 4, pp. 325–376, 1992. doi: 10.1145/146370.146374.

[16] B. Zitová and J. Flusser, "Image registration methods: A survey," *Image Vis. Comput.*, vol. 21, no. 11, pp. 977–1000, 2003. doi: 10.1016/S0262-8856(03)00137-9.

[17] M. Deshmukh and U. Bhosle, "A survey of image registration," *Int. J. Image Process.*, vol. 5, no. 3, p. 245, 2011.

[18] Z. Xiong and Y. Zhang, "A critical review of image registration methods," *Int. J. Image Data Fusion*, vol. 1, no. 2, pp. 137–158, 2010. doi: 10.1080/19479831003802790.

[19] M. V. Wyawahare, P. M. Patil, and H. K. Abhyankar, "Image registration techniques: An overview," *Int. J. Signal Process., Image Process. Pattern Recognit.*, vol. 2, no. 3, pp. 11–28, 2009.

[20] C. Dalmiya and V. Dharun, "A survey of registration techniques in remote sensing images," *Indian J. Sci. Technol.*, vol. 8, no. 26, pp. 1–7, 2015. doi: 10.17485/ijst/2015/v8i26/81048.

[21] R. M. Ezzeldeen, H. H. Ramadan, T. M. Nazmy, M. A. Yehia, and M. S. Abdel-Wahab, "Comparative study for image registration techniques of remote sensing images," *Egyptian J. Remote Sens. Space Sci.*, vol. 13, no. 1, pp. 31–36, 2010. doi: 10.1016/j.ejrs.2010.07.004.

[22] M. P. S. Tondewad and M. M. P. Dale, "Remote sensing image registration methodology: Review and discussion," *Proc. Comput. Sci.*, vol. 171, pp. 2390–2399, June 2020. doi: 10.1016/j.procs.2020.04.259.

[23] P. E. Anuta, "Spatial registration of multispectral and multitemporal digital imagery using fast Fourier transform techniques," *IEEE Trans. Geosci. Electron.*, vol. 8, no. 4, pp. 353–368, 1970. doi: 10.1109/TGE.1970.271435.

[24] X. Xu, X. Li, X. Liu, H. Shen, and Q. Shi, "Multimodal registration of remotely sensed images based on Jeffrey's divergence," *ISPRS J. Photogram. Remote Sens.*, vol. 122, pp. 97–115, Dec. 2016. doi: 10.1016/j.isprsjprs.2016.10.005.

[25] J. Ma, H. Zhou, J. Zhao, Y. Gao, J. Jiang, and J. Tian, "Robust feature matching for remote sensing image registration via locally linear transforming," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6469–6481, 2015. doi: 10.1109/TGRS.2015.2441954.

[26] N. Hanaizumi and S. Fujimur, "An automated method for registration of satellite remote sensing images," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 1993, pp. 1348–1350. doi: 10.1109/IGARSS.1993.322087.

[27] W. F. Webber, "Techniques for image registration," in *Proc. LARS Symp.*, West Lafayette, IN, 1973, pp. 1–7.

[28] D. I. Barnea and H. F. Silverman, "A class of algorithms for fast digital image registration," *IEEE Trans. Comput.*, vol. C-21, no. 2, pp. 179–186, 1972. doi: 10.1109/TC.1972.5008923.

[29] S. i. Kaneko, Y. Satoh, and S. Igarashi, "Using selective correlation coefficient for robust image registration," *Pattern Recognit.*, vol. 36, no. 5, pp. 1165–1173, 2003. doi: 10.1016/S0031-3203(02)00081-X.

[30] H. Gonçalves, J. A. Gonçalves, L. Corte-Real, and A. C. Teodoro, "CHAIR: Automatic image registration based on correlation and Hough transform," *Int. J. Remote Sens.*, vol. 33, no. 24, pp. 7936–7968, Dec. 20, 2012. doi: 10.1080/01431161.2012.701345.

[31] J. Inglada and A. Giros, "On the possibility of automatic multisensor image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 10, pp. 2104–2120, 2004. doi: 10.1109/TGRS.2004.835294.

[32] J. Ma, J. C. Chan, and F. Canters, "Fully automatic subpixel image registration of multiangle CHRIS/Proba data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 7, pp. 2829–2839, 2010. doi: 10.1109/TGRS.2010.2042813.

[33] Y. Wu, W. Ma, Q. Su, S. Liu, and Y. Ge, "Remote sensing image registration based on local structural information and global constraint," *J. Appl. Remote Sens.*, vol. 13, no. 1, p. 1, 2019. doi: 10.1117/1.JRS.13.016518.

[34] G. Wolberg and S. Zokai, "Image registration for perspective deformation recovery," in *Proc. SPIE, Automatic Target Recognit. X*, Orlando, FL, 2000, vol. 4050, pp. 259–270.

[35] D. P. Huttenlocher, G. A. Klanderman, and W. J. Rucklidge, "Comparing images using the Hausdorff distance," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 15, no. 9, pp. 850–863, 1993. doi: 10.1109/34.232073.

[36] Y. Ye and J. Shan, "A local descriptor based registration method for multispectral remote sensing images with non-linear intensity differences," *ISPRS J. Photogram. Remote Sens.*, vol. 90, pp. 83–95, 2014. doi: 10.1016/j.isprsjprs.2014.01.009.

[37] Y. Hel-Or, H. Hel-Or, and E. David, "Fast template matching in non-linear tone-mapped images," in *Proc. Int. Conf. Comput. Vision (ICCV)*, Barcelona, Spain, 2011, pp. 1355–1362. doi: 10.1109/ICCV.2011.6126389.

[38] Y. Bentoutou, N. Taleb, K. Kpalma, and J. Ronsin, "An automatic image registration for applications in remote sensing," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 9, pp. 2127–2137, 2005. doi: 10.1109/TGRS.2005.853187.

[39] K. Taejung and I. Yong-Jo, "Automatic satellite image registration by combination of matching and random sample consensus," *IEEE Trans. Geosci. Remote Sens.*, vol. 41, no. 5, pp. 1111–1117, 2003. doi: 10.1109/TGRS.2003.811994.

[40] J. P. Kern and M. S. Pattichis, "Robust multispectral image registration using mutual-information models," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 5, pp. 1494–1505, 2007. doi: 10.1109/TGRS.2007.892599.

[41] H. m. Chen, M. K. Arora, and P. K. Varshney, "Mutual information-based image registration for remote sensing data," *Int. J. Remote Sens.*, vol. 24, no. 18, pp. 3701–3706, 2003. doi: 10.1080/0143116031000117047.

[42] A. A. Cole-Rhodes, K. L. Johnson, J. LeMoigne, and I. Zavorin, "Multiresolution registration of remote sensing imagery by optimization of mutual information using a stochastic gradient," *IEEE Trans. Image Process.*, vol. 12, no. 12, pp. 1495–1511, 2003. doi: 10.1109/TIP.2003.819237.

[43] D. Brunner, G. Lemoine, and L. Bruzzone, "Earthquake Damage assessment of buildings using VHR optical and SAR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 5, pp. 2403–2420, 2010. doi: 10.1109/TGRS.2009.2038274.

[44] X. Wang, W. Yang, A. Wheaton, N. Cooley, and B. Moran, "Efficient registration of optical and IR images for automatic plant water stress assessment," *Comput. Electron. Agriculture*, vol. 74, no. 2, pp. 230–237, 2010. doi: 10.1016/j.compag.2010.08.004.

[45] S. Chen, X. Li, L. Zhao, and H. Yang, "Medium-low resolution multisource remote sensing image registration based on SIFT and robust regional mutual information," *Int. J. Remote Sens.*, vol. 39, no. 10, pp. 3215–3242, 2018. doi: 10.1080/01431161.2018.1437295.

[46] L. Y. Zhao, B. Y. Lü, X. R. Li, and S. H. Chen, "Multi-source remote sensing image registration based on scale-invariant

feature transform and optimization of regional information," *Acta Phys. Sin.*, vol. 64, no. 12, pp. 124204, 1-11), 2015.

[47] G. Hermosillo, C. Chefd'Hotel, and O. Faugeras, "Variational methods for multimodal image matching," *Int. J. Comput. Vis.*, vol. 50, no. 3, pp. 329–343, Dec. 1, 2002. doi: 10.1023/A:1020830525823.

[48] R. N. Bracewell and R. N. Bracewell, *The Fourier Transform and Its Applications*. New York: McGraw-Hill, 1986.

[49] H. Foroosh, J. B. Zerubia, and M. Berthod, "Extension of phase correlation to subpixel registration," *IEEE Trans. Image Process.*, vol. 11, no. 3, pp. 188–200, Mar. 2002. doi: 10.1109/83.988953.

[50] X. Wan, J. G. Liu, and H. Yan, "The illumination robustness of phase correlation for image alignment," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5746–5759, 2015. doi: 10.1109/TGRS.2015.2429740.

[51] X. Wan, J. Liu, H. Yan, and G. L. K. Morgan, "Illumination-invariant image matching for autonomous UAV localisation based on optical sensing," *ISPRS J. Photogram. Remote Sens.*, vol. 119, pp. 198–213, Sept. 2016. doi: 10.1016/j.isprsjprs.2016.05.016.

[52] Z. Liu, L. Wang, X. Wang, X. Shen, and L. Li, "Secure remote sensing image registration based on compressed sensing in cloud setting," *IEEE Access*, vol. 7, pp. 36,516–36,526, Mar. 2019. doi: 10.1109/ACCESS.2019.2903826.

[53] M. Xu and P. K. Varshney, "A subspace method for Fourier-based image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 3, pp. 491–494, 2009. doi: 10.1109/LGRS.2009.2018705.

[54] L. Lucchese, S. Leorin, and G. M. Cortelazzo, "Estimation of two-dimensional affine transformations through polar curve matching and its application to image mosaicking and remote-sensing data registration," *IEEE Trans. Image Process.*, vol. 15, no. 10, pp. 3008–3019, 2006. doi: 10.1109/TIP.2006.877519.

[55] P. Bao and D. Xu, "Complex wavelet-based image mosaics using edge-preserving visual perception modeling," *Comput. Graph.*, vol. 23, no. 3, pp. 309–321, 1999. doi: 10.1016/S0097-8493(99)00040-0.

[56] H. Gang and Z. Yun, "Combination of feature-based and area-based image registration technique for high resolution remote sensing image," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, 2007, pp. 377–380. doi: 10.1109/IGARSS.2007.4422809.

[57] Y. Ye, J. Shan, L. Bruzzone, and L. Shen, "Robust registration of multimodal remote sensing images based on structural similarity," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 5, pp. 2941–2958, 2017. doi: 10.1109/TGRS.2017.2656380.

[58] H. Yang, X. Li, L. Zhao, and S. Chen, "A novel coarse-to-fine scheme for remote sensing image registration based on SIFT and phase correlation," *Remote Sens.*, vol. 11, no. 15, p. 1833, 2019. doi: 10.3390/rs11151833.

[59] Y. Han, F. Bovolo, and L. Bruzzone, "An approach to fine coregistration between very high resolution multispectral images based on registration noise distribution," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6650–6662, 2015. doi: 10.1109/TGRS.2015.2445632.

[60] A. Plyer, E. Colin-Koeniguer, and F. Weissgerber, "A new coregistration algorithm for recent applications on urban SAR images," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 11, pp. 2198–2202, 2015. doi: 10.1109/LGRS.2015.2455071.

[61] G. Brigot, E. Colin-Koeniguer, A. Plyer, and F. Janez, "Adaptation and evaluation of an optical flow method applied to coregistration of forest remote sensing images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 7, pp. 2923–2939, 2016. doi: 10.1109/JSTARS.2016.2578362.

[62] R. Feng, X. Li, and H. Shen, "Mountainous remote sensing images registration based on improved optical flow estimation," *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. IV-2/W5, pp. 479–484, June 2019. doi: 10.5194/isprs-annals-IV-2-W5-479-2019.

[63] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artif. Intell.*, vol. 17, nos. 1–3, pp. 185–203, 1981. doi: 10.1016/0004-3702(81)90024-2.

[64] J. J. Gibson, *The Perception of the Visual World*. Oxford: Houghton Mifflin, 1950.

[65] B. D. Lucas and T. Kanade, "An iterative image registration technique with an application to stereo vision," in *Proc. Imag. Understanding Workshop*, 1981, pp. 121–130.

[66] Z. Tu et al., "A survey of variational and CNN-based optical flow techniques," *Signal Processing: Image Commun.*, vol. 72, pp. 9–24, Mar. 2019. doi: 10.1016/j.image.2018.12.002.

[67] M. J. Black and P. Anandan, "The robust estimation of multiple motions: Parametric and piecewise-smooth flow fields," *Comput. Vision Image Understand.*, vol. 63, no. 1, pp. 75–104, 1996. doi: 10.1006/cviu.1996.0006.

[68] C. Liu, J. Yuen, and A. Torralba, "SIFT Flow: Dense correspondence across scenes and its applications," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 33, no. 5, pp. 978–994, 2011. doi: 10.1109/TPAMI.2010.147.

[69] T. Brox, A. Bruhn, N. Papenberg, and J. Weickert, "High accuracy optical flow estimation based on a theory for warping," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, 2004, pp. 25–36. doi: 10.1007/978-3-540-24673-2_3.

[70] J.-Y. Xiong, Y.-P. Luo, and G.-R. Tang, "An improved optical flow method for image registration with large-scale movements," *Acta Autom. Sin.*, vol. 34, no. 7, pp. 760–764, 2008. doi: 10.3724/SP.J.1004.2008.00760.

[71] A. Plyer, G. Le Besnerais, and F. Champagnat, "Massively parallel Lucas Kanade optical flow for real-time video processing applications," *J. Real-Time Image Process.*, vol. 11, no. 4, pp. 713–730, 2016. doi: 10.1007/s11554-014-0423-0.

[72] Y. Xiang, F. Wang, L. Wan, N. Jiao, and H. You, "OS-Flow: A robust algorithm for dense optical and SAR image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 1–20, 2019. doi: 10.1109/TGRS.2019.2905585.

[73] C. Huo, C. Pan, L. Huo, and Z. Zhou, "Multilevel SIFT matching for large-size VHR image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 2, pp. 171–175, 2012. doi: 10.1109/LGRS.2011.2163491.

[74] L. Yu, D. Zhang, and E.-J. Holden, "A fast and fully automatic registration approach based on point features for multi-source remote-sensing images," *Comput. Geosci.*, vol. 34, no. 7, pp. 838–848, 2008. doi: 10.1016/j.cageo.2007.10.005.

[75] L. Hui, B. S. Manjunath, and S. K. Mitra, "A contour-based approach to multisensor image registration," *IEEE Trans. Image Process.*, vol. 4, no. 3, pp. 320–334, 1995. doi: 10.1109/83.366480.

[76] H. Goncalves, L. Corte-Real, and J. A. Goncalves, "Automatic image registration through image segmentation and SIFT," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 7, pp. 2589–2600, 2011. doi: 10.1109/TGRS.2011.2109389.

[77] H. P. Moravec, "Techniques towards automatic visual obstacle avoidance," no. 2, p. 584, 1977. [Online]. Available: https://frc.ri.cmu.edu/~hpm/project.archive/robot.papers/1977/aip.txt

[78] C. Harris and M. Stephens, "A combined corner and edge detector," in *Proc. Alvey Vision Conf.*, Manchester, U.K., 1988, vol. 15, pp. 147–151.

[79] Y. Xiang, F. Wang, and H. You, "OS-SIFT: A robust SIFT-like algorithm for high-resolution optical-to-SAR image registration in suburban areas," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3078–3090, 2018. doi: 10.1109/TGRS.2018.2790483.

[80] I. Misra, S. M. Moorthi, D. Dhar, and R. Ramakrishnan, "An automatic satellite image registration technique based on Harris corner detection and Random Sample Consensus (RANSAC) outlier rejection model," in *1st Int. Conf. on Recent Advances in Information Technology (RAIT)*, 2012, pp. 68–73.

[81] S. M. Smith and J. M. Brady, "SUSAN–A new approach to low level image processing," *Int. J. Comput. Vis.*, vol. 23, no. 1, pp. 45–78, 1997. doi: 10.1023/A:1007963824710.

[82] C. Leng, H. Zhang, B. Li, G. Cai, Z. Pei, and L. He, "Local feature descriptor for image matching: A survey," *IEEE Access*, vol. 7, pp. 6424–6434, 2019. doi: 10.1109/ACCESS.2018.2888856.

[83] W. He and X. Deng, "A modified SUSAN corner detection algorithm based on adaptive gradient threshold for remote sensing image," in *Proc. Int. Conf. Optoelectron. Image Process.*, 2010, vol. 1, pp. 40–43.

[84] R. Feng, X. Li, W. Zou, and H. Shen, "Registration of multitemporal GF-1 remote sensing images with weighting perspective transformation model," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, 2017, pp. 2264–2268. doi: 10.1109/ICIP.2017.8296685.

[85] R. Feng, Q. Du, X. Li, and H. Shen, "Robust registration for remote sensing images by combining and localizing feature- and area-based methods," *ISPRS J. Photogram. Remote Sens.*, vol. 151, pp. 15–26, May 2019. doi: 10.1016/j.isprsjprs.2019.03.002.

[86] Y. Duan, X. Huang, J. Xiong, Y. Zhang, and B. Wang, "A combined image matching method for Chinese optical satellite imagery," *Int. J. Digital Earth*, vol. 9, no. 9, pp. 851–872, 2016. doi: 10.1080/17538947.2016.1151955.

[87] P. K. Konugurthi, R. Kune, R. Nooka, and V. Sarma, "Autonomous ortho-rectification of very high resolution imagery using SIFT and genetic algorithm," *Photogram. Eng. Remote Sens.*, vol. 82, no. 5, pp. 377–388, 2016. doi: 10.14358/PERS.82.5.377.

[88] Q. Li, G. Wang, J. Liu, and S. Chen, "Robust scale-invariant feature matching for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 287–291, 2009. doi: 10.1109/LGRS.2008.2011751.

[89] W. Ma et al., "Remote sensing image registration with modified SIFT and enhanced feature matching," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 1, pp. 3–7, 2017. doi: 10.1109/LGRS.2016.2600858.

[90] M. Gong, S. Zhao, L. Jiao, D. Tian, and S. Wang, "A novel coarse-to-fine scheme for automatic image registration based on SIFT and mutual information," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 7, pp. 4328–4338, 2014. doi: 10.1109/TGRS.2013.2281391.

[91] C. A. Lee, S. D. Gasster, A. Plaza, C. Chang, and B. Huang, "Recent developments in high performance computing for remote sensing: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 4, no. 3, pp. 508–527, 2011. doi: 10.1109/JSTARS.2011.2162643.

[92] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91–110, 2004. doi: 10.1023/B:VISI.0000029664.99615.94.

[93] K. Mikolajczyk and C. Schmid, "A performance evaluation of local descriptors," (in English), *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 10, pp. 1615–1630, Oct. 2005. doi: 10.1109/TPAMI.2005.188.

[94] K. Yan and R. Sukthankar, "PCA-SIFT: A more distinctive representation for local image descriptors," in *Proc. IEEE Comput. Soc. Conf. Comput. Vision Pattern Recognit. (CVPR)*, Washington, D. C., 2004, vol. 2, pp. 506–513. doi: 10.1109/CVPR.2004.1315206.

[95] Y. Zheng, Z. Cao, and Y. Xiao, "Multi-spectral remote image registration based on SIFT," *Electron. Lett.*, vol. 44, no. 2, pp. 107–108, 2008.

[96] J. Morel and G. Yu, "ASIFT: A new framework for fully affine invariant image comparison," *SIAM J. Imag. Sci.*, vol. 2, no. 2, pp. 438–469, 2009. doi: 10.1137/080732730.

[97] A. Sedaghat, M. Mokhtarzade, and H. Ebadi, "Uniform robust scale-invariant feature matching for optical remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 11, pp. 4516–4527, 2011. doi: 10.1109/TGRS.2011.2144607.

[98] A. Sedaghat and H. Ebadi, "Distinctive order based self-similarity descriptor for multi-sensor remote sensing image matching," *ISPRS J. Photogram. Remote Sens.*, vol. 108, pp. 62–71, Oct. 2015. doi: 10.1016/j.isprsjprs.2015.06.003.

[99] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Graz, Austria, 2006, pp. 404–417.

[100] W. Yan, H. She, and Z. Yuan, "Robust registration of remote sensing image based on SURF and KCCA," *J. Indian Soc. Remote Sens.*, vol. 42, no. 2, pp. 291–299, 2014. doi: 10.1007/s12524-013-0324-x.

[101] X. Yuan, S. Chen, W. Yuan, and Y. Cai, "Poor textural image tie point matching via graph theory," *ISPRS J. Photogram. Remote Sens.*, vol. 129, pp. 21–31, July 2017. doi: 10.1016/j.isprsjprs.2017.04.015.

[102] R. Bouchiha and K. Besbes, "Automatic remote-sensing image registration using SURF," *Int. J. Comput. Theory Eng.*, vol. 5, no. 1, pp. 88–92, 2013. doi: 10.7763/IJCTE.2013.V5.653.

[103] J. Chen et al., "WLD: A robust local image descriptor," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 9, pp. 1705–1720, 2010. doi: 10.1109/TPAMI.2009.155.

[104] E. Rosten and T. Drummond, "Machine learning for high-speed corner detection," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Graz, Austria, 2006, pp. 430–443.

[105] M. Calonder, V. Lepetit, C. Strecha, and P. Fua, "BRIEF: Binary robust independent elementary features," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Crete, Greece, 2010, pp. 778–792.

[106] E. Rublee, V. Rabaud, K. Konolige, and G. Bradski, "ORB: An efficient alternative to SIFT or SURF," in *Proc. Int. Conf. Comput. Vision (ICCV)*, Barcelona, Spain, 2011, pp. 2564–2571.

[107] D. Ma and H. Lai, "Remote sensing image matching based improved ORB in NSCT domain," *J. Indian soc. Remote Sens.*, vol. 47, no. 5, pp. 801–807, 2019. doi: 10.1007/s12524-019-00958-y.

[108] P. F. Alcantarilla, A. Bartoli, and A. J. Davison, "KAZE Features," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Florence, Italy, 2012, pp. 214–227.

[109] P. Alcantarilla, J. Nuevo, and A. Bartoli, "Fast explicit diffusion for accelerated features in nonlinear scale spaces," in *Proc. Brit. Mach. Vision Conf. (BMVC)*, Bristol, U.K., 2013, pp. 1–11.

[110] Y. Ye, M. Wang, S. Hao, and Q. Zhu, "A novel keypoint detector combining corners and blobs for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 18, no. 3, pp. 451–455, Mar. 31, 2020. doi: 10.1109/LGRS.2020. 2980620.

[111] X. Liu, Y. Ai, J. Zhang, and Z. Wang, "A novel affine and contrast invariant descriptor for infrared and visible image registration," *Remote Sens.*, vol. 10, no. 4, p. 658, 2018. doi: 10.3390/rs10040658.

[112] Z. Ye et al., "Robust fine registration of multisensor remote sensing images based on enhanced subpixel phase correlation," *Sensors*, vol. 20, no. 15, p. 4338, Aug. 4, 2020. doi: 10.3390/s20154338.

[113] Y. C. Hsieh, D. M. McKeown, and F. P. Perlant, "Performance evaluation of scene registration and stereo matching for cartographic feature extraction," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 14, no. 2, pp. 214–238, 1992. doi: 10.1109/34.121790.

[114] S. Dongseok, J. K. Pollard, and J. Muller, "Accurate geometric correction of ATSR images," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 4, pp. 997–1006, 1997. doi: 10.1109/36.602542.

[115] J. Inglada and F. Adragna, "Automatic multi-sensor image registration by edge matching using genetic algorithms," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, Sydney, NSW, Australia, 2001, vol. 5, pp. 2313–2315.

[116] W. Shi and A. Shaker, "The Line-Based Transformation Model (LBTM) for image-to-image registration of high-resolution satellite image data," *Int. J. Remote Sens.*, vol. 27, no. 14, pp. 3001–3012, 2006. doi: 10.1080/01431160500486716.

[117] T.-Z. Xiang, G.-S. Xia, X. Bai, and L. Zhang, "Image stitching by line-guided local warping with global similarity constraint," *Pattern Recognit.*, vol. 83, pp. 481–497, Nov. 2018. doi: 10.1016/j.patcog.2018.06.013.

[118] C. Zhao and A. A. Goshtasby, "Registration of multitemporal aerial optical images using line features," *ISPRS J. Photogram. Remote Sens.*, vol. 117, pp. 149–160, July 2016. doi: 10.1016/j.isprsjprs.2016.04.002.

[119] C. Li and W. Shi, "The generalized-line-based iterative transformation model for imagery registration and rectification," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 8, pp. 1394–1398, 2014. doi: 10.1109/LGRS.2013.2293844.

[120] A. O. Ok, J. D. Wegner, C. Heipke, F. Rottensteiner, U. Soergel, and V. Toprak, "Matching of straight line segments from aerial stereo images of urban areas," *ISPRS J. Photogram. Remote Sens.*, vol. 74, pp. 133–152, Nov. 2012. doi: 10.1016/j.isprsjprs.2012.09.003.

[121] J. Canny, "A computational approach to edge detection," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-8, no. 6, pp. 679–698, 1986. doi: 10.1109/TPAMI.1986.4767851.

[122] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. Roy. Soc. Ser. B-Biol. Sci.*, vol. 207, no. 1167, pp. 187–217, 1980. doi: 10.1098/rspb.1980.0020.

[123] R. G. v. Gioi, J. Jakubowicz, J. Morel, and G. Randall, "LSD: A fast line segment detector with a false detection control," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 32, no. 4, pp. 722–732, 2010. doi: 10.1109/TPAMI.2008.300.

[124] C. Akinlar and C. Topal, "EDLines: A real-time line segment detector with a false detection control," *Pattern Recog. Lett.*, vol. 32, no. 13, pp. 1633–1642, 2011. doi: 10.1016/j.patrec.2011.06.001.

[125] A. Goshtasby and G. C. Stockman, "Point pattern matching using convex hull edges," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-15, no. 5, pp. 631–637, 1985. doi: 10.1109/TSMC.1985.6313439.

[126] W. Dorigo, M. Hollaus, W. Wagner, and K. Schadauer, "An application-oriented automated approach for co-registration of forest inventory and airborne laser scanning data," *Int. J. Remote Sens.*, vol. 31, no. 5, pp. 1133–1153, 2010. doi: 10.1080/01431160903380581.

[127] B. Sirmacek and C. Unsalan, "Urban-area and building detection using SIFT keypoints and graph theory," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1156–1167, 2009. doi: 10.1109/TGRS.2008.2008440.

[128] J. Flusser and T. Suk, "A moment-based approach to registration of images with affine geometric distortion," *IEEE Trans. Geosci. Remote Sens.*, vol. 32, no. 2, pp. 382–387, 1994. doi: 10.1109/36.295052.

[129] N. R. Pal and S. K. Pal, "A review on image segmentation techniques," *Pattern Recognit.*, vol. 26, no. 9, pp. 1277–1294, 1993. doi: 10.1016/0031-3203(93)90135-J.

[130] D. Xiaolong and S. Khorram, "Development of a feature-based approach to automated image registration for multitemporal and multisensor remotely sensed imagery," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. Proc. Remote Sens.-A Sci. Vision Sustainable Develop. (IGARSS)*, 1997, vol. 1, pp. 243–245.

[131] L. M. Fonseca and B. Manjunath, "Registration techniques for multisensor remotely sensed imagery," *Photogram. Eng. Remote Sensing (PERS)*, vol. 62, no. 9, pp. 1049–1056, 1996.

[132] A. Goshtasby, G. C. Stockman, and C. V. Page, "A region-based approach to digital image registration with subpixel accuracy," *IEEE Trans. Geosci. Remote Sens.*, vol. GE-24, no. 3, pp. 390–399, 1986. doi: 10.1109/TGRS.1986.289597.

[133] J. Ton and A. K. Jain, "Registering Landsat images by point matching," *IEEE Trans. Geosci. Remote Sens.*, vol. 27, no. 5, pp. 642–651, 1989. doi: 10.1109/TGRS.1989.35948.

[134] Y. Chen, X. Zhang, Y. Zhang, S. J. Maybank, and Z. Fu, "Visible and infrared image registration based on region features and edginess," *Mach. Vis. Appl.*, vol. 29, no. 1, pp. 113–123, 2018. doi: 10.1007/s00138-017-0879-6.

[135] A. Irani Rahaghi, U. Lemmin, D. Sage, and D. A. Barry, "Achieving high-resolution thermal imagery in low-contrast lake surface waters by aerial remote sensing and image registration," *Remote Sens. Environ.*, vol. 221, pp. 773–783, Feb. 2019. doi: 10.1016/j.rse.2018.12.018.

[136] A. Li, X. Cheng, H. Guan, T. Feng, and Z. Guan, "Novel image registration method based on local structure constraints," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 9, pp. 1584–1588, 2014. doi: 10.1109/LGRS.2014.2305982.

[137] S. Jiang and W. Jiang, "Hierarchical motion consistency constraint for efficient geometrical verification in UAV stereo image matching," *ISPRS J. Photogram. Remote Sens.*, vol. 142, pp. 222–242, Aug. 2018. doi: 10.1016/j.isprsjprs.2018.06.009.

[138] J. S. Beis and D. G. Lowe, "Shape indexing using approximate nearest-neighbour search in high-dimensional spaces," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 1997, vol. 97, pp. 1000–1006. doi: 10.1109/CVPR.1997.609451.

[139] Y. Ma et al., "Remote sensing big data computing: Challenges and opportunities," *Future Gener. Comput. Syst.*, vol. 51, pp. 47–60, Oct. 2015. doi: 10.1016/j.future.2014.10.029.

[140] G. Stockman, S. Kopstein, and S. Benett, "Matching images to models for registration and object detection via clustering," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. PAMI-4, no. 3, pp. 229–241, 1982. doi: 10.1109/TPAMI.1982.4767240.

[141] G. Borgefors, "Hierarchical chamfer matching: A parametric edge matching algorithm," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 10, no. 6, pp. 849–865, 1988. doi: 10.1109/34.9107.

[142] L. Livi and A. Rizzi, "The graph matching problem," *Pattern Anal. Appl.*, vol. 16, no. 3, pp. 253–283, 2013. doi: 10.1007/s10044-012-0284-8.

[143] L. Torresani, V. Kolmogorov, and C. Rother, "Feature correspondence via graph matching: models and global optimization," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Berlin, Heidelberg, 2008, pp. 596–609.

[144] Z. Liu, J. An, and Y. Jing, "A simple and robust feature point matching algorithm based on restricted spatial order constraints for aerial image registration," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 2, pp. 514–527, 2012. doi: 10.1109/TGRS.2011.2160645.

[145] Z. Xiong and Y. Zhang, "A novel interest-point-matching algorithm for high-resolution satellite images," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 12, pp. 4189–4200, 2009. doi: 10.1109/TGRS.2009.2023794.

[146] H. Chang, G. Wu, and M. Chiang, "Remote sensing image registration based on modified SIFT and feature slope grouping," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 9, pp. 1363–1367, 2019. doi: 10.1109/LGRS.2019.2899123.

[147] S. Zhili and Z. Jiaqi, "Image registration approach with scale-invariant feature transform algorithm and tangent-crossing-point feature," *J. Electron. Imag.*, vol. 29, no. 2, pp. 1–14, Mar. 2020. doi: 10.1117/1.JEI.29.2.023010.

[148] M. A. Fischler and R. C. Bolles, "Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981. doi: 10.1145/358669.358692.

[149] K. Zhang, X. Li, and J. Zhang, "A robust point-matching algorithm for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 2, pp. 469–473, 2014. doi: 10.1109/LGRS.2013.2267771.

[150] Y. Wu, W. Ma, M. Gong, L. Su, and L. Jiao, "A novel point-matching algorithm based on fast sample consensus for image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 43–47, 2015. doi: 10.1109/LGRS.2014.2325970.

[151] B. Li and H. Ye, "RSCJ: Robust sample consensus judging algorithm for remote sensing image registration," *IEEE Geosci. Remote Sens. Lett.*, vol. 9, no. 4, pp. 574–578, 2012. doi: 10.1109/LGRS.2011.2175434.

[152] H. Zhang et al., "Remote sensing image registration based on local affine constraint with circle descriptor," *IEEE Geosci. Remote Sens. Lett.*, early access, 2020. doi: 10.1109/LGRS.2020.3027096.

[153] F. Ye, Y. Su, H. Xiao, X. Zhao, and W. Min, "Remote sensing image registration using convolutional neural network features," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 2, pp. 232–236, 2018. doi: 10.1109/LGRS.2017.2781741.

[154] C. Farabet, C. Couprie, L. Najman, and Y. LeCun, "Learning hierarchical features for scene labeling," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 8, pp. 1915–1929, 2013. doi: 10.1109/TPAMI.2012.231.

[155] W. Huang, L. Xiao, Z. Wei, H. Liu, and S. Tang, "A new pan-sharpening method with deep neural networks," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 5, pp. 1037–1041, 2015. doi: 10.1109/LGRS.2014.2376034.

[156] Y. Xing, M. Wang, S. Yang, and L. Jiao, "Pan-sharpening via deep metric learning," *ISPRS J. Photogram. Remote Sens.*, vol. 145, pp. 165–183, Nov. 2018. doi: 10.1016/j.isprsjprs.2018.01.016.

[157] G. J. Scott, M. R. England, W. A. Starms, R. A. Marcum, and C. H. Davis, "Training deep convolutional neural networks for land–cover classification of high-resolution imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 4, pp. 549–553, 2017. doi: 10.1109/LGRS.2017.2657778.

[158] L. Ma, Y. Liu, X. Zhang, Y. Ye, G. Yin, and B. A. Johnson, "Deep learning in remote sensing applications: A meta-analysis and review," *ISPRS J. Photogram. Remote Sens.*, vol. 152, pp. 166–177, June 2019. doi: 10.1016/j.isprsjprs.2019.04.015.

[159] Y. Liu, D. Minh Nguyen, N. Deligiannis, W. Ding, and A. Munteanu, "Hourglass-shapenetwork based semantic segmentation for high resolution aerial imagery," *Remote Sens.*, vol. 9, no. 6, p. 522, 2017. doi: 10.3390/rs9060522.

[160] H. Zhang et al., "Registration of multimodal remote sensing image based on deep fully convolutional neural network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 8, pp. 3028–3042, 2019. doi: 10.1109/JSTARS.2019.2916560.

[161] N. Merkle, W. Luo, S. Auer, R. Müller, and R. Urtasun, "Exploiting deep matching and SAR data for the geo-localization accuracy improvement of optical satellite images," *Remote Sensing*, vol. 9, no. 6, p. 586, 2017. doi: 10.3390/rs9060586.

[162] H. He, M. Chen, T. Chen, and D. Li, "Matching of remote sensing images with complex background variations via Siamese convolutional neural network," *Remote Sens.*, vol. 10, no. 3, p. 355, 2018. doi: 10.3390/rs10020355.

[163] L. H. Hughes, M. Schmitt, L. Mou, Y. Wang, and X. X. Zhu, "Identifying corresponding patches in SAR and optical images with a pseudo-Siamese CNN," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 5, pp. 784–788, 2018. doi: 10.1109/LGRS.2018.2799232.

[164] S. Wang, D. Quan, X. Liang, M. Ning, Y. Guo, and L. Jiao, "A deep learning framework for remote sensing image registration," *ISPRS J. Photogram. Remote Sens.*, vol. 145, pp. 148–164, Nov. 2018. doi: 10.1016/j.isprsjprs.2017.12.012.

[165] R. Fan, B. Hou, J. Liu, J. Yang, and Z. Hong, "Registration of multi-resolution remote sensing images based on L2-Siamese model," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 1–1, Nov. 19, 2020. doi: 10.1109/JSTARS.2020.3038922.

[166] N. Merkle, S. Auer, R. Müller, and P. Reinartz, "Exploring the potential of conditional adversarial networks for optical and SAR image matching," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 6, pp. 1811–1820, 2018. doi: 10.1109/JSTARS.2018.2803212.

[167] H. L. Hughes, M. Schmitt, and X. X. Zhu, "Mining hard negative samples for SAR-optical image matching using generative adversarial networks," *Remote Sens.*, vol. 10, no. 10, p. 1552, 2018. doi: 10.3390/rs10101552.

[168] J. Zhang, W. Ma, Y. Wu, and L. Jiao, "Multimodal remote sensing image registration based on image transfer and local features," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 8, pp. 1210–1214, 2019. doi: 10.1109/LGRS.2019.2896341.

[169] N. Girard, G. Charpiat, and Y. Tarabalka, "Aligning and updating cadaster maps with aerial images by multi-task, multi-resolution deep learning," in *Proc. Asian Conf. Comput. Vision (ACCV 2018)*, Cham, 2019, pp. 675–690.

[170] L. Li, L. Han, M. Ding, Z. Liu, and H. Cao, "Remote sensing image registration based on deep learning regression model," *IEEE Geosci. Remote Sens. Lett.*, early access, 2020. doi: 10.1109/LGRS.2020.3032439.

[171] F. Liu, F. Bi, L. Chen, H. Shi, and W. Liu, "Feature-area optimization: A novel SAR image registration method," *IEEE Geosci. Remote Sens. Lett.*, vol. 13, no. 2, pp. 242–246, 2016. doi: 10.1109/LGRS.2015.2507982.

[172] X. Huang, Y. Sun, D. Metaxas, F. Sauer, and C. Xu, "Hybrid image registration based on configural matching of scale-invariant salient region features," in *Proc. IEEE Comput. Society Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Washington, D. C., 2004, pp. 167–167. doi: 10.1109/CVPR.2004.362.

[173] G. Hong and Y. Zhang, "Combination of feature-based and area-based image registration technique for high resolution remote sensing image," in *Proc. Int. Geosci. Remote Sens. Symp. (IGARSS)*, Barcelona, Spain, 2007, pp. 377–380.

[174] N. E. Mekky, F. E.-Z. Abou-Chadi, and S. Kishk, "Wavelet-based image registration techniques: A study of performance," *Int. J. Comput. Sci. Netw. Security*, vol. 11, no. 2, pp. 188–196, 2011.

[175] S. Suri, P. Schwind, P. Reinartz, and J. Uhl, "Combining mutual information and scale invariant feature transform for fast and robust multisensor SAR image registration," in *Proc. 75th Annu. ASPRS Conf.*, 2009.

[176] Y. S. Heo, K. M. Lee, and S. U. Lee, "Joint depth map and color consistency estimation for stereo images with different illuminations and cameras," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 35, no. 5, pp. 1094–1106, 2013. doi: 10.1109/TPAMI.2012.167.

[177] R. Feng, Q. Du, H. Shen, and X. Li, "Region-by-region registration combining feature-based and optical flow methods for remote sensing images," *Remote Sens.*, vol. 13, no. 8, p. 1475, 2021. doi: 10.3390/rs13081475.

[178] C. Xing, J. Wang, and Y. Xu, "A method for building a mosaic with UAV images," *Int. J. Inform. Eng. Electron. Bus.*, vol. 2, no. 1, pp. 9–15, 2010. doi: 10.5815/ijieeb.2010.01.02.

[179] Z. Kang, L. Zhang, S. Zlatanova, and J. Li, "An automatic mosaicking method for building facade texture mapping using a monocular close-range image sequence," *ISPRS J. Photogram. Remote Sens.*, vol. 65, no. 3, pp. 282–293, 2010. doi: 10.1016/j.isprsjprs.2009.11.003.

[180] S. R. Lee, "A coarse-to-fine approach for remote-sensing image registration based on a local method," *Int. J. Smart Sens. Intell. Syst.*, vol. 3, no. 4, 2010.

[181] K. Sharma and A. Goyal, "Very high resolution image registration based on two step Harris-Laplace detector and SIFT descriptor," in *2013 4th Int. Conf. Comput., Commun. Netw. Technol. (ICCCNT)*, pp. 1–5. doi: 10.1109/ICCCNT.2013.6726632.

[182] W. Ma, J. Zhang, Y. Wu, L. Jiao, H. Zhu, and W. Zhao, "A novel two-step registration method for remote sensing images based on deep and local features," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4834–4843, 2019. doi: 10.1109/TGRS.2019.2893310.

[183] S. Li, L. Yuan, J. Sun, and L. Quan, "Dual-feature warping-based motion model estimation," in *Proc. IEEE Int. Conf. Comput. Vision (ICCV)*, 2015, pp. 4283–4291.

[184] S. Nag, "Image registration techniques: A survey," Nov. 28, 2017, arXiv:1712.07540.

[185] J. S. Bhatt and N. Padmanabhan, "Image Registration for meteorological applications: Development of a generalized software for sensor data registration at ISRO," *IEEE Geosci. Remote Sens. Mag. (replaces Newslett.)*, vol. 8, no. 4, pp. 23–37, 2020. doi: 10.1109/MGRS.2019.2949382.

[186] A. Sedaghat and N. Mohammadi, "High-resolution image registration based on improved SURF detector and localized GTM," *Int. J. Remote Sens.*, vol. 40, no. 7, pp. 2576–2601, Apr. 2019. doi: 10.1080/01431161.2018.1528402.

[187] Y. Ma, J. Wang, H. Xu, S. Zhang, X. Mei, and J. Ma, "Robust image feature matching via progressive sparse spatial consensus," *IEEE Access*, vol. 5, pp. 24,568–24,579, Oct. 2017. doi: 10.1109/ACCESS.2017.2768078.

[188] H. Goncalves, J. A. Goncalves, and L. Corte-Real, "Measures for an objective evaluation of the geometric correction process quality," *IEEE Geosci. Remote Sens. Lett.*, vol. 6, no. 2, pp. 292–296, 2009. doi: 10.1109/LGRS.2008.2012441.

GRS