Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Fast uniform content-based satellite image registration using the scale-invariant feature transform descriptor

Hamed BOZORGI[1], Ali JAFARI[†‡2]

([1]*Electrical Engineering Department, University of Guilan, Rasht 41635-3756, Iran*)

([2]*School of Electrical & Electronic Engineering, Malek Ashtar University of Technology, Tehran 15875-1774, Iran*)

[†]E-mail: iustuser@aut.ac.ir

Received Sept. 5, 2015; Revision accepted Feb. 17, 2016; Crosschecked Aug. 1, 2017

Abstract: Content-based satellite image registration is a difficult issue in the fields of remote sensing and image processing. The difficulty is more significant in the case of matching multisource remote sensing images which suffer from illumination, rotation, and source differences. The scale-invariant feature transform (SIFT) algorithm has been used successfully in satellite image registration problems. Also, many researchers have applied a local SIFT descriptor to improve the image retrieval process. Despite its robustness, this algorithm has some difficulties with the quality and quantity of the extracted local feature points in multisource remote sensing. Furthermore, high dimensionality of the local features extracted by SIFT results in time-consuming computational processes alongside high storage requirements for saving the relevant information, which are important factors in content-based image retrieval (CBIR) applications. In this paper, a novel method is introduced to transform the local SIFT features to global features for multisource remote sensing. The quality and quantity of SIFT local features have been enhanced by applying contrast equalization on images in a pre-processing stage. Considering the local features of each image in the reference database as a separate class, linear discriminant analysis (LDA) is used to transform the local features to global features while reducing dimensionality of the feature space. This will also significantly reduce the computational time and storage required. Applying the trained kernel on verification data and mapping them showed a successful retrieval rate of 91.67% for test feature points.

Key words: Content-based image retrieval; Feature point distribution; Image registration; Linear discriminant analysis; Remote sensing; Scale-invariant feature transform
http://dx.doi.org/10.1631/FITEE.1500295 **CLC number:** TP391

1 Introduction

With the advances in the field of image capturing devices, both the quantity and quality of images are continuing to grow at a rapid pace. Description of image content and extraction of information from images have become essential tasks in the field of image processing (Li and Narayanan, 2004; Bhatia *et al.*, 2007). Furthermore, satellite images in the terabyte range are being generated everyday by Earth observation satellites. These images are interpreted by a variety of users in fields such as agriculture, geology, ecology, and security (Eugenio and Marqués, 2003; Blanchart and Datcu, 2010; Bhattacharya and Bhaskar, 2013; Liu *et al.*, 2013). In the field of satellite image registration, there is a great necessity to retrieve images based on specified criteria.

Content-based image registration (CBIR) methods use image contents (such as color, shape, texture, or contrast) to create features for an image in a database (Goswami *et al.*, 2007; Maloo *et al.*, 2010; Nileshsingh and Koli, 2010; Irtazaa *et al.*, 2013). The features extracted from a test image are then matched against the stored index to determine the degree of match between the test and each of the database images. Generally, there are two types of image features:

‡ Corresponding author
ORCID: Ali JAFARI, http://orcid.org/0000-0003-2644-1755

global and local. Global features are global image properties, such as color, intensity histogram, texture patterns, and shape (Torres *et al.*, 2003; Nileshsingh and Koli, 2010; Bakar *et al.*, 2013; Irtazaa *et al.*, 2013). In most CBIR applications, global features are successful. They provide powerful descriptors that represent image properties (Stehling *et al.*, 2002; Arica and Vural, 2003; Torres *et al.*, 2003; Bakar *et al.*, 2013; Irtazaa *et al.*, 2013). The most important advantages of global features are their low computational costs and acceptable performances (Bakar *et al.*, 2013). However, these global features may not be able to satisfy other specific image requirements. In the case of satellite images, global features face greater difficulties. Lack of color in some satellite images, differences in scale and viewpoint, rotation, and difficulties in determining specific objects as features are some problems limiting the use of global features in remote sensing (Stehling *et al.*, 2002; Arica and Vural, 2003; Torres *et al.*, 2003; Veganzones *et al.*, 2008; Bakar *et al.*, 2013). On the other hand, local features are local image properties related to local image regions, such as edges, corners, lines, and curves (Mikolajczyk and Schmid, 2005; Meena *et al.*, 2014). CBIR systems that use local features are able to represent details of an object inside an image; however, it requires a time-consuming computational process and a huge amount of storage. Such systems are able to detect the object under occlusions, and most of them are invariant to changes in terms of scale and rotation. Recently, combined systems that use both local and global features have been introduced for better image retrieval (Wu and Wu, 2009; Li *et al.*, 2013; Singaravelan and Murugan, 2013). Although such systems have shown promising results, they still suffer from local feature problems.

The scale-invariant feature transform (SIFT) is one of the most successful techniques. It has been used for local feature point detection and description (Lowe, 2004). Dagher *et al.* (2014) used the SIFT algorithm in the field of face recognition to extract local features from different images of a person and divided the results into different regions using the *k*-means clustering method. On the basis of the features they obtained, they proposed a method to obtain the most representative images of each face. Even though it has been proved that the SIFT algorithm is successful in casual image matching (Lowe, 2004; Dagher *et al.*, 2014; Okade and Biswas, 2014), this algorithm suffers from problems with the quality and quantity of extracted features, particularly in multi-source remote sensing imaging. Different methods have been applied to improve the performance of the SIFT algorithm in this field (Wen *et al.*, 2008; Sedaghat *et al.*, 2011; Bakar *et al.*, 2013; Kupfer *et al.*, 2015). Yu *et al.* (2008) applied affine transformation and the SIFT algorithm in a preregistration stage followed by pyramidal Harris point extraction and correspondence for fully automatic and fast multisource satellite image registration (Manuel *et al.*, 2013).

Furthermore, the SIFT algorithm, along with several different methods, has been proposed recently to achieve the goals of CBIR (Chu *et al.*, 1996; Andre *et al.*, 2012; Bakar *et al.*, 2013; Manuel *et al.*, 2013). Most methods in this field suffer from high dimensionality, which leads to computationally intensive processes and the demand of subsequently massive amounts of memory. To counteract this problem, some CBIR methods have been proposed to reduce the dimensionality of feature extraction methods (Li *et al.*, 2011; Manuel *et al.*, 2013; Chandrakanth *et al.*, 2014). Valenzuela *et al.* (2012) applied different dimensionality reduction techniques to reduce the dimensions of SIFT and SURF feature vectors while maintaining the image retrieval accuracy.

In this paper, we introduce a novel method that benefits from the advantages of both local and global features in the field of satellite image retrieval. The SIFT algorithm is used for extracting and describing local features from images captured from various sensors and satellites. As thresholds in this algorithm cause serious differences in the number of features extracted from different types of remote sensor images, a pre-processing stage is considered. Linear discriminant analysis (LDA) is used to transform the local image features to global while reducing their dimensionality. Thus, instead of individually matching each feature of a query image against features of reference images, its transformed feature will be matched in a new feature space. This also decreases the required computational time and storage for CBIR applications.

2 Scale-invariant feature transform algorithm review

The SIFT algorithm proposed by Lowe (2004) consists of three main steps: local feature extraction and selection, feature description, and feature matching.

In the first step, points that are invariant to the scale change of the image are detected as potential feature points. This is carried out by searching for stable features among all possible scales using a scale function known as the scale space. It is created by applying a Gaussian kernel function $G(x, y, \sigma)$ with different values of σ onto the original image. The scale space consists of blurred and subsampled versions of the reference image. The blurred image, $L(x, y, \sigma)$, which is the result of convolution between the original image $I(x, y)$ and the Gaussian kernel function $G(x, y, \sigma)$, is given by

$$L(x,y,\sigma)=G(x,y,\sigma)\cdot I(x,y), \tag{1}$$

where

$$G(x,y,\sigma)=\frac{1}{2\pi\sigma^2}\exp\left(\frac{-(x^2+y^2)}{2\sigma^2}\right). \tag{2}$$

To create more blurred images, the operation continues with $G(x, y, k\sigma)$. A difference-of-Gaussian (DOG) image, which is the result of the difference between two blurred images, is given by

$$D(x,y,\sigma)=L(x,y,k\sigma)-L(x,y,\sigma). \tag{3}$$

The scale space of the SIFT algorithm is shown in Fig. 1. In each octave, scale-space images are separated by a constant factor k. Blurred images that are next to each other are then subtracted to produce the DOG images. After each octave, the blurred image that has twice the value of σ is downsampled by a factor of two, and the next octave is generated by repeating the steps used to generate the first octave.

The initial value for σ and the number of scale layers in each octave suggested by Lowe (2004) are 1.6 and 3, respectively, and the constant factor is $k=2^{1/L_N}$, where L_N is the number of scale layers in each octave. The middle layer of the DOG images in each octave is used for feature extraction. Detecting potential feature points is accomplished by using the scale-space extremum (Lowe, 2004). The value of each pixel in a scale layer is compared against its eight adjacent pixels in the same layer, nine neighbors in the higher scale layer, and nine neighbors in the lower scale layer. In the case of there being an extremum among these 26 pixels, the pixels will be chosen as a local potential feature point. After the local potential feature points are determined, a detailed fit will be applied to their nearby data. The results will show the exact location, scale, and ratio of principle curvatures. Using this information, the candidate feature points that have a low contrast or are poorly located along an edge will be removed. By applying a 3D quadratic polynomial to the scale space DoG function and finding the extremum, the subsampled accurate position and scale for each feature point will be computed. The absolute DoG extremum value illustrates the contrast of the candidate feature point. Low-contrast feature points whose extremum of sensitivity is less than a threshold (e.g., T_C=0.3) are unstable and sensitive to noise and will be rejected. In the next step, extrema corresponding to the edges are discarded using curvature analysis.

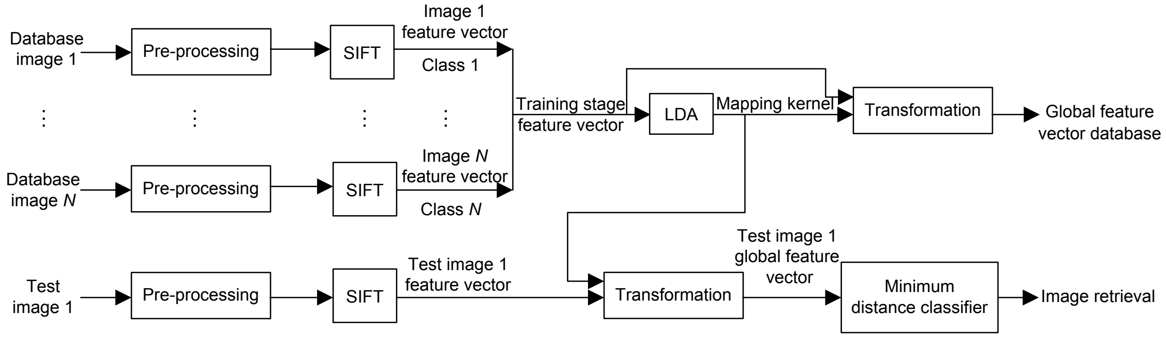

Fig. 1 Block diagram of the proposed method (LDA: linear discriminant analysis; SIFT: scale-invariant feature transform)

Potential feature points that are located along an edge will have a large principle curvature, while the ones that are on the perpendicular direction will have small values. Lowe (2004) suggested a threshold, T_r=10, to eliminate the candidate feature points that are poorly localized along an edge.

To satisfy the image invariance characteristics of the SIFT algorithm, a dominant orientation created using local gradient direction, is assigned to each feature. Dominant orientation is determined by building a histogram of gradient orientations from feature point neighbors, weighted by a Gaussian and the gradient magnitude.

The second step of the algorithm is the SIFT description. In this stage, a representation for each local key point is built on the basis of a patch of pixels in its neighborhood. The SIFT descriptor is a histogram of gradient locations and orientations where the location is quantized into a 4×4 location grid around the key point, and the gradient angle is quantized into eight main orientations. The results would be made up of a 128-dimensional and highly distinctive descriptor matrix. Then, the dominant orientation calculated in the first stage is used to rotate the descriptor coordinates and gradient orientations to achieve the orientation invariant features.

As a final step, to match the reference and the test image, a minimum Euclidean distance criterion is used to determine the degree of correspondence among feature descriptors. Further details about the SIFT algorithm can be found in Lowe (2004).

3 Methodology

In this section, an efficient and robust automatic method for accurate content-based image retrieval in optical satellite remote sensing images is presented. The proposed method can be divided into three major steps: pre-processing, feature extraction and transformation, and feature classification. Fig. 1 illustrates a block diagram relating to the proposed method. Inspired by the SIFT algorithm, the proposed technique uses a methodology which contains a pre-processing stage and a transforming stage in which local features produced by a SIFT key point detector and descriptor are transformed into general-type image features using LDA.

3.1 Pre-processing

The complexity of satellite images makes it difficult for the SIFT algorithm to determine a sufficient number of feature points. This has led to the use of SIFT in the preregistration stage for many remote sensing imagery related works (Yu *et al.*, 2008; Sedaghat *et al.*, 2012).

The number of features extracted from some satellite image sensors is usually redundant, while for others this number would be much less than required. This is due to the sensitivity of SIFT to its parameters, particularly to the contrast threshold (T_C), which controls the number of extracted feature points.

In Sedaghat *et al.* (2011), the sensitivity of T_C to the number of extracted feature points was shown for different remote sensing images from different scene conditions, such as homogenous regions and structured urban scenes. The results proved that the SIFT algorithm is highly sensitive to this threshold, and that it has no optimum value.

In the case of satellite imaging, the distribution of feature points has a great influence on reliability registration. The SIFT algorithm suffers from lack of controllability of the spatial distribution of extracted feature points. This problem especially intensifies in the case of images of terrain, such as deserts which have few landmarks.

Extracting a sufficient number of feature points at all possible scales is essential for a successful image registration. In this study, uniform and more robust feature extraction from different satellite image sensors is achieved through a re-processing stage. In this stage, to reduce the SIFT sensitivity, the contrast of each image is normalized and equalized. This method usually increases the global contrast of many images, especially when the usable data of the image are represented by similar contrast values. This is very important in the case of remote sensing images. It allows areas of lower local contrast to gain a higher contrast. It can be accomplished by effectively spreading out the most frequent intensity values.

Consider a discrete grayscale image $\{x\}$ represented by an $r×c$ matrix of integer pixel intensities ranging from n=0 to $L-1$. Let P_n be the probability of an occurrence of a pixel with an intensity level of n in image $\{x\}$. Considering all possible intensities:

$$P_n = \frac{n}{\text{Total number of pixels}}, \tag{4}$$

where n=0, 1, …, L−1. A transformed image is given by

$$g_{i,j} = \left\lfloor (L-1)\sum_{n=0}^{x_{i,j}} P_n \right\rfloor. \tag{5}$$

A key advantage of this method is that it is a fairly straightforward technique and the calculation is not computationally intensive.

Figs. 2 and 3 illustrate the distribution of local SIFT features in both the image space and the scale space extracted from raw and pre-processed IRS-P6 images, respectively.

3.2 Feature extraction, transformation, and classification

The purpose of the feature extraction stage is to obtain the best features to represent an image and to include the most amount information about it. In this study, we present a new method to transform the local features to robust global features for satellite image retrieval. In the first step, each image in a reference database is pre-processed, and then its related local key points are determined using the SIFT key point detector. Using the SIFT descriptor, a 128-dimensional feature vector is created for each local feature point. Considering N as the number of reference images in the database and L_N as the number of feature points extracted from the Nth image, in a normal SIFT matching stage, the descriptor of each local feature point obtained from the query image has

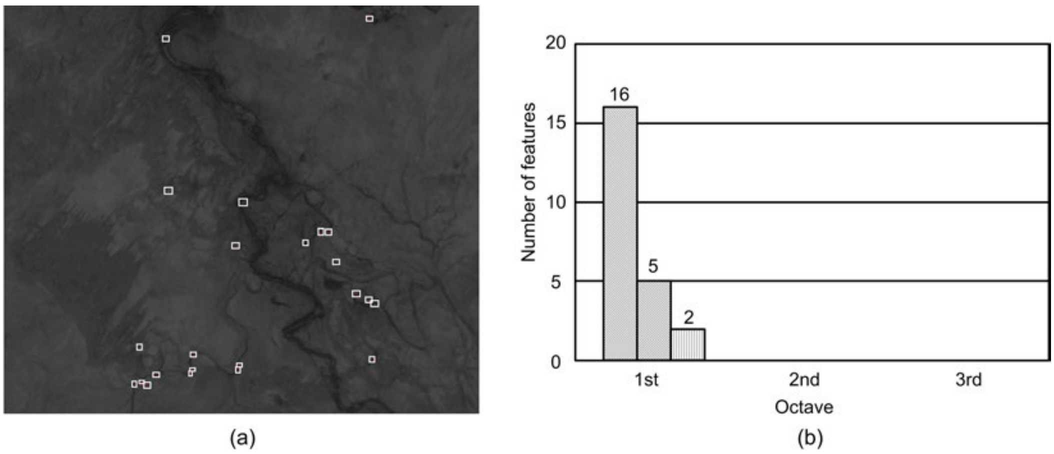

Fig. 2 Insufficient and unevenly distributed features extracted from a raw IRS image captured with the LISS III sensor using a standard scale-invariant feature transform algorithm in the image space (a) and the scale space (b) (The image size is 445×518)

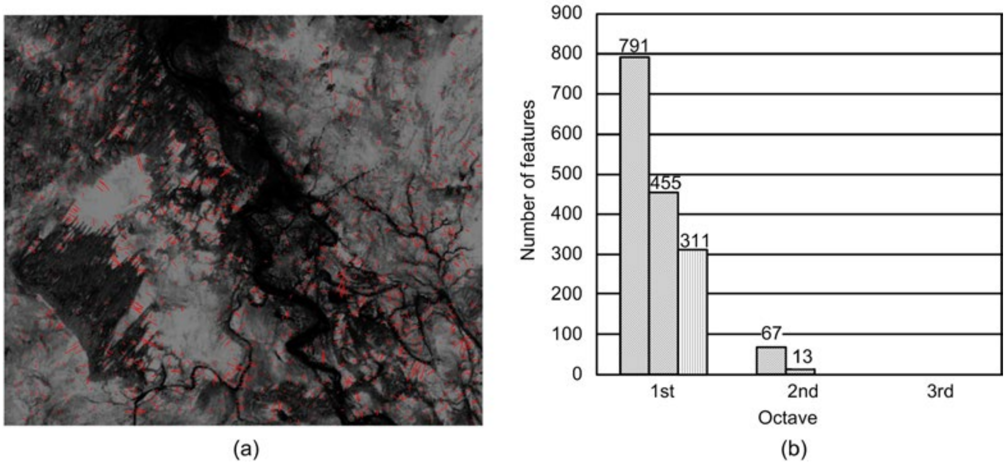

Fig. 3 Image-space (a) and scale-space (b) distributions of features extracted from a pre-processed image used in Fig. 1 with the standard scale-invariant feature transform algorithm (The image size is 445×518)

to be checked and matched against all feature points from L_1 to L_N. Even though the SIFT algorithm is successful in many cases, it usually produces redundant features, especially in the case of satellite images, which causes high computational costs in other stages, such as feature description and mismatch elimination. Also, these features have high dimensionality. In CBIR applications where large numbers of images are considered as a database, these problems will result in a time-consuming registration stage as well as a large storage requirement for saving relevant information.

To overcome such problems, a projection kernel has been created to transform the SIFT local features to global features. LDA is used to map the high-dimensional local feature vector extracted from the training database and to learn the projection kernel. Once the projection kernel is computed, the transformed local features related to each image in the database can be considered to be general features representing that image as a unique class. Our experiments show that in case of satellite images, the first two dimensions of the reduced feature vector include the most information about the image, and using higher dimensions will result in only a time-consuming computational training and classification process. Although LDA is able to mark a feature vector as belonging to a specific class, a huge number of extracted features produced by the SIFT algorithm will make it difficult to clearly identify the boundaries between some of the classes. On the other hand, LDA reduces only the dimensionality of the extracted features. Storing large amounts of data related to these feature points is still a problem.

To solve this problem, the mean global feature vectors for each image are considered representative for that class. Thus, each image will have a representative feature point in a 2D feature space.

In the verification stage, a query image is pre-processed first, and its local features are extracted using the SIFT algorithm. Local feature vectors constructed using the SIFT descriptor are mapped into a 2D general feature space using the transformation kernel calculated during the training process. The mean general feature vector related to a query image is classified using a minimum distance classifier to find its match image in the reference database.

4 Experimental results

To assess the proposed methodology, 24 images from 12 different locations in Iran and the UAE taken from IRS-P6, with LISS III, LISS IV, and AWIFS remote sensors and the Landsat satellite are used in this study. For each individual location, a pair of satellite images taken from different image sensors is selected. The selected image pairs cover a variety of spatial resolutions from 5.80 to 57.00 m and are chosen from both urban and rural areas with different textural patterns and rotations. In each pair, the first image is chosen as the reference database for system training, and the second is used as a query image. The pre-processing stage is applied to both reference and query images.

As mentioned before, the SIFT algorithm is applied for key point detection and feature description for both raw and pre-processed images. In this study, pre-defined parameters and thresholds of this algorithm are equal to the values proposed by Lowe (2004). The input image specification and the results of applying the SIFT algorithm to both groups of images are shown in Table 1. Table 1 also shows a superior SIFT performance in the case of pre-processed satellite images. Comparison of the results shows that the SIFT algorithm with thresholds proposed by Lowe (2004) is not capable of producing an adequate number of features for image registration purposes; however, after pre-processing, a suitable number of features are extracted from these images. For Landsat images, although the SIFT algorithm has an acceptable performance, it can be seen that the pre-processed images have higher match points which increase their registration reliability. In this regard, Fig. 2 illustrates a non-uniform spatial distribution of local SIFT features in both the image space and the scale space extracted from an IRS-P6 image, which is also possible for other types of satellite image sensors. Fig. 3 shows the feature distribution results on a pre-processed version of the first image. Comparison between the two figures shows that the method may produce unrealistic effects in photographs; however, the results prove that it is very useful for remote sensing images.

In this study, LDA is used to transform the local SIFT features extracted from the training dataset into

Table 1 Input image pairs and registration quality of standard scale-invariant feature transform (SIFT) on raw and pre-processed images

Pair number	Satellite	Image size	Pixel size (m/pixel)	Location	Number of features extracted		Number of correct matches obtained	
					Raw	Pre-processed	Raw	Pre-processed
1	IRS-P6 LISS III	858×873	24.00	Iran–Kerman	332	14 330	24	1214
	IRS-P6 AWIFS	354×399	57.00		35	3169		
2	IRS-P6 LISS IV	2152×1920	5.80	Iran–Sistan & Baloochestan	2	13 527	0	75
	IRS-P6 AWIFS	236×212	57.00		89	1107		
3	Landsat ETM+	647×592	14.25	Iran–Hormozgan	4034	6677	0	236
	IRS-P6 LISS IV	1310×1367	5.80		0	5708		
4	IRS-P6 LISS III	489×361	24.00	UAE	873	2899	32	229
	Landsat ETM+	787×585	14.25		5038	8074		
5	IRS-P6 LISS III	858×873	24.00	Iran–Fars	345	5681	57	781
	IRS-P6AWIFS	516×435	57.00		166	2125		
6	Landsat ETM+	492×584	14.25	Iran–Hormozgan	2314	4359	14	127
	IRS-P6 LISS III	312×254	24.00		322	2396		
7	Landsat ETM+	2224×1526	14.25	UAE	10 796	12 574	519	1413
	IRS-P6 AWIFS	527×331	57.00		1046	1216		
8	IRS-P6 AWIFS	422×438	57.00	Iran–Fars	136	1725	43	674
	IRS-P6 LISS III	709×679	24.00		432	4481		
9	IRS-P6 LISS IV	1005×1106	5.80	Iran–Kerman	0	4708	0	237
	Landsat ETM+	524×511	14.25		2134	5126		
10	IRS-P6 LISS III	568×730	24.00	Iran–Kerman	621	5413	75	316
	Landsat ETM+	748×983	14.25		4072	8153		
11	IRS-P6 AWIFS	543×642	57.00	Iran–Sistan & Baloochestan	891	3603	93	563
	IRS-P6 LISS III	1030×1326	24.00		2910	7469		
12	IRS-P6 LISS III	858×873	24.00	Iran–Fars	692	4531	47	394
	IRS-P6 AWIFS	858×873	57.00		192	1021		

global features. Consequently, local feature vectors are mapped onto a new 2D feature space, and the projection kernel of this transformation is learned.

Once the projection kernel is computed, each image in the training database can be considered as a separate class. Using this kernel, local feature vectors related to the second image in each pair, which are considered as a test dataset, can be projected onto a new global feature space. The mean of global features related to each image is considered to be representative of that image. Classifying these representatives using a minimum distance classifier gives a classification rate of 91.67%.

The results of mapping local feature points onto a 2D space and considering representatives for each image are shown in Fig. 4. Here, unfilled markers indicate representative global features of training vectors, while the filled markers relate to the verification images. The results show a successful retrieval rate while significantly reducing required computational time and storage. Also, application of the SIFT key point detector and descriptor creates robustness to changes in image scale, noise, and rotation.

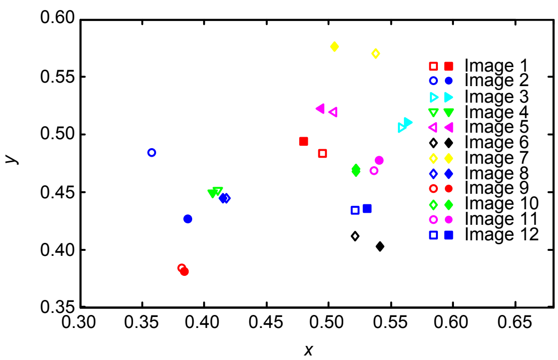

Fig. 4 Linear discriminant analysis mapping of the feature vectors of pre-processed images onto a 2D feature space
Unfilled and filled markers represent the vectors related to the training and verification images, respectively. References to color refer to the online version of this figure

5 Conclusions

Content-based image retrieval applications usually compute features of reference images to identify query images. In this paper, a new algorithm for multisource content-based satellite image retrieval inspired by SIFT (Lowe, 2004) has been introduced. This algorithm has been applicable to various remote sensing images with different sensors, resolutions, rotations, and scales. The proposed method consists of pre-processing and feature transformation stages. In the pre-processing stage, application of a contrast enhancement technique improves the distribution of extracted local features in both the image space and the scale space, while reducing the illumination effect results in better matched local features among images taken from different sensors. The core of the proposed approach is to transform the local features produced by the SIFT key point detector and descriptor into general-type image features using LDA. By considering each image in the training database as a separate class, a 128-dementional local SIFT feature space was mapped onto a 2D space. LDA not only transforms the form of image features, but also reduces their dimensionality, decreasing the computational time significantly for the registration stage and the storage requirement for database features. These factors are both very important in CBIR applications.

The experimental results from the training and query databases, including a variety of multisource remote sensing images, proved the algorithm's advantages in terms of retrieval rate, computational efficiency, sensitivity to noise, and reliability. The results indicate that the proposed method combines the advantages of both local and global features while greatly reducing their disadvantages.

References

Andre, B., Vercauteren, T., Buchner, A.M., *et al*., 2012. Learning semantic and visual similarity for endomicroscopy video retrieval. *IEEE Trans. Med. Imag.*, **31**(6): 1276-1288. https://doi.org/10.1109/TMI.2012.2188301

Arica, N., Vural, F.T.Y., 2003. BAS: a perceptual shape descriptor based on the beam angle statistics. *Patt. Recogn. Lett.*, **24**(9-10):1627-1639. https://doi.org/10.1016/S0167-8655(03)00002-3

Bakar, S.A., Hitam, M.S., Wan, Y., *et al*., 2013. Content-based image retrieval using SIFT for binary and greyscale images. IEEE Int. Conf. on Signal and Image Applications, p.83-88. https://doi.org/10.1109/ICSIPA.2013.6707982

Bhatia, S.K., Samal, A., Vadlamani, P., 2007. RISE-SIMR: a robust image search engine for satellite image matching and retrieval. Int. Symp. on Visual Computting, p.245-254. https://doi.org/10.1007/978-3-540-76856-2_24

Bhattacharya, D., Bhaskar, K., 2013. Efficient aerial image matching algorithm for autonomous navigation of aerial vehicle. *Int. J. Sci. Eng. Technol.*, **2**(12):1204-1207.

Blanchart, P., Datcu, M., 2010. A semi-supervised algorithm for auto-annotation and unknown structures discovery in satellite image databases. *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, **3**(4):698-717. https://doi.org/10.1109/JSTARS.2010.2058794

Chandrakanth, R., Saibaba, J., Varadan, G., *et al*., 2014. A novel image fusion system for multisensor and multiband remote sensing data. *IETE J. Res.*, **60**(2):168-182. https://doi.org/10.1080/03772063.2014.914697

Chu, L., Jiang, S., Wang, S., *et al*., 1996. Robust spatial consistency graph model for partial duplicate image retrieval. *IEEE Trans. Multim.*, **15**(8):1982-1996. https://doi.org/10.1109/TMM.2013.2270455

Dagher, I., Sallak, N., Hazim, H., 2014. Face recognition using the most representative SIFT images. *Int. J. Signal Process. Imag. Process. Patt. Recogn.*, **7**(1):225-236. https://doi.org/10.14257/ijsip.2014.7.1.21

Eugenio, F., Marqués, F., 2003. Automatic satellite image georeferencing using a contour-matching approach. *IEEE Trans. Geosci. Remote Sens.*, **41**(12):2869-2879. https://doi.org/10.1109/TGRS.2003.817226

Goswami, D., Bhatia, S.K., Samal, A., 2007. RISE: a robust image search engine. IEEE Int. Conf. on Electron/Information Technology, p.354-359. https://doi.org/10.1109/EIT.2006.252181

Irtazaa, A., Jaffarb, M.A., Aleisab, E., 2013. Correlated networks for content based image retrieval. *Int. J. Comput. Intell. Syst.*, **6**(6):1189-1205. https://doi.org/10.1080/18756891.2013.823005

Kupfer, B., Netanyahu, N.S., Shimshoni, I., 2015. An efficient SIFT-based mode-seeking algorithm for sub-pixel registration of remotely sensed images. *IEEE Geosci. Remote Sens. Lett.*, **12**(2):379-383. https://doi.org/10.1109/LGRS.2014.2343471

Li, H., Wang, X., Tang, J., *et al*., 2013. Combining global and local matching of multiple features for precise item image retrieval. *Multim. Syst.*, **19**(1):37-49. https://doi.org/10.1007/s00530-012-0265-1

Li, J., Narayanan, R.M., 2004. Integrated spectral and spatial information mining in remote sensing imagery. *IEEE Trans. Geosci. Remote Sens.*, **42**(3):673-685. https://doi.org/10.1109/TGRS.2004.824221

Li, Z., Park, U., Jain, A.K., 2011. A discriminative model for age invariant face recognition. *IEEE Trans. Inform. Forens. Secur.*, **6**(3):1028-1037. https://doi.org/10.1109/TIFS.2011.2156787

Liu, G., Sun, X., Fu, K., *et al*., 2013. Aircraft recognition in high-resolution satellite images using coarse-to-fine shape prior. *IEEE Geosci. Remote. Sens. Lett.*, **10**(3): 573-577. https://doi.org/10.1109/LGRS.2012.2214022

Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.*, **60**(2):91-110. https://doi.org/10.1023/B:VISI.0000029664.99615.94

Maloo, A., Thepade, S.D., Kekre, H., 2010. CBIR feature vector dimension reduction with eigenvectors of covariance matrix using row, column and diagonal mean sequences. *Int. J. Comput. Appl.*, **3**(12):39-46. https://doi.org/10.5120/786-1114

Manuel, I., Tavares, R.S., Natal, J., 2013. Computational Vision and Medical Image Processing IV. CRC Press, London, UK.

Meena, K., Suruliandi, A., Reena, R.R., 2014. Face recognition based on local derivate ternary pattern. *IETE J. Res.*, **60**(1):20-32. https://doi.org/10.1080/03772063.2014.890811

Mikolajczyk, K., Schmid, C., 2005. A performance evaluation of local descriptors. *IEEE Trans. Patt. Anal. Mach. Intell.*, **27**(10):1615-1630. https://doi.org/10.1109/TPAMI.2005.188

Nileshsingh, V., Koli, N., 2010. An overview of feature extraction methods in CBIR. *Int. J. Comput. Sci. Appl.*, **3**(2):71-76.

Okade, M., Biswas, P.K., 2014. Improving video stabilization using multi-resolution MSER features. *IETE J. Res.*, **60**(5):373-380. https://doi.org/10.1080/03772063.2014.962627

Sedaghat, A., Mokhtarzade, M., Ebadi, H., 2011. Uniform robust scale-invariant feature matching for optical remote sensing images. *IEEE Trans. Geosci. Remote Sens.*, **49**(11):4516-4527. https://doi.org/10.1109/TGRS.2011.2144607

Sedaghat, A., Mokhtarzade, M., Ebadi, H., 2012. Image matching of satellite data based on quadrilateral control networks. *Photogr. Rec.*, **27**(140):423-442. https://doi.org/10.1111/j.1477-9730.2012.00699.x

Singaravelan, S., Murugan, D., 2013. Combined global–local specialized feature descriptor for content based image retrieval under noisy query. Int. Conf. on Advanced Computing and Communication Systems, p.1-6. https://doi.org/10.1109/ICACCS.2013.6938716

Stehling, R.O., Nascimento, M.A., Falcão, A.X., 2002. A compact and efficient image retrieval approach based on border/interior pixel classification. Proc. 11th Int. Conf. on Information and Knowledge Management, p.102-109. https://doi.org/10.1145/584792.584812

Torres, R.S., Picado, E.M., Falcao, A.X., *et al.*, 2003. Effective image retrieval by shape saliences. 16th Brazilian Symp. on Computer Graphics and Image Processing, p.167-174. https://doi.org/10.1109/SIBGRA.2003.1241005

Valenzuela, R.E.G., Schwartz, W.R., Pedrini, H., 2012. Dimensionality reduction through PCA over SIFT and SURF descriptors. IEEE 11th Int. Conf. on Cybernetic Intelligent Systems, p.58-63. https://doi.org/10.1109/CIS.2013.6782160

Veganzones, M.A., Maldonado, J.O., Grana, M., 2008. On content-based image retrieval for hyper spectral remote sensing images. *In*: Graña, M., Duro, R.J. (Eds.), Computational Intelligence for Remote Sensing. Springer-Verlag, Berlin, Germany, p.125-144. https://doi.org/10.1007/978-3-540-79353-3_5

Wen, G.J., Lv, J.J., Yu, W., *et al.*, 2008. A high-performance feature-matching method for image registration by combining spatial and similarity information. *IEEE Trans. Geosci. Remote Sens.*, **46**(4):1266-1277. https://doi.org/10.1109/TGRS.2007.912443

Wu, Y.Y., Wu, Y.Q., 2009. Shape based image retrieval using combining global and local shape features. 2nd Int. Congress on Image and Signal Processing, p.12-20. https://doi.org/10.1109/CISP.2009.5304693

Yu, I., Zhang, D., Holden, E.J., 2008. A fast and fully automatic registration approach based on point features for multi-source remote-sensing images. *Comput. Geosci.*, **34**(7):838-848. https://doi.org/10.1016/j.cageo.2007.10.005