Korean Journal of Remote Sensing, Vol.26, No.3, 2010, pp.325~333

The Comparison of the SIFT Image Descriptor by Contrast Enhancement Algorithms with Various Types of High-resolution Satellite Imagery

Jaewan Choi*, Daesung Kim*, Yongmin Kim*, Dongyeob Han, and Yongil Kim*†**

*Department of Civil & Environmental Engineering, Seoul National University

**Department of Civil & Environmental Engineering, Chonnam National University

Abstract : Image registration involves overlapping images of an identical region and assigning the data into one coordinate system. Image registration has proved important in remote sensing, enabling registered satellite imagery to be used in various applications such as image fusion, change detection and the generation of digital maps. The image descriptor, which extracts matching points from each image, is necessary for automatic registration of remotely sensed data. Using contrast enhancement algorithms such as histogram equalization and image stretching, the normalized data are applied to the image descriptor. Drawing on the different spectral characteristics of high resolution satellite imagery based on sensor type and acquisition date, the applied normalization method can be used to change the results of matching interest point descriptors. In this paper, the matching points by scale invariant feature transformation (SIFT) are extracted using various contrast enhancement algorithms and injection of Gaussian noise. The results of the extracted matching points are compared with the number of correct matching points and matching rates for each point.

Key Words : contrast enhancement, image registration, image descriptor, matching point, normalized data, SIFT.

1. Introduction

Mapping image data from a broad area using high-resolution satellite imagery requires the use of an effective geo-referencing algorithm to eliminate geometric distortion and assign the same coordinate to each image. The use of automatic image registration, which overlaps images of identical regions and assigns same coordinate to the images, has been studied in various fields such as remote sensing, computer vision and image processing (Mikolajczyk and Schmid, 2005). The automatic image registration method has been used at bothe global and local levels (Zitová and Flusser, 2003).

The global method measures global similarities between the reference and target image. The correlation-based method, which uses template matching and fast Fourier transformation (FFT), is an

Received June 13, 2010; Revised June 18, 2010; Accepted June 20, 2010.

† Corresponding Author: Yongil Kim (yik@snu.ac.kr)

example of a global registration algorithm (Reddy and Chatterji, 1996). Although global methods are an effective means for image shift or translation between the reference and target image, they are highly sensitive to geometric distortion and the scale of the image.

To overcome problems associated with the global method, many researchers have proposed the use of local methods, which are based on features extracted from such as point, line and area (Zitová and Flusser, 2003; Habib, 2005; Xiong and Zhang, 2009). Interest features by point have advantages over line and area features in terms of registration correction and robustness in image scale. Point features for image registration are extracted by various corner detectors such as Moravec, SUSAN and Harris; registration among images is based on the relationship between peculiarities in the corresponding point features. (Geverekci and Gunturk, 2009). The SIFT method proposed by Lowe (2004) is one of the representative image descriptors for automatic image registration. The features of SIFT invariant to image scale and rotation yield robust matching across a substantial range of affine distortion, in addition to changes in 3-D viewpoint, addition of noise, and changes in illumination (Mikolajczyk and Schmid, 2004). On the other hand, the result of matching points by SIFT can be changed by the applied normalization method because high resolution satellite imagery have different spectral characteristics according to the sensor type and acquisition date. In this paper, the matching points by SIFT are extracted according to normalized data by using various contrast enhancement algorithm and injection of gaussian noise. And then, we compared the result of extracted matching points based on number of correct matching point and matching rate by each point.

2. Contrast enhancement algorithm

In image normalization, the pixel values of the image are effectively adjusted via transformation of the histogram or contrast in the radiometric resolution. In this paper, in applying the image normalization method, we selected efficient methods such as linear stretching, 2% linear stretching, illumination robust feature extraction transform (IRFET) and dynamic histogram equalization (DHE). The images used in the test are assumed to be transformed to range[0-255].

1) Linear stretching

Linear stretching is a commonly used method in image normalization. It transforms the dynamic range used in the input image to a full range of radiometric resolution. When $DN(x, y)$ represents the pixel value of the location (x, y) in the input image, the pixel transformation $DN'(x, y)$ is defined as

$$DN'_{(x,y)} = \frac{DN'_{(x,y)} - DN_{\min}}{DN_{\max} - DN_{\min}} \times 255 \qquad (1)$$

where, $DN_{\max}$ and $DN_{\min}$ represent the maximum and minimum value in the input image.

2) 2% linear stretching

The 2% linear stretching method differs from the linear stretching method in that pixel values that are 2% apart from both end points in the cumulative histogram are defined as maximum and minimum values in the input image. This method is primarily used to visually improve the image in software for satellite image analysis like ENVI software (ENVI User's Guide, 2003).

3) Illumination robust feature extraction transform(IRFET)

IRFET (Geverekci and Gunturk, 2009) involves the extraction of interest points and combats effects

from a change in the light source through gamma correction. In IRFET, center point c is defined as Eq. (2). The latest image extracting interest points is made by integrating c like Eq. (3).

$$DN'_{(x,y;c)} = \frac{1}{1 + e^{-\gamma}(\mathrm{DN}_{(x,y)} - c)} \quad (2)$$

$$DN'_{(x,y)} = \int_0^1 DN'_{(x,y;c)} dc \quad (3)$$

4) Histogram equalization

The histogram equalization method improves the contrast of the input image by reassigning pixels that are concentrated in a specific range to a total range of radiometric resolution. Therefore, the output image shows the equalized histogram. This method has the advantage of being easily and simply applied. When the cumulated distribution function (CDF) is defined as Eq. (4), histogram equalization is defined as Eq. (5).

$$CDF(X_n) = \sum_{i=0}^{k} p(X_n) \quad (4)$$

$$Y_n = f(x) = \frac{L-1}{N} CDF(X_n) \quad (5)$$

where, $p(X_n)$ means the PDF(Probability density function) of X_n, L is a radiometric resolution of input image and N means the pixel number of input image.

5) Dynamic histogram equalization(DHE)

DHE (Wadud *et al.*, 2007) preserves characteristics of the input histogram and reassigns pixel values that have relatively absolute frequency compared with other pixel values. It involves partitioning the input histogram into various histograms so that there is not dominant component. The range of the sub-histogram is then rearranged according to its cumulated frequency. Finally, each sub-histogram is subjected to histogram equalization. The progress involved in setting up the range of sub-histogram is defined as

$$factor_i = span_i * (\log CF_i)^x \quad (6)$$

$$range_i = \frac{factor_i}{\sum_{k=1}^{n} factor_k} \quad (7)$$

where, *span* represents the range of the sub-histogram and CF represents the sum of the frequencies in sub-histogram. x represents optional value to give CF a weight.

3. SIFT

SIFT is considered one of the most powerful image descriptor due to its invariance to image scale, rotation and illumination. There are four stages in the SIFT algorithm (Lowe, 2004).

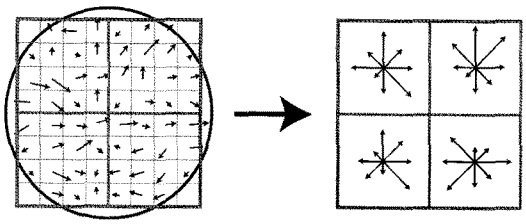
a. Scale-space extrema detection: The first stage of computation searches over all scales and image locations. It is implemented efficiently using a difference-of-Gaussian (DoG) function to identify potential interest points that are invariant to scale and orientation.

b. Keypoint localization: A detailed model about each candidate points is fit to determine location and scale.

c. Orientation assignment: One or more orientations are assigned to each keypoint location based on local image gradient directions. All potential interest points are transformed relative to the assigned orientation, scale, and location to provide invariance to these transformations.

d. Keypoint descriptor: The local image gradients are measured at the selected scale in the region around each keypoint. These are transformed into a representation to minimize the local shape distortion and changes in illumination.

The point feature by SIFT is composed of 128 dimensional vectors, according to the scale and rotation of each feature. Fig. 1 shows the SIFT image descriptor. In specific locations of keypoints, the



(a) image gradient (b) SIFT descriptor

Fig. 1. SIFT image descriptor.

gradient histogram on the normalized area is calculated such as in Fig. 1 (a); the gradient histogram is the divided into an 8-directional element. Finally, 128 dimensional vectors ($4 \times 4 \times 8$) such as Fig. 1(b), are generated as SIFT image descriptor vectors (For efficiency, the SIFT descriptor of Fig. 1(b) is described as a dimensional vector). In the process of automatic image registration, the Euclidian distance between the extracted point features in the reference and target image is calculated. The potential matching point is supposed to the point set, which has a nearest neighbor Euclidian distance at the point pair between the reference and target image. To minimize matching errors by its nearest neighbor, the matching point is obtained by comparing the distance of the closest neighbor with that of its second-closest neighbor (Lowe, 2004). Previous work on automatic image registration has shown that re-organization of extracted matching points using nearest neighbor Euclidean distance is necessary to increase the matching rate by post-processing such as RANdom SAmple Consensus (RANSAC) (Fischler and Bolles, 1981). As our experiments focus solely on the analysis and accuracy of initial matching points in SIFT, we excluded post-processing to increase the true matching rate.

4. Experiments

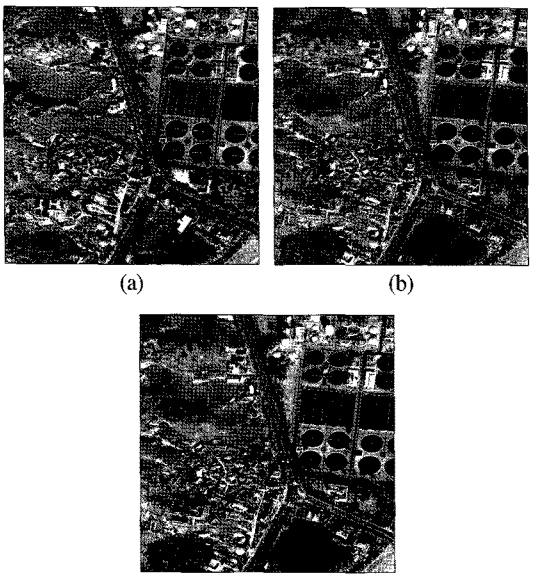
QuickBird images acquired on December 31, 2002, and October 31, 2004, and a KOMPSAT-2 image acquired on October 7, 2007, were used to determine the performance of the SIFT descriptor using the contrast enhancement algorithm. To evaluate the influence of the spatial resolution and spectral characteristics of satellite image, as well as noise errors, we organized the experimental and estimation methods.

1) Experimental methods

Five contrast enhancement algorithms (linear stretching, 2% linear stretching, IRFET, histogram equalization, DHE) were applied to the selected reference and target data. The coordinates of the target data are registered by using the reference data. The reference and target images were transformed to 8-bit dynamic range by contrast enhancement. The SIFT descriptor was then applied to the enhanced reference and target images.

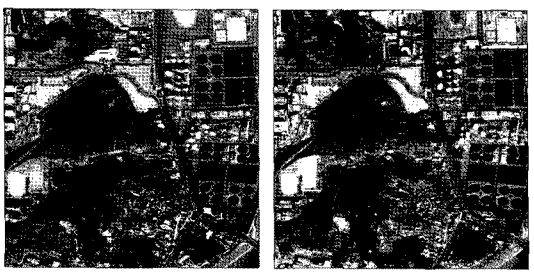
To estimate the influence of various conditions on the image, we proceeded as follows: We added noise error to the QuickBird original panchromatic image with 0.6 m spatial resolution. The reference data were acquired on October 30, 2004, and target data with noise were acquired on December 31, 2002 (Fig. 2 (a) - (b)). Its purpose is to analyze the matching result of SIFT descriptor according to data noise of satellite imagery and examine the sensitivity of contrast enhancement algorithms to noise. Our aim was to determine the effects of noise on the extraction of matching points. The results of the original QuickBird panchromatic image excluding the addition of noise are presented in Fig. 2 (a) and (c). Our second experiment focused on the characteristics of satellite imagery such as spectral resolution, nadir angle and sensor types. The data are presented in Fig. 3, 4 and 5.

a. QuickBird synthetic image with the same spatial resolution (2.4m). The location and dates of



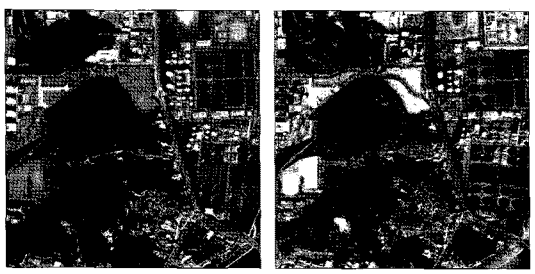
(a) (b) (c)

Fig. 2. Experimental datasets for influence of noise: (a) reference data, (b) target data with noise injection, (c) target data without noise.



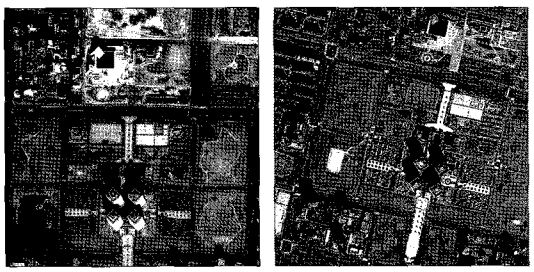
(a) (b)

Fig. 3. Experimental datasets for the characteristics of satellite imagery: (a) the reference data (b) target data.



(a) (b)

Fig. 4. Experimental datasets for the characteristics of satellite imagery: (a) the reference data of green band (b) target data of red band.



(a) (b)

Fig. 5. Experimental datasets for the characteristics of satellite imagery: (a) the reference data (IKONOS) (b) target data (KOMPSAT-2).

reference and target data are equal to those in the panchromatic image. The synthetic image was generated using the average of multispectral bands with 2.4m spatial resolution (Fig. 3).

b. QuickBird imagery with different spectral characteristics. The reference data were acquired on October 30, 2004, at 0.52-0.60μm wavelength, and target data were acquired on December 31, 2002, at 0.63-0.69μm wavelength (Fig. 4).

c. Panchromatic imagery with different sensor types. The reference data for the IKONOS sensor were acquired on November 19, 2001, and the target data for the KOMPSAT-2 sensor were acquired on October 5, 2007 (Fig 5).

2) Estimating the matching rate

We extracted the interest points within the reference and target dataset by SIFT. In the process of SIFT, the interest point, in which the Euclidean distance ratio is greater than 0.8, was rejected among all matches. The registration process used affine transformation in which the matching points consist of a control point. The positional error of registration was calculated based on the geometric characteristics of the image. When relief displacement was virtually non-existent or translation was only considered between the reference and target data such as in

detailed in Fig. 2, 3 and 4, correct matching points of the target data were equal to the coordinated of the reference data. Therefore, we calculated the matching rate by counting the number of matching points that had an identical location among all matching points. If relief displacement occurred due to differences in the nadir angle such as is shown if Fig. 5, we manually checked the true matching point and then computed the matching rate.

5. Experiment results

Table 1 shows the results of extracted and matched interest points by SIFT in automatic image registration irrespective of whether noise data were added to the reference and target data. When noise data were added to the target image, the matching rate and extracted matching points decreased generally. However, the matching points by IREFT show a similar total number of 172 and 160 after the addition of noise. IRFET may have included matching points for various light conditions; therefore, some of the effects of noise are eliminated using the IRFET method. Nevertheless, the general matching performance in the target data following the injection of noise showed a tendency to decrease.

Table 2 shows the results of matching points by SIFT when various contrast enhancement algorithms were applied according to image conditions. In the case of QuickBird synthetic imagery with same spatial resolution (2.4m), the highest correct matching points number of 313 (62.62 matching rate) was derived from histogram equalization. Linear stretching yielded the least number of correct matching points (4 points). DHE-based SIFT showed more correct matching points than 2% linear stretching; however, 2% linear stretching yielded the highest correct matching rate. IRFET yielded a lower matching rate of 54.62 percentage than the other contrast enhancement algorithms, although it had the most extracted point. The IRFET data are due to the method excessively extracting interest points that have similar characteristics, resulting in a low

Table 1. The result of matching point extraction of contrast enhanced image according to noise data

		panchromatic image (0.7m)		panchromatic image with noise	
		reference	target	reference	target (with noise)
linear stretch	extracted points	1468	2205	1468	2484
	matching points	120		53	
	correct matching point (rate)	35 (29.17%)		21 (39.51%)	
2% linear stretch	extracted points	23892	19215	23892	10626
	matching points	498		387	
	correct matching point (rate)	197 (39.6%)		118 (30.57%)	
IRFET	extracted points	32035	33684	32035	26239
	matching points	928		713	
	correct matching point (rate)	172 (18.53%)		160 (22.44%)	
histogram equalization	extracted points	25270	24424	25270	13052
	matching points	574		395	
	correct matching point (rate)	225 (39.2%)		129 (32.74%)	
DHE	extracted points	24492	23004	24492	12953
	matching points	512		342	
	correct matching point (rate)	191 (37.3%)		101 (29.53%)	

Table 2. The result of matching point extraction of contrast enhanced image according to satellite imagery condition

		synthetic panchromatic image (2.4m)		multispectral image		IKONOS-KOMPSAT	
		reference	target	reference (green)	target (red)	reference (IKONOS)	target (KOMPSAT)
linear stretch	extracted points	1974	134	1200	132	10638	1402
	matching points	8		10		168	
	correct matching point (rate)	4 (50%)		0 (0%)		5 (2.98%)	
2% linear stretch	extracted points	5998	5699	4743	5428	29055	11189
	matching points	281		157		331	
	correct matching point (rate)	176 (62.63%)		57 (36.31%)		31 (9.37%)	
IRFET	extracted points	10715	6634	8654	6224	20766	51689
	matching points	249		199		344	
	correct matching point (rate)	135 (54.62%)		74 (37.18%)		34 (9.88%)	
histogram equalization	extracted points	6284	6093	5850	5791	30031	13261
	matching points	313		162		311	
	correct matching point (rate)	196 (62.62%)		57 (35.18%)		42 (13.50%)	
DHE	extracted points	6274	6076	5678	5749	28792	11456
	matching points	304		167		290	
	correct matching point (rate)	182 (59.87%)		61 (36.52%)		43 (14.83%)	

matching rate.

When the reference and target data had different spectral characteristics, linear stretching failed to extract correct matching points. Relative to other contrast enhancement algorithms, the images subjected to linear stretching show a narrower dynamic range of pixels. Therefore, matching points obtained by linear stretching do not reflect relative spectral dissimilarities between the reference and target data. The IRFET method yielded the highest matching rate(37.18%) and number of correct matching points (74 points). In common with finding resulting from the addition of noise, IRFET appeared less sensitive to matching point extraction than other contrast enhancement algorithms. Histogram equalization yielded a lower matching rate than the original panchromatic data. Fig. 6 presents the results of automatic registration between red and green band multispectral data by histogram equalization.

Overall, matching points on areas of vegetation showed more matching error than built-up areas or complex areas (Fig. 4). Histogram-based algorithm such as histogram equalization and DHE, therefore, appear to be more sensitive to spectral differences than geometric or spatial differences.

To estimate the influence of different geometry data, image registration produced by IKONOS and KOMPST-2 image was assessed. As shown in Table 2, histogram equalization yielded a high matching rate (13.50%) and high number of correct matching points (42 points). In addition, the histogram-based algorithm showed a higher matching rate and more correct matching points than either the linear stretching-based algorithm or IRFET. However, linear stretching-based methods may make it difficult to extract the correct matching point efficiently.

Thus, histogram-based contrast enhancement algorithms are more capable of overcoming spatial and geometric distortion than other contrast enhancement methods.

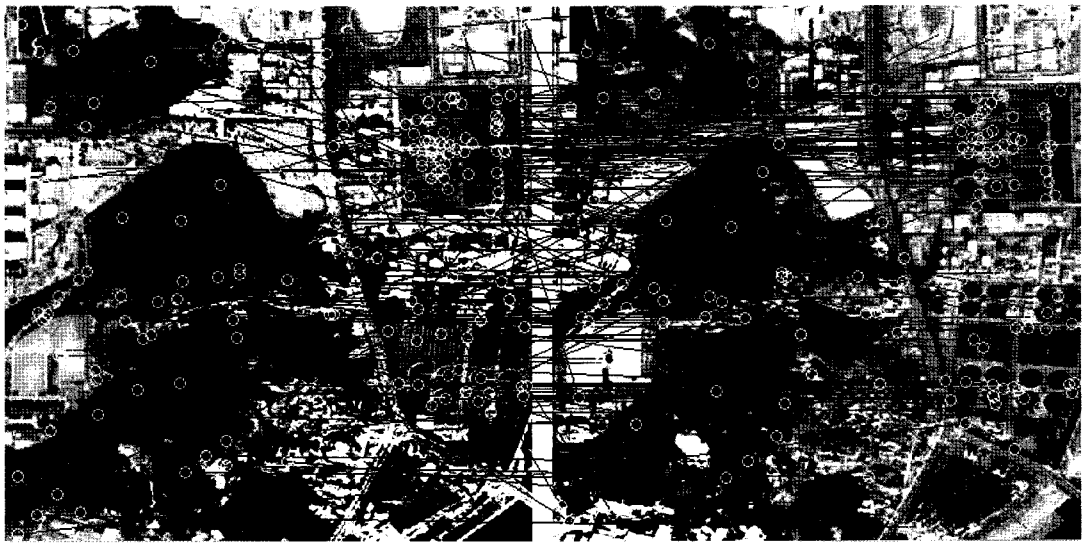


Fig 6. The example of image registration between red and green band of QuickBird imagery using SIFT based on histogram equalization method.

6. Conclusion

Image registration is important in applications such as image fusion, change detection and the generation of digital maps. Our analysis of the performance of the SIFT descriptor using various contrast enhancement algorithms suggests that the histogram-based contrast enhancement algorithm is more capable of combating spatial and geometric distortion compared with the other methods. We also concluded that IRFET is less sensitive to spectral dissimilarities. In addition, correct matching points and matching rate increase as the spectral information between reference and target data is similar to each other, generally. Further experimentation with the methods discussed here should prove worthwhile in the area of satellite imagery and analysis the effect of automatic image registration between pan-sharpened satellite image or panchromatic and each multispectral or hyperspectral imagery. Further research is needed on the development of new contrast enhancement algorithm, which is more insensitive to both spectral and spatial resolution.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MEST) (No. 20090085932).

References

Mikolajczyk, K. and C. Schmid, 2005. A performance evaluation of local descriptors, *IEEE Transcations on Pattern Analysis and Machine Intelligence*, 27(10): 1615-1630.

Zitová, B. and J. Flusser, 2003. Image registration methods: a survey, *Image and Vision Computing*, 21(11): 977-1000.

Lowe, D. G, 2004. Distinctive image features from scale-invariant keypoints, *International Journal of Computer Vision*, 60(2): 91-110.

Reddy, B. S. and B. N. Chatterji, 1996. An FFT-Based Technique for Translation, Rotation, and Scale-Invariant Image Registration, *IEEE Transactions on Image Processing*, 5(8):

1266-1271.

Habib, A. G. Mwafag, M. Michel, and R. Al-Ruzouq, 2005. Photogrammetric and LIDAR data registration using linear features, *Photogrammetric Engineering & Remote Sensing*, 71(6): 699-707.

Xiong, Z. and Y. Zhang, 2009. A novel interest-point-matching algorithm for high-resolution satellite images, *IEEE Transactions on Geoscience and Remote Sensing*, 47(12): 4189-4200.

Geverekci, M. and B. K. Gunturk, 2009. Illumination robust interest point detection, *Computer Vision and Image Understanding*, 113: 565-571.

Mikolajczyk, K. and C. Schmid, 2004. Scale & Affine Invariant Interest Point Detectors, *International Journal of Computer Vision*, 60(1): 63-86.

ENVI User's Guide, Version 4, 2003. Research System, Inc.

Wadud, M. A.-A., Md. H. Kabir, M. A. A. Dewan, O. Chae, 2007. A Dynamic Histogram Equalization for Image Contrast Enhancement, *IEEE Transactions on Consumer Electronics*, 53(2): 593-600.

Fischler, M. A., and R. C. Bolles, 1981. Random Sample consensus: a paradigm ofr model fitting with applications to image analysis and automated cartography, *Commun. ACM*, 24(6): 381-395.