한국측량학회지, 제 29권 제 5호, 2011. 10 : pp. 449~457

A Performance Analysis of the SIFT Matching on Simulated Geospatial Image Differences

공간 영상 처리를 위한 SIFT 매칭 기법의 성능 분석

Oh, Jaehong[1)] · Lee, Hyoseong[2)]
오재홍 · 이효성

Abstract

As automated image processing techniques have been required in multi-temporal/multi-sensor geospatial image applications, use of automated but highly invariant image matching technique has been a critical ingredient. Note that there is high possibility of geometric and spectral differences between multi-temporal/multi-sensor geospatial images due to differences in sensor, acquisition geometry, season, and weather, etc. Among many image matching techniques, the SIFT (Scale Invariant Feature Transform) is a popular method since it has been recognized to be very robust to diverse imaging conditions. Therefore, the SIFT has high potential for the geospatial image processing. This paper presents a performance test results of the SIFT on geospatial imagery by simulating various image differences such as shear, scale, rotation, intensity, noise, and spectral differences. Since a geospatial image application often requires a number of good matching points over the images, the number of matching points was analyzed with its matching positional accuracy. The test results show that the SIFT is highly invariant but could not overcome significant image differences. In addition, it guarantees no outlier-free matching such that it is highly recommended to use outlier removal techniques such as RANSAC (RANdom SAmple Consensus).

Keywords : SIFT, Matching, Matching performance, RANSAC, Geospatial image

1. Introduction

Image matching is to automatically extract conjugate image points from an overlapping image pair such as a stereo image pair or a multi-temporal image pair. Image matching is a critical process for various image applications such as identifying same features from multiple looking images, change detection, feature tracking, and image alignment. In the geospatial image processing, image matching is used for DEM (Digital Elevation Model) generation from stereo images (Maas, 1996; Zhang and Gruen, 2006) and image-to-image matching based automated georegistration (Chen and Arora, 2003; Wong and Clausi, 2007; Cariou and Chehdi, 2008; Oh et al., 2010). Image matching techniques can be classified into the area-based, the feature-based, and the hybrid method (Wolf and Dewitt, 2000). The area-based methods perform the matching by comparing pixel values from each image. The normalized cross correlation matching technique is a well-known area-based matching, and it is the most common method which is adopted in many photogrammetry systems. The least square matching is also one of the area-based matching but it is invariant to intensity change and affine deformations. The feature-based matching techniques are relatively new and have been developing in the computer science field, and it is more complicated and uses feature characteristics such as size and shape. The hybrid methods involve the combination of both approaches.

A matching task typically has two steps, i.e., extraction of distinct points (or features) from one image, and the search for the counterpart in the corresponding image. The critical factor

1) Researcher · U-space Research Team · Electronics and Telecommunications Research Institute (E-mail:ojh@etri.re.kr)
2) Corresponding author · Associate Professor · Civil Engineering · Suncheon National University (E-mail:hslee@sunchon.ac.kr)

that affects the quality of image matching is the similarity between the two images in addition to image quality. In the geospatial image processing, the paired images often have different geometry and different spectral information as they are obtained in different environments i.e. different acquisition times, angles, and different sensors. Therefore, robust matching techniques should be used to account for the large imaging condition difference between the multi-temporal/multi-sensor geospatial image pair.

In the field of computer vision, the SIFT (Scale Invariant Feature Transform, Lowe, 1999) is popularly used because it has been recognized to be very reliable and invariant to imaging condition changes such as scale, rotation, affine distortion and intensity changes. There are also some modifications to SIFT to make it more effective: PCA-SIFT (Ke and Sukthankar, 2004), GLOH (Gradient Location-Orientation Histogram) (Mikolajczyk and Schmid, 2005), CSIFT (Abdel-Hakim and Farag, 2006), SR-SIFT (Yi et al., 2008), SURF (Speeded-Up Robust Features) (Bay et al., 2008) and Robust SIFT (Li et al., 2009), though, they are conceptually similar. Even though the SIFT is popularly used in the computer vision application, characteristics and positional accuracy of the SIFT should be analyzed first for geospatial images. Therefore, in this paper, the SIFT was tested based on simulation for various geospatial image conditions such as shear distortion, scale, rotation, noise, intensity, and spectra difference. Then matching positional accuracy was checked using the known image transform because the positional accuracy of matching is very important in geospatial image processing contrast to generic computer vision applications.

This paper is structured as follows: in section 2, the SIFT algorithm and concept are briefly described and section 3 presents analysis on the SIFT performance and matching positional accuracy for various simulated geospatial image conditions, followed by the summary and conclusion in section 4.

2. SIFT (Scale Invariant Feature Transform)

Scale Invariant Feature Transform (SIFT) matching is designed to extract invariant features from images and to perform matching. For effective image matching, the features should be invariant to scale, rotation, affine distortion and intensity changes. The brief summary of the SIFT feature extraction algorithm is written and paraphrased in the below based on the authors' understanding; refer to (Lowe, 2004) for more detail.

2.1 Scale-space extrema detection

At this stage, keypoints are extracted using difference of Gaussian filters (DoG) at different image scale. Features invariant to image scale change are searched using different image scales as shown in Fig. 1. At the first octave, Gaussian blur shown in Eq.(1) is used to generate a set of scaled images, then DoG images are obtained. By down-sampling the first octave image by a factor of two, the next octave image is obtained and the same procedure is repeated.

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

Where σ is the standard deviation of the Gaussian distribution, $L(x, y, \sigma)$ is the scale space of an image, blurred with σ, $G(x, y, \sigma)$ is the Gaussian at the standard deviation σ, $I(x, y)$ is the input image, and * is the convolution operator.

The difference of the Gaussian-blurred images, generated using σ and $k\sigma$, is simply expressed as Eq.(2).

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma) \tag{2}$$

Where k is a constant multiplicative factor for DoG.

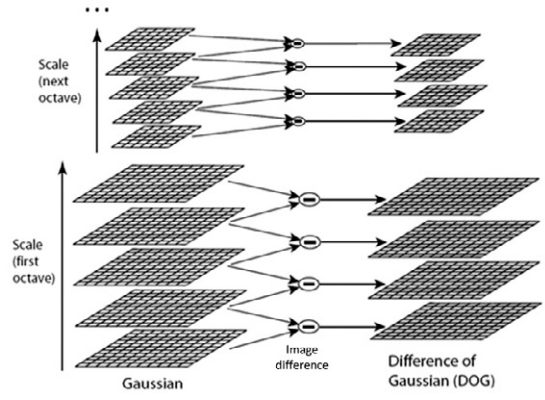


Fig. 1. The Gaussian-blurred (scaled) images at the first octave shown in the left-bottom, and difference of Gaussian (DoG) images are in the right. Using down-sampling by factor of two, the next octave is obtained (Lowe, 2004).

After DoG images across image scales are generated, as shown in the right-bottom of Fig. 1, keypoints are located by comparing each sample point to its eight neighboring pixels at the same scale, and to nine neighbors in the scales above and below (total 26 neighbors), as shown in Fig. 2. In the figure, the point marked with 'X' is a sample point, and it is compared to the neighbors (marked with circles). If the sample point is the local minimum or the local maximum of the neighborhood, it is selected as a keypoint.

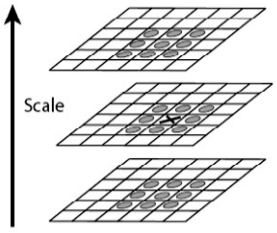


Fig. 2. Maximum and minimum of the DoG images are detected by comparing a pixel (marked with X) to its 26 neighbors in 3x3 regions at the current and adjacent scales (marked with circles) (Lowe, 2004).

2.2 Keypoint localization

The next step is to remove keypoints with low contrast or poorly localized along the edges. At the initial implementation of SIFT, the location and scale of keypoints obtained in the previous step were used as directly obtained at that step. Later, more accurate localization was performed by fitting a 3D quadratic function to the keypoints to determine the interpolated location of the maximum (Lowe, 2004).

The keypoint with low contrast can be removed using the value of $D(x, y, \sigma)$. Lowe (2004) used 0.03 as the threshold for $|D(x, y, \sigma)|$. Removing keypoints poorly localized along the edges can be done using the 2 x 2 Hessian matrix, obtained by taking second derivatives of D as shown in Eq.(3). Then, the keypoints not satisfying Eq. (4) are removed.

$$H = \begin{bmatrix} D_{xx} & D_{xy} \\ D_{xy} & D_{yy} \end{bmatrix} \tag{3}$$

With $trace(H) = \lambda_1 + \lambda_2$, $\det(H) = \lambda_1 \lambda_2$, and $\lambda_1 = r\lambda_2$

Where λ_1, λ_2 are the eigenvalues and r is the ratio between the eigenvalues.

$$\frac{trace(H)^2}{\det(H)} < \frac{(1+r)^2}{r} \tag{4}$$

2.3 Orientation assignment

For the next step, each keypoint needs to be assigned an orientation. Magnitude and orientation can be computed using pixel difference of neighbors, as shown in Eqs.(5) and (6), respectively. To determine the orientation, an orientation histogram, which has 35 bins for 360 deg orientation range, is formed around the keypoint. And then, each neighboring pixel is weighted by the gradient magnitude and a Gaussian window with σ that is 1.5 times the scale of the key points. In the histogram, peaks correspond to dominant orientations. Multiple keypoints are created for the direction to the histogram peak and any others within the 80% of highest peaks.

$$m(x, y) = \sqrt{\left(L(x+1, y) - L(x-1, y)\right)^2 + \left(L(x, y+1) - L(x, y-1)\right)^2} \tag{5}$$

$$\theta(x, y) = \tan^{-1}\left(\frac{L(x, y+1) - L(x, y-1)}{L(x+1, y) - L(x-1, y)}\right) \tag{6}$$

Where m(*x, y)* is a magnitude, θ(*x, y)* is an orientation, and L(*x, y)* is a scale space of an image at the scale of the keypoint.

2.3 Generation of keypoint descriptors

After orientation is assigned to a keypoint, the keypoint descriptor is computed as a set of orientation histogram in 4×4 pixel neighborhoods. Gradient magnitude and orientation of each sample point around the keypoint are computed and weighted by a Gaussian window. Then, the information is accumulated into the orientation histogram of 4×4 subregions. Each histogram has 8 bins and each descriptor has an

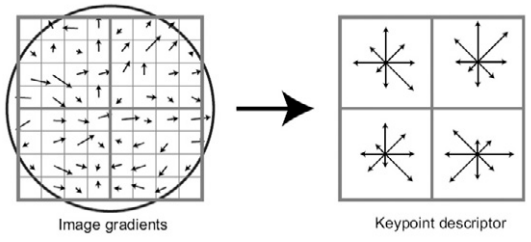


Fig. 3. Generation of keypoint descriptors (Lowe, 2004).

array of 4×4 histograms around the keypoint. This leads to a feature vector with 4×4×8=128 elements. This descriptor is shown in Fig. 3. Note that the descriptor in Fig. 3 has an array of 2×2 histograms around the keypoint. Conventionally the feature vector of 128 elements is normalized and used for SIFT matching.

2.4 SIFT matching

SIFT feature matching is performed by comparing each keypoint descriptor from one image to all the key point descriptors from the counterpart image. In other words, the matching counterpart is determined when the most similar feature vector is found. Similarity can be obtained by computing a dot product of two feature vectors, e.g. if two features are same, their similarity will be one. One concern is that if a feature vector is the most similar, it does not mean it is a correct matching point. Therefore, SIFT uses a threshold to filter out similar but wrong matching points. SIFT utilizes the ratio of the maximum similarity to the second similarity to obtain only matching points with outstanding similarity.

3. SIFT performance analysis

SIFT matching performance is analyzed for simulated image differences such as shear distortion, scale, rotation, noise, intensity, and spectral difference. Fig. 4 depicts the flowchart of the SIFT matching based on simulation. From a test image, three images are subset which geometric and spectral distortion are simulated to. Then the SIFT matching was carried out between the test image and simulated images. For the experiment, three test images, shown in Fig. 5(b), which represent building, residential, and flat areas, are subsets from an aerial hyperspectral image shown in Fig. 5(a), and the simulated image differences are applied to those sub-images shown in Fig. 5(c). The analysis was carried out for the SIFT matching between the simulated images shown in Fig. 5(c) and the original images shown in Fig. 5(a). The geometric image differences between the original and simulated images are modeled using the affine transform Eq.(7), with any combinations of Eqs.(8),(9), and (10).

$$\begin{bmatrix} x' & y' & 1 \end{bmatrix}^T = T \begin{bmatrix} x & y & 1 \end{bmatrix}^T \tag{7}$$

Where T is an affine transformation matrix, x, y are the image coordinates in the original image, and x', y' are the image coordinates of the simulated (distorted) image.

$$Shear: T = \begin{bmatrix} 1 & 0 & 0 \\ \alpha & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \; or \; T = \begin{bmatrix} 1 & \alpha & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{8}$$

Where α is the shear parameter.

$$Scale: T = \begin{bmatrix} s_x & 0 & 0 \\ 0 & s_y & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{9}$$

Where s_x, s_y are the scale parameters in x and y directions, respectively.

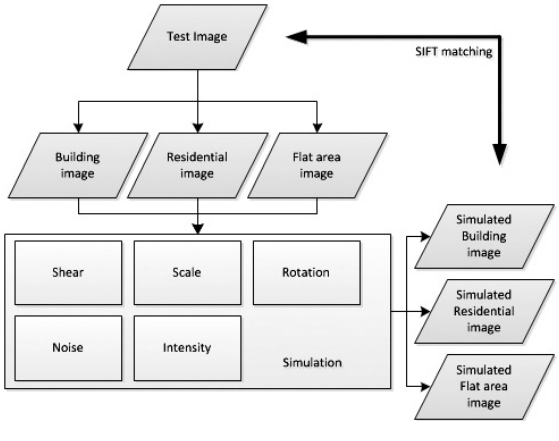


Fig. 4. Simulation and SIFT matching flow chart.

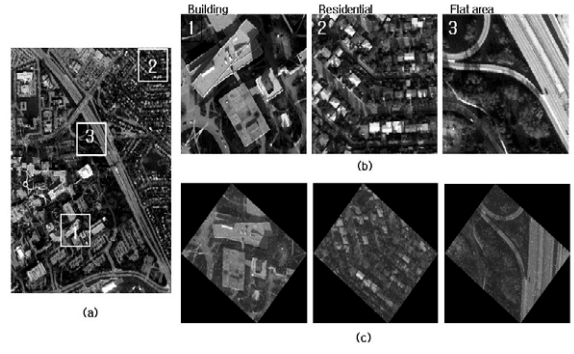


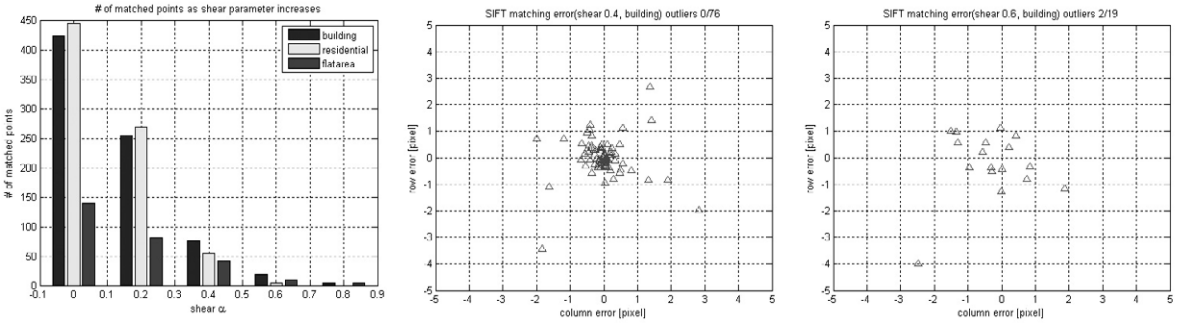
Fig. 5. Test images, CASI-1500 (Courtesy of ITRES Research) (a) full scene (b) subset scenes, each 300•™300 pixels (c) simulated images using scale difference of 0.8, shear of 0.2, rotation by 45 deg and noise at 1%.

$$Rotation: T = \begin{bmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (10)$$

Where θ is the rotation angle.

3.1 Shear distortion

With changing shear parameter α, the number of matching points of SIFT matching is counted and plotted in Fig. 6(a). Note that shear distortion including the affine is common distortion which often appears in the aerial images and satellite images due to camera orientation. The figure shows that as shear distortion α increases, the number of matching points significantly decrease, meaning that SIFT matching does not overcome significant shear distortion. The test images of buildings and residential area show more matching points are obtained than the image of a flat area. Note that the building and residential images contain more distinct features than the flat area image and they produce large number of corner points in the keypoint localization step of SIFT. Next, the image matching positional accuracy was investigated since the correct position of the counterpart points is known from the distortion parameter. The matching errors within the five pixels are plotted in Fig. 6(b)(c) and the image points with errors larger than five pixels are marked as outliers and counted. The number of outliers out of the total number of matching points is also presented at the caption of each figure. Fig. 6 demonstrates that with increasing shear distortion, the overall positional accuracy tends to decrease. Ideally, the matching error should be zero for the no-distortion case, but the test showed that the errors for some matching points are not zero. The likely reason is due to the scale-space approach of SIFT; i.e., different octave requires down-sample image resampling which might affect accurate localization of keypoints. Similarly, zero error is not obtained when an image is matched to a subset of itself. At shear of 0.0~0.4, no outliers were observed; for shear of 0.6, two outlier matching points were observed.

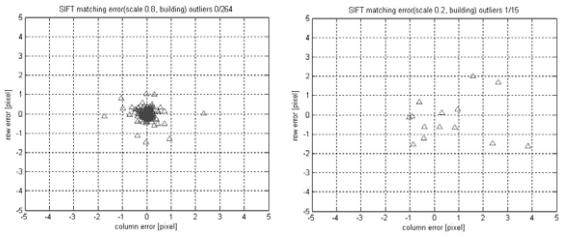


(a) (b) shear 0.4: outliers= 0/76 (c) shear 0.6: outliers= 2/19

Fig. 6. SIFT invariance and positional accuracy as a function of shear distortions

3.2 Scale difference

Geospatial images such as aerial images and satellite images often have different image scales, i.e. resolution, depending on camera focal length and altitude; therefore, image matching should be highly invariant to scale. In the test, the scale ratio of one is for the same scales of both images and smaller scale parameters mean larger scale difference. Obviously, smaller scale parameters (larger scale difference) should produce fewer matching points because the size of the scaled image is reduced, i.e. the total number of image pixels is reduced. Therefore, the effect of image scale difference cannot be investigated based on the number of the matching points. Thus, only the positional accuracy was tested for the scale difference as shown in Fig. 7, the small scale difference tends to insignificantly affect the accuracy while few inaccurate and outlier points are observed in large scale differences such as in the scale ratio of 0.2. Note that the positional accuracy was computed in the reference image scale (full scene). The SIFT matching seems highly invariant to the scale difference.

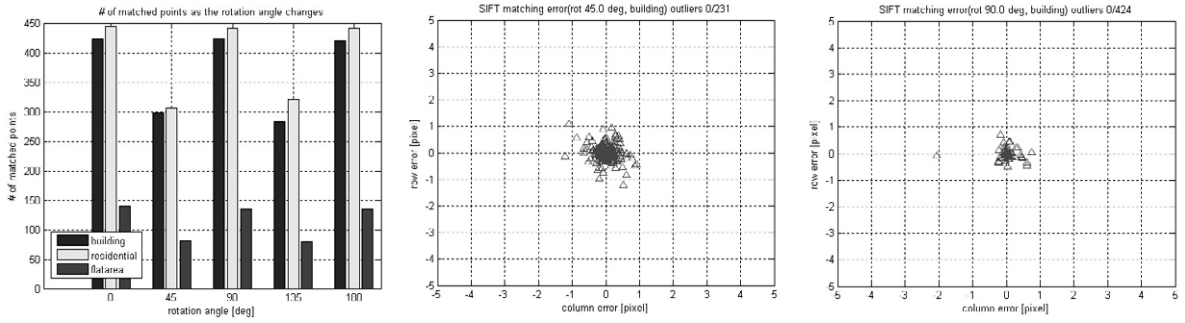


(a) scale 0.8: outliers= 0/264 (b) scale 0.2: outliers= 1/15

Fig. 7. SIFT positional accuracy as a function of scale difference.

3.3 Image rotation

Fig. 8(a) presents the number of matching points as a function of varying image rotation. Note that aerial images often have different rotation (orientation) with respect to the reference image due to the flight direction. In Fig. 8(a), it is observed that the same number of matching points is acquired when rotating the image by 90 and 180 deg. The probable reason that 45 and 135 deg rotations provide different results is that 45 and 135 deg rotations require interpolation of pixel values such as cubic convolution, which often result in slightly different pixel values in the output image than the original image. In contrast, 90 and 180 deg rotations do not require the interpolation. In the matching positional accuracy test shown in Fig. 8(b)(c), there were no matching points showing significantly low accuracy. In other words, SIFT matching has good invariance to image rotations.

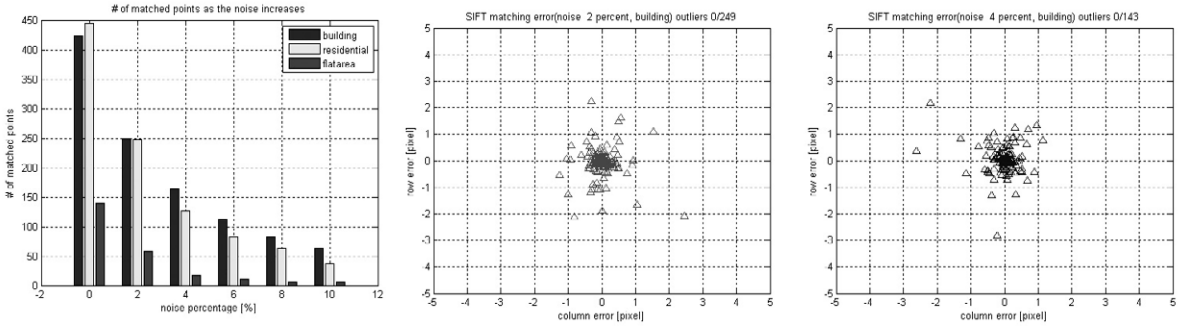


(a) (b) rotation 45 deg: outliers= 0/231 (c) rotation 90 deg: outliers= 0/424

Fig. 8. SIFT invariance and positional accuracy as a function of image rotations.

3.4 Noise

Fig. 9(a) presents the SIFT matching results for 'salt and pepper' image noise, which is also called impulse noise. This unwanted noise corrupts an image by randomly appearing as white and black dots superimposed on the image. It is also introduced to the aerial and satellite images due to a noisy channel, errors during the measurement process and during quantization of the data for digital storage (Chahal and Singh,



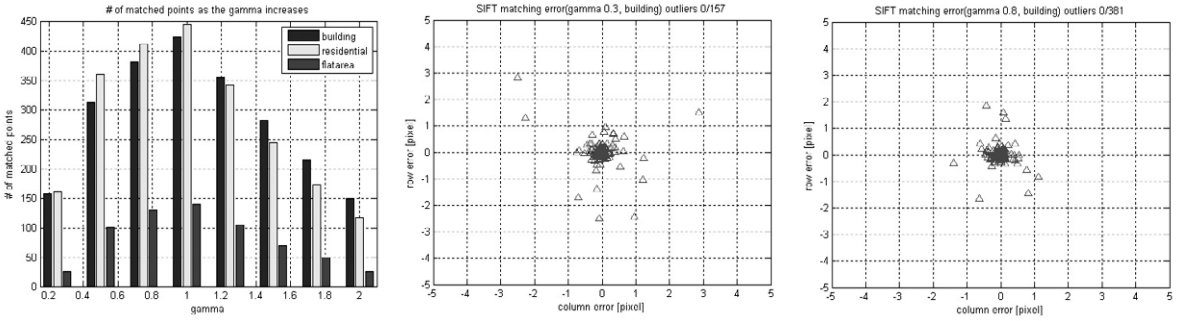
(a) (b) noise: 2 %: outliers= 0/249 (c) noise 4 %: outliers= 0/143

Fig. 9. SIFT invariance and positional accuracy a function of salt & pepper noises.

2010). Therefore, the image matching should be robust to this phenomenon too. In Fig. 9(a), it is observed that the number of matching points significantly decreases for the images with significant noise. Fig. 9(b)(c) show that the matching positional accuracy only slightly degrades, as the noise increases. In other words, the noise does not have much impact on the SIFT matching accuracy.

3.5 Intensity difference

The intensity in aerial images and satellite images can be different depending on the image acquisition time, season, and weather. Therefore, invariance to intensity change is also required for robust image matching. Fig. 10(a) shows the number of matching points from the SIFT matching as a function of intensity differences, which was simulated using the power-law equation (Gonzalez and Woods, 2001). Note that gamma is the coefficient in the power-law equation. Gamma of 1 represents no intensity difference and smaller gamma tends to produce the brighter image from an input image. Fig. 10(a) shows that the number of matching points decreases as the image intensity difference increases. Matching positional accuracy slightly degrades as the intensity difference increases in Fig. 10(b)(c). The test result indicates that SIFT is highly invariant to the intensity differences.



(a) (b) gamma 0.3: outliers= 0/157 (c) gamma 0.8: outliers= 0/381

Fig. 10. SIFT invariance and positional accuracy a function of intensity differences.

3.6 Spectral difference

Current trend in aerial and satellite sensors is to provide multispectral information, not only the high-resolution panchromatic image; for example, WorldView-2 satellite (DigitalGlobe) has eight spectral bands. Also, hyperspectral images such as CASI-1500 by ITRES are gaining more interest. Therefore, image matching across the spectral bands will be required in many applications. Fig. 11 shows the result when SIFT matching is performed between the spectral band 14 (878 nanometers wavelength, near-infrared) and bands 1-18 (385~1030 nanometers wavelength) of the test image. The figure clearly shows that the SIFT matching between visible spectrum and near-infrared spectrum produces only a few matching points. Note that the spectral bands from 1 to 10 are in visible spectrums. In addition, the number of matching point decreases significantly as the spectral gap increases. The test results indicate that SIFT cannot handle the large spectral gap. Therefore, it will be very important to select spectral bands of similar spectral ranges for redundant image matching points.

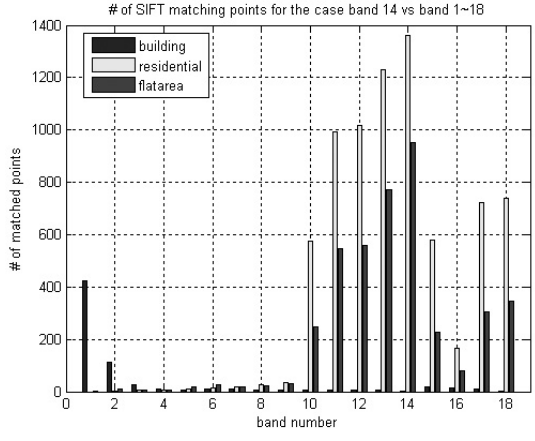
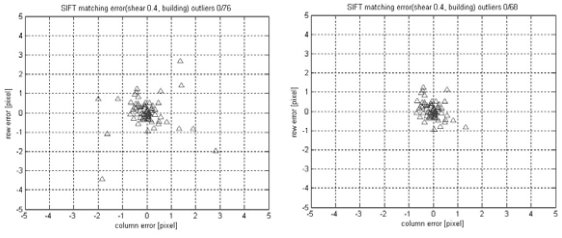


Fig. 11. SIFT invariance test on spectrum difference (band 14 vs. the others).

3.7 Need of RANSAC

In the SIFT performance analysis, it was shown that SIFT matching may produce inaccurate matching points with the positional accuracy lower than a few pixel level, which can be significant in the perspective of mapping. In addition, some outlier matching points were observed as error could directly propagate to the quality of aerial image georeferencing. Therefore, these inaccurate matching points should be removed. RANSAC (Fischler and Bolles, 1981) is a technique to estimate parameters of a model through iterations from a set of observations that contain outliers. Model parameters are estimated from a randomly selected observation subset and then every observation is tested if it fits to the model, and it is added to the consensus set. Through the iteration procedure, a new consensus set is obtained, and a better model is estimated. RANSAC is useful especially when the number of outliers is large, where other robust techniques, such as the least squares residual check or Baarda's data snooping (Baarda, 1968) have practical limitations. RANSAC requires appropriate selection of model to fit data. Depending on the applications, different model should be selected such as line and plane equation, homography, fundamental matrix, and affine model etc.

As a test, affine model-based RANSAC (6 DoF) was applied to attenuate an inaccurate SIFT matching points for the shear distortion as shown in Fig. 6. A number of low-positional-accuracy matching points are observed in Fig. 12(a), but Fig. 12 (b) shows that RANSAC successfully pruned inaccurate matching points.



(a) before outlier removal (b) after outlier removal

Fig. 12. RANSAC test to attenuate SIFT matching errors from shear distortion.

4. Conclusion

As automated image processing techniques have been required in multi-temporal/multi-sensor geospatial image processing such as georegistration, use of automated but highly invariant image matching technique has been a critical ingredient. SIFT is one of the most popular point feature matching methods but its characteristics and matching positional accuracy has not been investigated for geospatial image of diverse imaging conditions. This study carried out analysis of the SIFT matching performance for geospatial imagery as a function of varying image parameters and conditions such as shear, scale, rotation, intensity, noise, and spectral differences. The conclusion derived from the experiment can be summarized as below.

- the SIFT matching can provide acceptable number of matching points under insignificant image geometric distortion.
- the noise (salt and pepper) does not have much impact on the SIFT matching positional accuracy.
- the SIFT matching is highly invariant to scale, rotation and intensity differences.
- the SIFT matching may produce matching points with inaccurate positional matching accuracy, which may not be enough for large-scale mapping.
- the SIFT matching guarantees no outlier-free matching results.
- the SIFT matching requires extensive post-processing to secure satisfactory matching positional accuracy.
- the SIFT matching is not invariant to spectral differences, e.g. matching between visible and infrared bands. Therefore, it is important to select optimal spectral bands when matching is carried out between different types of geospatial imagery

In this study, simulated geometric distortions are limited to 2-dimensional case. For the future study, the effect of relief displacement of ground features on the image matching needs to be investigated and more experiments on real data sets are highly required. In addition, more diverse outlier removal algorithms need to be investigated for their efficiency and

robustness such as MLESAC (Maximum Likelihood Estimation SAmple and Consensus).

References

Abdel-Hakim, A.E., and Farag, A.A. (2006), CSIFT: A SIFT Descriptor with Color Invariant Characteristics, *In: Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, Vol. 2, pp. 1978-1983.

Baarda, W. (1968), *A testing procedure for use in geodetic networks*, Netherlands Geodetic Commission, Publications on Geodesy, New Series, Vol. 2, No. 5, Delft.

Bay, H., Ess, A., Tuytelaars, T., and Gool, L.V. (2008), SURF: Speeded Up Robust Features, *Computer Vision and Image Understanding (CVIU)*, 110(3): 346-359.

Cariou, C. and Chehdi,K. (2008), Automatic Georeferencing of Airborne Pushbroom Scanner Images With Missing Ancillary Data Using Mutual Information, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 46, No. 5, pp. 1290-1300.

Chahal, G. and Singh H. (2010), Robust Statistics based Filter to Remove Salt and Pepper Noise in Digital Images, *International Journal of Information Technology and Knowledge Management*, Vol. 2, No. 2, pp. 601-604.

Chen, H.M., Arora, M.K., and Varshney, P.K. (2003), Mutual Information-Based Image Registration for Remote Sensing Data, *International Journal of Remote Sensing*, Vol. 24, No. 18, pp. 3701-3706.

Fischler, M.A. and Bolles, R.C. (1981), Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography, *Communications of the ACM*. Vol. 24, pp. 381-395.

Gonzalez, R.C. and Woods, R.E. (2001), *Digital image processing*, second edition, Prentice Hall, New Jersey, pp. 80-81.

Ke, Y and Sukthankar, R. (2004), PCA-SIFT: A more distinctive representation for local image descriptors, *In: Proceedings International Conferences on Computer Vision*, Washington DC, 2004, pp. 506-513.

Li, Q., Wang, G., Liu, J., and Chen, S. (2009), Robust Scale-Invariant Feature Matching for Remote Sensing Image Registration, *IEEE Geoscience and Remote Sensing Letters*, Vol. 6, No. 2, pp. 287-291.

Lowe, D.G. (1999), Object recognition from local scale-invariant features. *In: Proceedings International Conferences on Computer Vision*, Corfu, Greece, pp. 1150-1157.

Lowe, D.G. (2004), Distinctive Image Features from Scale-Invariant Keypoints, *International Journal of Computer Vision*, Vol. 60, No. 2, pp. 91-110.

Maas, H.G. (1996), Automatic DEM generation by multi-image feature based matching. *Int. Arch. Photogramm. Remote Sens*. Vol. 31 Part B3, pp. 484-489.

Mikolajczyk, K., and Schmid, C. (2005), A performance evaluation of local descriptors. *IEEE Trans. Pattern Anal. Mach. Intell*., Vol. 27, No. 10, pp. 1615-1630, Oct. 2005.

Oh, J.H., Toth, C.K. and Grejner-Brzezinska, D.A. (2010), Automatic Geo-referencing of Aerial images using High-resolution Stereo Satellite Images, *ASPRS 2010 Annual Conference*, San Diego, CA, April 26-30.

Wolf, P., and Dewitt, A. (2000), *Elements of Photogrammetry with Applications in GIS*, McGraw-Hill.

Wong, A., and Clausi, D.A. (2007), ARRSI: Automatic Registration of Remote-Sensing Images, *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 45, No. 5, pp. 1483-1493.

Yi, Z., Zhiguo, C. and Yang, X. (2008), Multi-spectral remote image registration based on SIFT, *Electron. Lett*., Vol. 44, No. 2, pp. 107-108.

Zhang, L. and Gruen, A., (2006), Multi-image matching for DSM generation from IKONOS imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 60, No. 3, pp. 195-211.

(접수일 2011. 08. 03, 심사일 2011. 09. 15, 심사완료일 2011. 09. 16)