EVALUATION OF MULTIPLE-DOMAIN IMAGERY MATCHING BASED ON DIFFERENT FEATURE SPACES

Hui Ju, PhD Candidate
Charles K. Toth, Senior Research Scientist
Dorota A. Grejner-Brzezinska, Professor
The Center for Mapping
The Ohio State University
Columbus, OH 43210
ju.32@osu.edu
toth@cfm.ohio-state.edu
dbrzezinska@osu.edu

ABSTRACT
This paper is focused on analyzing the performance of various matching methods that could be applied to multiple-domain imagery matching for photogrammetric and remote sensing purpose. More precisely, the matching between the LiDAR intensity imagery and optical satellite/airborne image domains is of interest, which is a challenging task due to substantial differences such as dissimilarity in sensing methodology (e.g., wavelength, active/passive image acquisition), geometric and radiometric differences, etc. Reviewing previous attempts on multiple-domain imagery matching, SIFT key point matching is the primary candidate to tackle this problem, due to its scale and rotation invariant property. However, the matching results appear to be less reliable for different image domains. For instance, the key points generated from LiDAR intensity and optical images are rather different, and thus, no common features may be found, and consequently, SIFT matching fails. The objective of this paper is to identify a feature space that can better support multiple-domain image matching, including the evaluation of different matching techniques and comparing their performance. In this paper, three region descriptor-based matching techniques, namely, probability density function (PDF) matching, covariance matching, and edge normalized cross correlation (NCC) matching are applied to the LiDAR intensity and satellite/aerial image pairs. The performance of the three methods is compared and evaluated under the similar conditions, including the analysis of their advantages and disadvantages. Initial results reveal that PDF feature space has promising potential for multiple-domain imagery matching in the tested image domains.

KEYWORDS: multiple-domain imagery matching, LiDAR, satellite/airborne imagery

INTRODUCTION

Image matching is regarded as a pattern matching problem. Based on how patterns are described and compared, matching methods can be generally categorized into two classes: feature-based and area-based matching. It is common to combine techniques from both classes to achieve better and reliable matching results, such as in the digital photogrammetric approach, where aerial image matching accuracy can achieve sub-pixel level by starting from coarse feature-based matching and continuing to fine area-based matching (combined with blunder detection). Generally speaking, matching performance among images acquired from the same sensor or similar sensors is rather reliable, even for image pairs of the same scene taken from arbitrary positions. In particular, research in computer vision has improved performance in this area in past decade. One of these techniques is the Scale-Invariant Feature Transformation (SIFT), published by Lowe in 1999, a very robust feature-based matching method widely used in the computer vision community. Besides the impressive developments in matching techniques, image sensors have been rapid evolving too. Consequently, the availability of imagery from different domains covering the same ground region is increasingly becoming common, such as high-resolution satellite imagery, multispectral aerial imagery, and elevation/intensity imagery rasterized from point cloud acquired by LiDAR/IfSAR systems, etc. Therefore, interest is growing to co-register all these data to provide reliable information for remote sensing applications, such as change analysis. Multiple-domain imagery matching is still an undergoing research topic; in this study, matching between LiDAR data and aerial/satellite optical data domains is investigated.

Since airborne LiDAR point cloud data is generally sparse, it is almost impossible to identify point features from the point cloud. Therefore, other primitives, such as 3D straight lines and/or surface patches are considered for matching features (Habib, *et al.*, 2007; Habib *et al.*, 2004; Habib, *et al.*, 2004; Kim and Habib, 2009). On the other hand, with increasing laser point density, better LiDAR intensity data is becoming available. This brings a new opportunity to match LiDAR data to other optical imagery. As a first approach, it is natural to consider applying SIFT to this task (Abedini, *et al.*, 2008). According to our earlier experiences with a limited data set, SIFT was successfully applied to aerial and satellite imagery registration (Toth, *et al.*, 2010). However, in our extended tests, SIFT failed to provide robust registration between LiDAR intensity and optical images. Therefore, other possible approaches are considered in this study, including three other region descriptors, the probability density function (PDF) descriptor (Comaniciu, *et al.*, 2003), covariance descriptor (Porikli, *et al.*, 2006a, 2006b) and the descriptor using edges. The purpose of this investigation is to identify which feature space is most suitable for the LiDAR intensity and optical multiple-domain image matching task.

For the tests, the PDF matching and covariance matching are implemented in MATLAB and the normalized cross correlation edges matching is implemented in OpenCV-supported C++ environment. The performance comparison of different methods is based on using identical data.

SIFT MATCHING CHALLENGES

SIFT matching is a well-known robust, scale and rotation invariant and semi-affine invariant matching technique; a comprehensive description on SIFT can be found in (Lowe, 2004). Generally, it provides an excellent solution to identify the correspondences between the image pair with large perspective and scale changes as well as intensity variations. Since mismatches cannot be ruled out, blunder detection based on RANSAC (RANdom Sample Consensus) (Fischler and Bolles, 1981), is always a necessary follow-on task. Figure 1, shows an example where SIFT worked well on satellite/aerial multiple-domain imagery matching. However, based on our limited data sets, SIFT matching between intensity and optical imagery is not reliable. Using several LiDAR intensity and optical imagery data sets, SIFT failed in most cases. Figure 2 illustrates the best SIFT matching result between satellite and intensity images. Compared to the satellite/aerial image pair, the number of matched features is much smaller, and, the correct matches are all in the parking lot, where the special ground marks are recognizable in both images. Typical SIFT matching result between satellite and intensity images is illustrated in Figure 3; none of the matches is correct for that image pair.

SIFT matching is challenged due to the substantial differences between the LiDAR intensity and satellite/aerial image domains, such as different sensing methodology (e.g., wavelength, passive/active image acquisition), geometric and radiometric differences, etc. These factors can cause the extracted key points to be quite different between intensity and aerial/satellite images; even for those key points extracted from similar locations, their descriptors could be still rather different. Note SIFT key point descriptor is based on the gradient magnitude and orientation in a region around the key point, and then weighted by a Gaussian window, which is, eventually, accumulated into orientation histograms. Usually, a 4×4 descriptor with 8 orientations, resulting in feature vector of 128 dimensions, is used. While it is a very creative and efficient way to describe a feature point, it may not be the best choice for intensity and satellite/aerial image pairs.

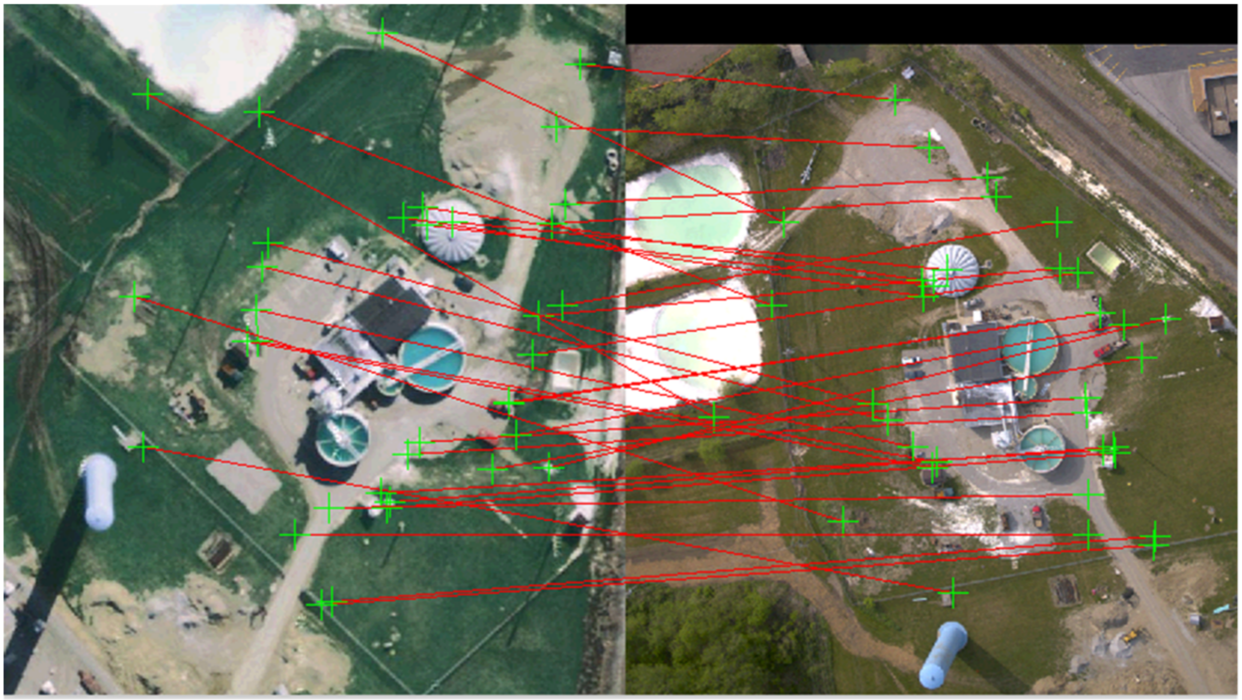

Figure 1. Aerial and satellite image pair with matched SIFT features;
aerial image (left) and satellite image (right).

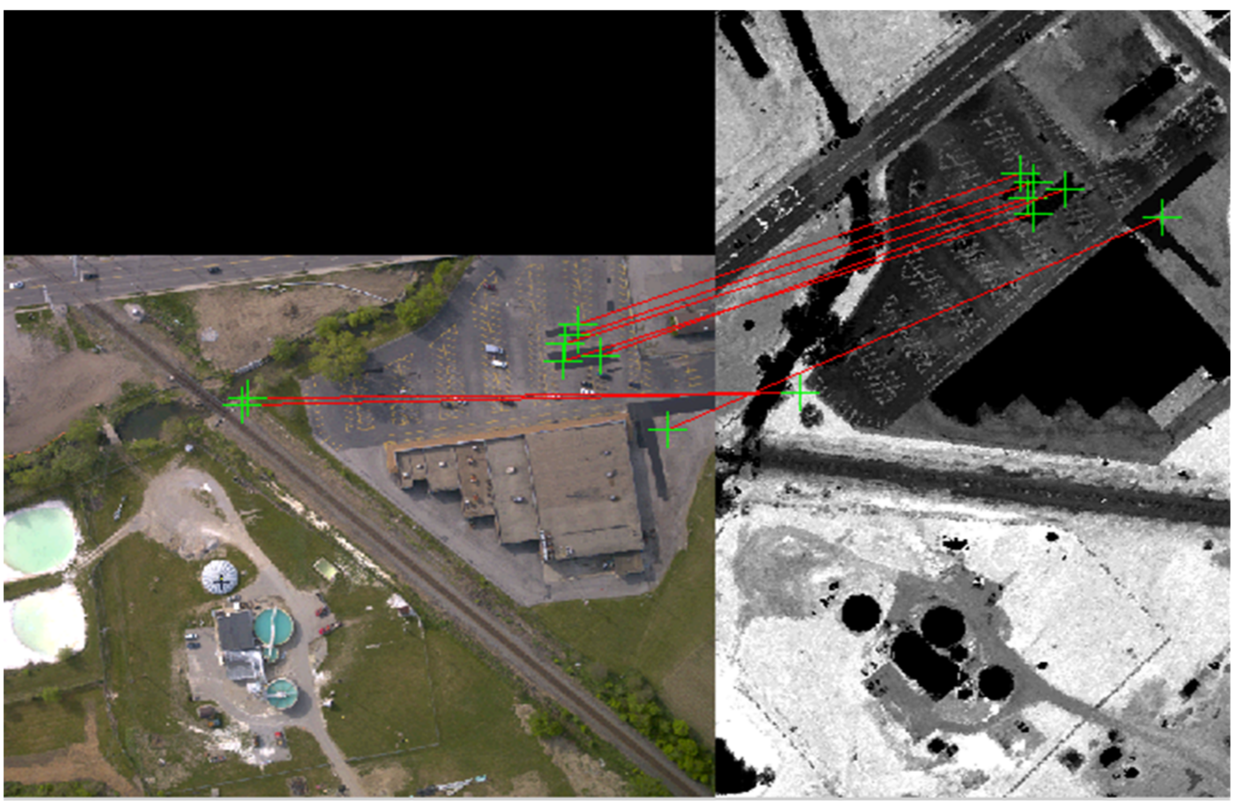

Figure 2. Satellite and LiDAR intensity image pair with matched SIFT features;
satellite image (left) and intensity image (right)**.**

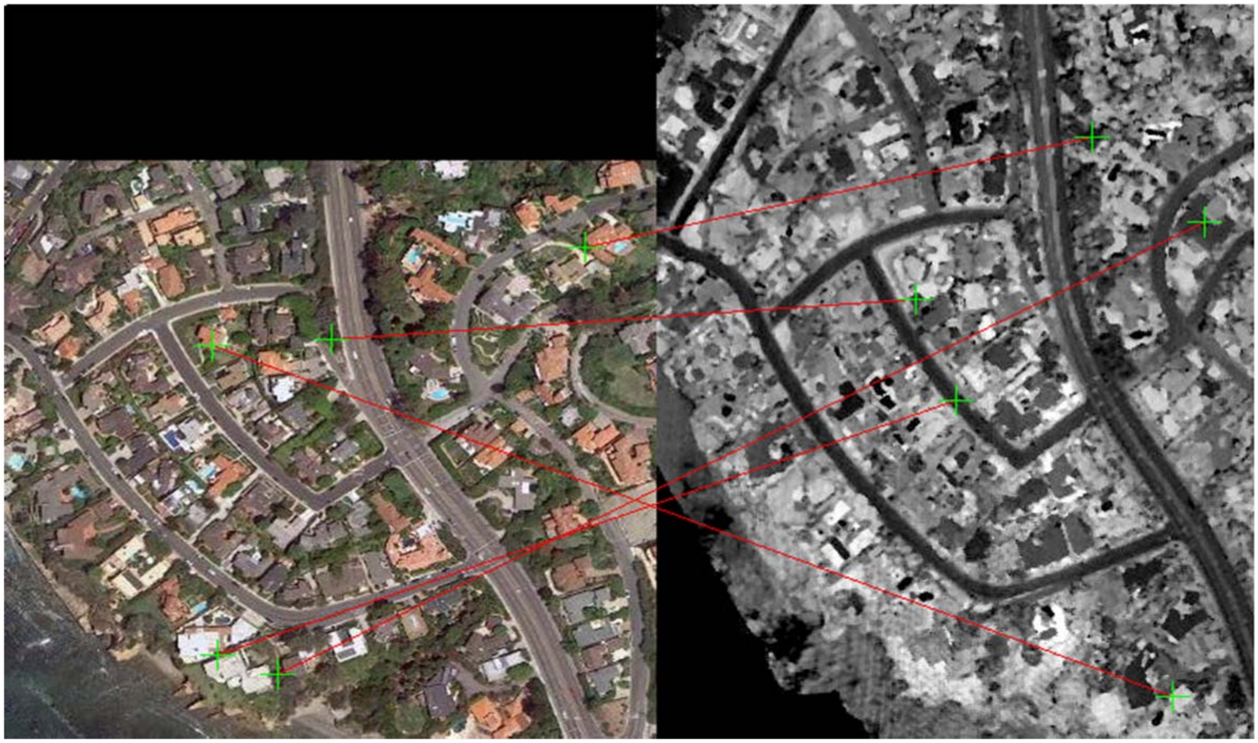

Figure 3. Satellite and LiDAR intensity image pair with matched SIFT features;
satellite image (left) and intensity image (right)**.**

TESTED METHODS

In this study, image rectangle region is used as the feature, and different feature spaces and descriptors are tested to identify a suitable approach to perform matching between intensity and satellite/aerial images. The methods considered are PDF matching, covariance matching and NCC-based edge matching.

PDF Matching

PDF (Probability Density Function) matching is typically used in the mean-shift based target tracking. The target model is represented by its PDF in the feature space. Feature space could be the intensity value, RGB value, etc. In our case, only the intensity PDF is available for both domains. The intensity PDF of a region feature can be approximated by the normalized histogram (Comaniciu, *et al.*, 2003), and represented in a 256-dimension feature descriptor. The similarity between two PDF descriptors is computed via the Bhattacharyya Coefficient, which is the cosine of angle correlation between the two PDF descriptors, defined as:

$$\rho = \rho(p,q) = \sum_{u=1}^{m} \sqrt{p_u \cdot q_u} = \cos\theta \geq 0$$

where

$$\sum_{u=1}^{m} p_u = 1 \text{ and } \sum_{u=1}^{m} q_u = 1$$

ρ is always non-negative, as the two normalized unit vectors p and q (PDFs) are both non-negative. The maximal similarity score is 1 which means the two feature descriptors are exactly the same. The minimal similarity score is 0 which means the two feature descriptors are orthogonal to each other; in other words, they don't have any relation. To give an example, a reference patch is selected in the satellite image, and then its PDF descriptor is computed as the target model. Subsequently, a candidate patch with the same size of the reference one is shifted in the intensity image suing a simple brute force searching for the maximum similarity.

Figure 4 illustrates the matching result; the green rectangle region is matched, the region size is 100 100 pixels. The lower left corner subfigure represents the similarity score surface and the lower right corner subfigure shows the two PDF descriptors together; blue one is from the reference patch, and red one is from the best matched patch in the intensity image.

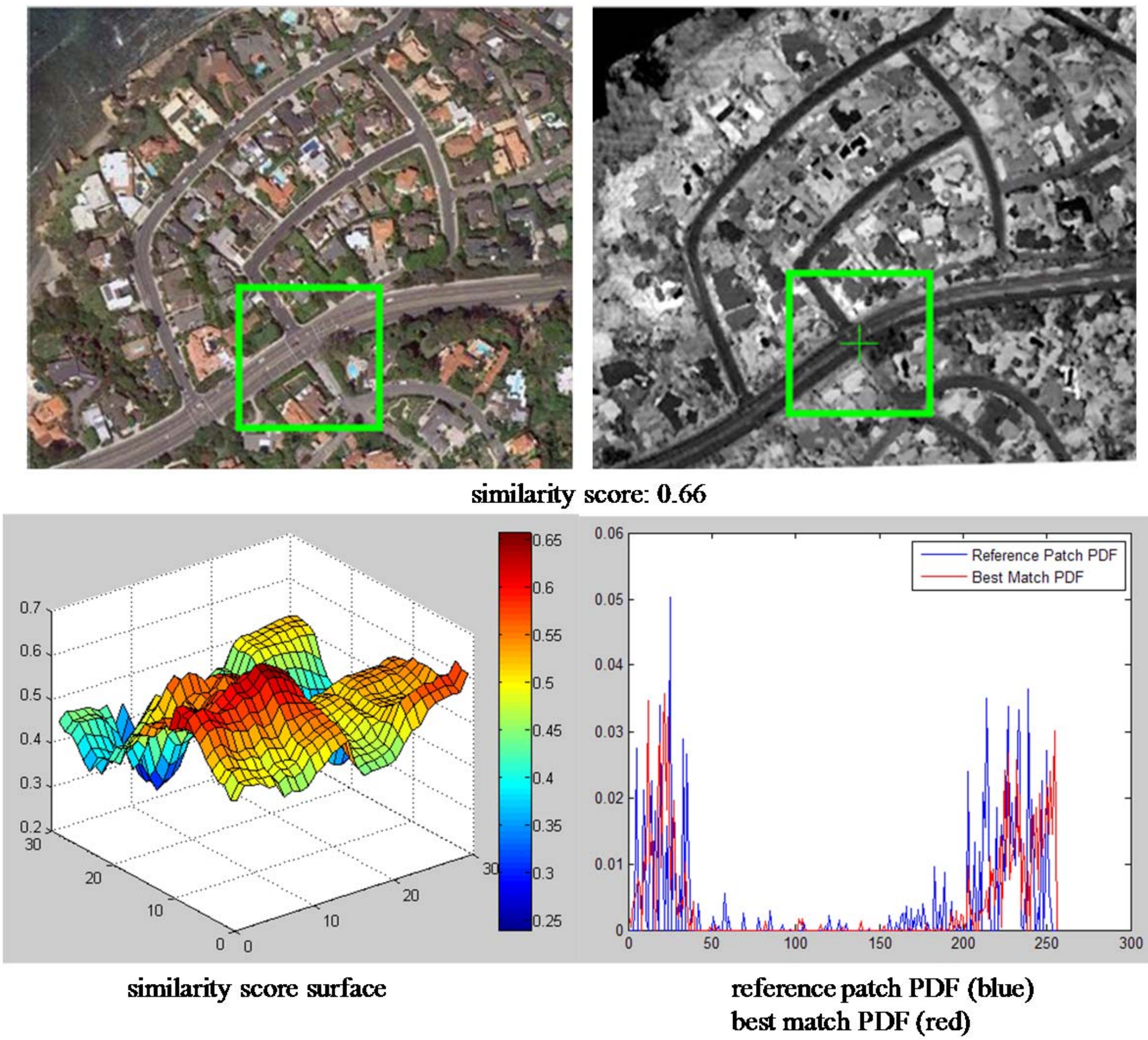

Figure 4. PDF matching example; aerial image (top left) and intensity image (top right).

Covariance Matching

Covariance matching is often used in target tracking; the target is enclosed in a rectangle region and the target model is computed based on the covariance in the feature space. The feature space could be intensity value, RGB value, the norm of first and second derivatives of intensity with respect to row and column directions, etc. For our case, the feature space is [row, column, intensity], and the feature descriptor of the region is the 3 3 covariance matrix of the 3-dimensional measurements. The similarity between two feature descriptors is computed by the distance between the two matrices, defined in a Riemannian Manifold (Porikli, *et al*., 2006a, 2006b):

$$\rho(C_{Model}, C_{Candiate}) = \sqrt{\sum_{i=1}^{r} \{\ln[\lambda_i (C_{Model}, C_{Candiate})]\}^2}$$

where $\lambda_i()$ is the generalized eigenvalues. The minimum distance indicates the best match. For example, a reference patch is selected in the aerial image, and then its feature covariance matrix is computed. Subsequently, the candidate patch with the same size of the reference one is shifted in the intensity image, and a brute force searching for the minimum distance in the searching region is performed. Figure 5 shows the covariance matching example; the green rectangle region is matched, the region size is 100   100 pixels. The right subfigure shows the distance surface, the minimum distance is 0.32.

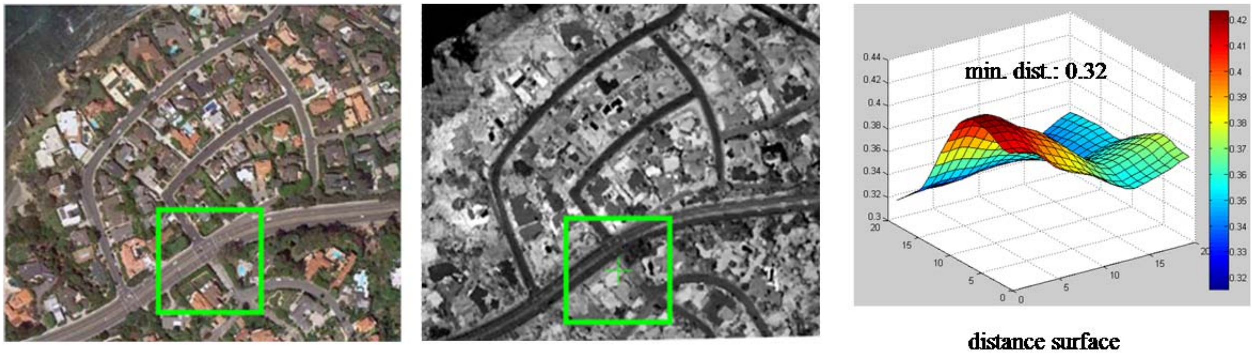

Figure 5. Covariance matching example, aerial image (left), intensity image (middle), distance surface (right).

NCC Edge Matching

As mentioned earlier, it is almost impossible to identify and extract point features from the point cloud. Even when the point cloud is rasterized into intensity image, point features are generally not good primitives due to the rather high level of noise. Consequently, linear features are often used to co-register the LiDAR data and other optical data. This idea is adapted to develop NCC edge matching. In this case, the feature space is the binary edges in the selected region. For example, a reference patch is selected in the satellite image with edges extracted by the Canny algorithm. Subsequently, the intensity image is converted to a binary edge image and a brute force searching for the maximal NCC value in the whole intensity image is performed.

Figure 6 represents the NCC edge matching example; the yellow and red rectangle regions are matched. Their corresponding NCC maps are showed in the bottom row of the figure. This method has an obvious limitation in flat regions where no edges exist. In order to avoid those patches, a check on the entropy and number of white pixels may be performed.

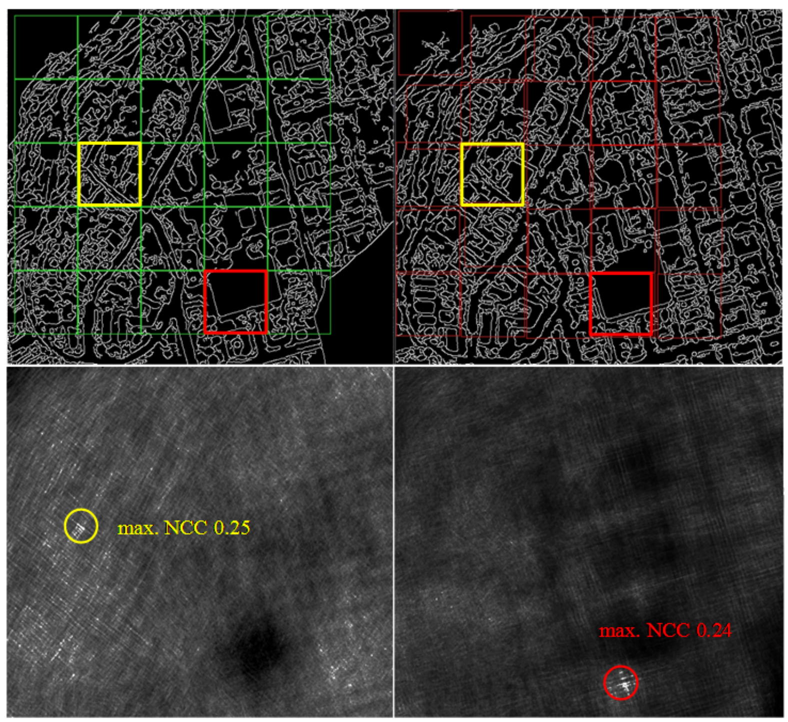

Figure 6. NCC edge matching example result; edge image from intensity image (top left), from satellite image (top right), NCC map for yellow patch (bottom left), and for red patch (bottom right).

TEST RESULTS AND ANALYSIS

Data Preparation

LAS format LiDAR data acquired by Fugro-EarthData in 2009 is used to generate 1m GSD intensity raster image. The maximum intensity value of all points that fall into the 1m by 1m region is taken as the pixel intensity value. If there is no laser point in the region, the pixel intensity is set 0 (black). Black holes are usually seen in the rasterized LiDAR intensity image, which can be reduced by using a median filter.

As for the optical sensor imagery, IKONOS satellite imagery acquired by GeoEye in 2010 and aerial imagery scanned from aerial photos from the National Aerial Photography Program (NAPP) by the U.S. Geological Survey (USGS) are used. The satellite image is orthorectified with 1m GSD. The aerial imagery has about a 3.3m GSD.

For our tests, all satellite and aerial images are rescaled to the 8-bit intensity range. Each test image pair is resampled to the same spatial resolution and has the same main orientation. In other words, there is no significant scale and rotation difference between image pair. Since the aerial image has the coarsest resolution, intensity images have to be down-sampled, and the size of intensity/aerial image pair is smaller than intensity/satellite image pair. Five intensity/satellite image pairs and five intensity/aerial image pairs are used in our tests.

Evaluation Method

This study is aimed to identify which region descriptor is appropriate for the intensity and optical image matching, by comparing the performance of the three methods under the same conditions. Removing the scale and rotation difference between intensity and optical data sets assures that all three methods will work. For evaluating the performance, the reference patches are generated in the reference image; those reference patches seem like a grid in the reference image. This is ideal for evaluating the performance of the different region descriptor matching methods, since they all use the same reference patches. Then, for each reference patch, the best match is found in the searching image. Ideally, the best match patch should have the same position as the reference patch position, since the image pair is already aligned. The position error is then computed as $\delta p = \sqrt{\delta x^2 + \delta y^2}$. The search window is the region size with a 10-pixel extended boundary in width and height. For example, if the patch size is 50 pixels, and then the searching window has the size of 70 pixels. This way, the maximum position error detected is $\delta p_{\max} = \sqrt{10^2 + 10^2} = 14.14$ pixels. Figure 7 illustrates a sample reference grid (left) with best matches (right). Using such a test configuration, the accuracy of these three methods can be computed and compared under the same conditions. Each matching method has its own matching precision measure; e.g., the similarity score or minimum distance.

For the PDF matching, the similarity score, S_{sim} is between 0 and 1; the larger the better. The position error is from 0 to 14.14-pixel, the smaller the better. The position error is linearly transferred to the [0-1] range; i.e., $S_{pos} = 1 - \delta p / 14.14$, small is better. The matching score is defined as $S_{PDF} = 0.5 \cdot (S_{pos} + S_{sim})$; S_{pos} and S_{sim} have the same weight. The matching score S_{PDF} is in the range of 0 and 1; the larger the better. If S_{PDF} of a match is larger than the mean S_{PDF} of all matches, it is regarded as a good match.

For the covariance matching, the similarity score S_{sim} is the distance between the two covariance matrices; smaller is better. Similarly, position error is included in the matching score; a good match should have both small S_{sim} and S_{pos}. For each match, if both of them are smaller than the mean S_{sim} and S_{pos} of all matches, the match is regarded as a good one. If either one is bad, then the match is regarded bad.

For the NCC edge matching, the similarity score S_{sim} is in the [0-1] range; the same definition of PDF matching score is taken here too, which means $S_{NCC} = 0.5 \cdot (S_{pos} + S_{sim})$ where $S_{pos} = 1 - \delta p / 14.14$. If S_{NCC} of a match is larger than the mean S_{NCC} of all matches, it is regarded as a good match.

Using the above patch criteria, both position error (accuracy) and the similarity quantity (precision) are considered. Left subfigure in Figure 8 plots the position errors for the good matched patches of a LiDAR intensity and satellite image pair, while right subfigure in Figure 8 plots the position errors and the PDF similarity scores of the same image pair. Note that the plotted similarity score is enlarged 10 times for better visualization. The matches with large position errors have rather good similarity scores, which make them still good matches.

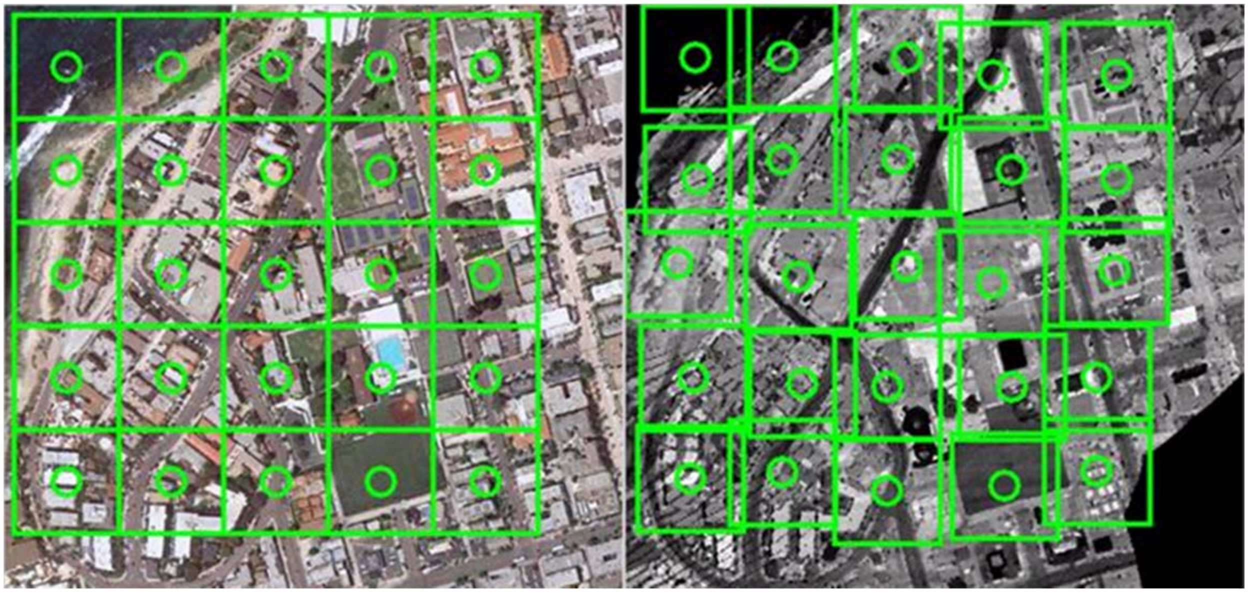

Figure 7. Evaluation gird, reference patches in satellite image (left),
matched patches in intensity image (right).

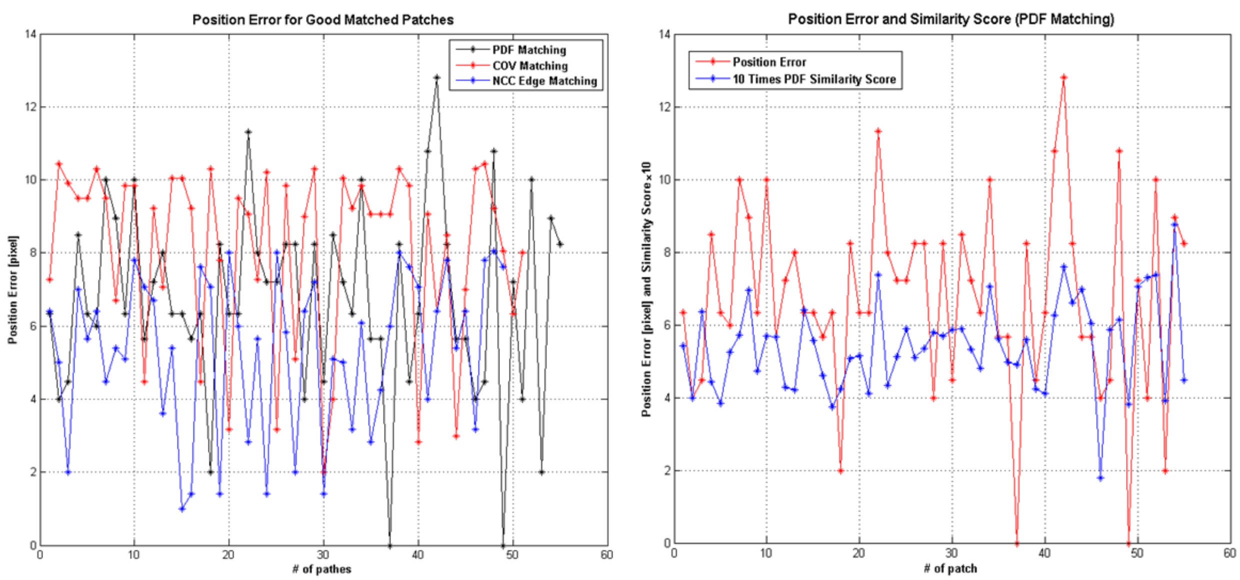

Figure 8. Position errors for good matched patches (left),
position errors and PDF similarity scores for good matched patches (right).

Matching Results and Analysis

Five LiDAR intensity/satellite image pairs and intensity/aerial image pairs covering both residential and non-residential areas are selected for this performance evaluation. After several experiments on optimizing the patch size and searching window size, 50-pixel patch size and 70-pixel searching window size are used for all image pairs. The matching results from LiDAR intensity/satellite image pairs are listed in Table 1, and results from LiDAR intensity/aerial image pairs are listed in Table 2.

Table 1. Matching Results (LiDAR intensity and satellite images)

Image Pair		A	B	C	D	E
# of Patches		68	40	78	175	115
Image Size (W×H) [pixel]		877×225	427×295	694×370	1299×375	1197×287
# of Good Matches (%)	PDF	31 (45.6%)	18 (45.0%)	39 (48.7%)	88 (50.3%)	55 (47.8%)
	COV	21 (30.9%)	15 (37.5%)	17 (21.8%)	66 (37.7%)	51 (44.4%)
	NCC	19 (27.9%)	17 (42.5%)	39 (50.0%)	77 (44.0%)	49 (42.6%)
Mean Similarity Score (PDF)		0.59	0.45	0.60	0.60	0.54
Mean Matrices' Distance (COV)		0.08	0.07	0.07	0.27	0.20
Mean NCC Score		0.30	0.28	0.29	0.25	0.26
Mean Pos. Error [pixel]	PDF	7.38	5.85	7.58	7.33	6.75
	COV	6.99	8.61	6.12	7.64	8.11
	NCC	5.83	6.33	4.09	4.52	5.37

Obviously, mismatches cannot be avoided for all three methods, because the reference patch could be from a rather flat region (LiDAR intensity); in other words, this region feature is not a good primitive in the image, see Figure 9. Since the evaluation of different feature descriptors for matching, and not how to extract strong primitives in image, is emphasized in this paper, such weak region features are acceptable. More importantly, same reference patches are used, which guarantee all three methods are compared under the same conditions.

Table 2. Matching Results (LiDAR intensity and satellite images)

Image Pair		A	B	C	D	E
# of Patches		64	30	24	24	26
Image Size (W×H) [pixel]		375×140	669×85	279×98	294×104	585×79
# of Good Matches (%)	PDF	29 (45.3%)	11 (36.7%)	12 (50%)	9 (37.5%)	12 (46.2%)
	COV	26 (40.6%)	12 (40%)	5 (20.8%)	6 (25%)	9 (34.6%)
	NCC	40 (62.5%)	16 (53.3%)	11 (45.8%)	15 (62.5%)	15 (57.7%)
Mean Similarity Score (PDF)		0.46	0.44	0.45	0.43	0.42
Mean Matrices' Distance (COV)		0.15	0.15	0.14	0.17	0.54
Mean NCC Score		0.38	0.37	0.45	0.42	0.40
Mean Pos. Error [pixel]	PDF	6.88	5.45	6.14	5.74	7.04
	COV	8.32	5.95	6.81	7.12	8.92
	NCC	2.53	3.39	1.31	2.67	3.33

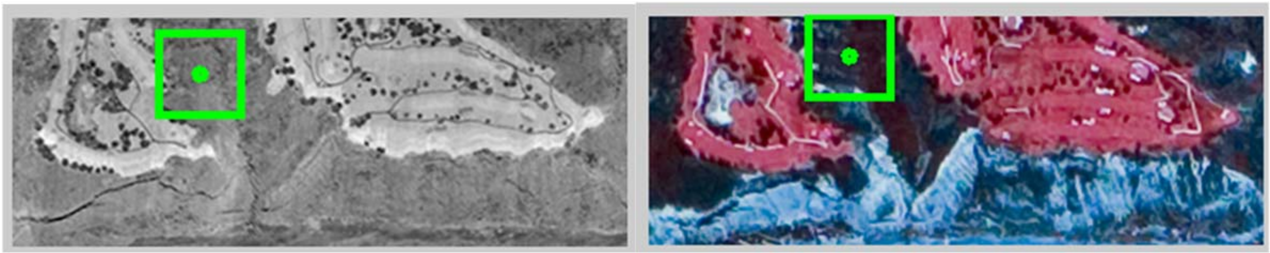

Figure 9. Example of mismatched patch (LiDAR intensity/aerial image pair).

An overview of the matching performance is listed in the Table 3. Position errors are generally large for all methods. First, the reference patch may not be an ideal feature region. Second, although image pair is manually aligned to remove translation, scale and rotation differences, the geometric deformation between the image pair may still exist. PDF and covariance matching have similar performance, since their feature spaces are related; i.e. the intensity value. Position errors have no significant difference between the two domain combinations for PDF and covariance matching. NCC Edge matching is a quite different approach; its feature space is not directly related to PDF and covariance feature space. Its performance is the best among all methods. However, the NCC method is known to be limited to no scale and rotation variation in the image pair. On the other hand, both PDF and covariance descriptors can be easily adapted to a rotation-invariant descriptor, as PDF and covariance descriptor of a symmetric circular region feature is rotation invariant. Comparing the similarity score, PDF matching is the best in all cases, which indicates that PDF descriptor is an appropriate choice for the intensity and optical image matching.

Table 3. **Overview of the Matching Performance**

	PDF Mean Similarity Score	Covariance Mean Similarity Distance	Edge NCC Mean Score	PDF Mean Position Error [pixel]	Covariance Matching Mean Position Error [pixel]	Edge NCC Matching Mean Position Error [pixel]
Intensity/Satellite Image Pair	0.55	0.14	0.28	6.98	7.49	5.23
Intensity/Aerial Image Pair	0.44	0.23	0.40	6.25	7.42	2.65

CONCLUSION AND FURTHER WORK

In this paper, the challenge the LiDAR intensity and optical imagery domain matching is discussed. Our earlier experiences with SIFT inspired us to investigate three other region feature descriptors for multiple-domain image matching. All these methods are evaluated and compared with their performance based on the five LiDAR intensity/satellite and LiDAR intensity/aerial image pairs. Initial results show that PDF descriptor is an appropriate choice, as it outperformed all the other techniques. Current research work is to adapt the PDF descriptor to a scale and rotation invariant variant to improve matching between LiDAR intensity and satellite image domains.

ACKNOWLEDGMENT

The satellite imagery provided by the GeoEye Foundation is greatly appreciated.

REFERENCES

Abedini, A., M. Hahn, and F. Samadzadegan, 2008. An investigation into the registration of LiDAR intensity data and aerial images using the SIFT approach, In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXVII Part B1*, pp. 169.

Comaniciu, D., V. Ramesh, and P. Meer, 2003. Kernel-based object tracking, *IEEE Transactions on Pattern Analysis and Machine Intelligence, 25*(5)*:564-577.*

Fischler, M. A., and R.C. Bolles, 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography, *Communications of ACM 24*: 381-395.

Habib, A. F., K.I. Bang., M. Aldelgawy, S.W. Shin, and K.O. Kim, 2007. Integration of photogrammetric and LiDAR data in a multi-primitive triangulation procedure, *Proceedings of ASPRS 2007 Annual Conference - Identifying Geospatial Solutions*, Tampa, FL-USA.

Habib, A. F., M.S. Ghanma, and M. Tait, 2004. Integration of LiDAR and photogrammetry for close range applications, In: *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXV, Part B5*.

Habib, A. F., M.S. Ghanma, M.F. Morgan, and E. Mitishita, 2004. Integration of laser and photogrammetric data for calibration purpose, *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. XXXV part B2*, pp. 170.

Kim, C., and A.F. Habib, 2009. Object-based integration of photogrammetric and LiDAR data for automated generation of complex polyhedral building models, *SENSORS, 9*(7): 5679-5701.

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

Porikli, F., O. Tuzel, and P. Meer, 2006a. Covariance tracking using model update base on Lie Algebra, In: *the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, Vol. 1, pp. 728-735.

Porikli, F., O. Tuzel, O., and P. Meer 2006b. Covariance tracking using model update based means on Riemannian Manifolds, In: Proceeding IEEE Conference on Computer Vision and Pattern Recognition.

Toth, C. K., H. Ju, and D.A. Grejner-Brzezinska, 2010. Experiences with using SIFT for mulitple image domain matching, In: Proceedings of ASPRS 2010 Annual Conference, San Diego, CA-USA.