

AUTOMATIC FEATURE DETECTION, DESCRIPTION AND MATCHING FROM MOBILE LASER SCANNING DATA AND AERIAL IMAGERY

Zille Hussnain*, Sander Oude Elberink, George Vosselman

ITC, Faculty of Geo-Information Science and Earth Observation of the University of Twente, Enschede, The Netherlands -
(s.z.hussnain, s.j.oudeelberink, george.vosselman)@utwente.nl

Commission I, ICWG I/Va

KEY WORDS: Feature detection, description, matching, mobile laser scanning, point cloud, aerial imagery

ABSTRACT:

In mobile laser scanning systems, the platform's position is measured by GNSS and IMU, which is often not reliable in urban areas. Consequently, derived Mobile Laser Scanning Point Cloud (MLSPC) lacks expected positioning reliability and accuracy. Many of the current solutions are either semi-automatic or unable to achieve pixel level accuracy. We propose an automatic feature extraction method which involves utilizing corresponding aerial images as a reference data set. The proposed method comprise three steps; image feature detection, description and matching between corresponding patches of nadir aerial and MLSPC ortho images. In the data pre-processing step the MLSPC is patch-wise cropped and converted to ortho images. Furthermore, each aerial image patch covering the area of the corresponding MLSPC patch is also cropped from the aerial image. For feature detection, we implemented an adaptive variant of Harris-operator to automatically detect corner feature points on the vertices of road markings. In feature description phase, we used the LATC binary descriptor, which is robust to data from different sensors. For descriptor matching, we developed an outlier filtering technique, which exploits the arrangements of relative Euclidean-distances and angles between corresponding sets of feature points. We found that the positioning accuracy of the computed correspondence has achieved the pixel level accuracy, where the image resolution is 12cm. Furthermore, the developed approach is reliable when enough road markings are available in the data sets. We conclude that, in urban areas, the developed approach can reliably extract features necessary to improve the MLSPC accuracy to pixel level.

1. INTRODUCTION

Over the past few years, use of the mobile mapping data products have been growing constantly. Data providers want to produce highly accurate data products, and to generate them more frequently at lower costs. However, it is always necessary to utilize manual ground control points for the data correction and adjustment due to the low positioning accuracy of the GNSS in urban areas.

Automatic feature extraction between Mobile Laser Scanning Point Cloud (MLSPC) and Aerial imagery is very advantageous for maintaining MLSPC product quality and evaluation. Especially, in the urban canyons, where GNSS based positioning inaccuracies are common. A solution to the problem is to use manually measured and handpicked Ground Control Points (GCPs), which could not achieve the pixel level accuracy. Moreover, acquiring the ground control point is a very labour intensive and tedious work and hinders the acquisition of the high quality data in automatic fashion. There are many techniques available, which can try to minimize the number of ground control points which are required to correct a data product. However, even acquisition of fewer ground control points requires manual interventions. This manual post processing step of data correction forces surveyors to survey a city site less frequently at the cost of more manual effort, while as a consequence customers use the outdated, imprecise and expensive data sets.

Previous research has shown that the automatic image feature extraction can be used for the registration between two data sets from different sensors. The obtained transformation can be used to correct a data set when the other already has a reliable accuracy. Jende, Hussnain et al. (2016) have shown the preliminary results of registration between aerial and ground (MLSPC and terrestrial ortho images) data sets. Similarly, Gao, Huang et al. (2015) have improved the Mobile Laser Scanning (MLS) data accuracy by its automatic registration with high –resolution –accurate UAV's imagery. They performed bundle adjustment between UAV imagery and rasterized MLSPC ortho image patches using the Harris corner keypoint detection and edge based template matching. This work reported the RMS $\Delta X=0.086m$ - $\Delta Y=0.063m$ - $\Delta Z=0.106m$ in the corrected data set. A SIFT based feature detection and matching approach for registration of aerial lidar to aerial image was described in Abedini, Hahn et al. (2008), however, the achieved accuracy of feature matching was not described.

Frueh and Zakhor (2003) proposed to utilize the aerial imagery as a reference global edge map for the lidar observation on the ground and achieved an absolute coordinate difference of ± 5 meters using Monte-Carlo-Localization (MCL). Kümmerle, Steder et al. (2011) developed a simultaneous localization and mapping (SLAM) approach that computes its global position for every node,

* Corresponding author

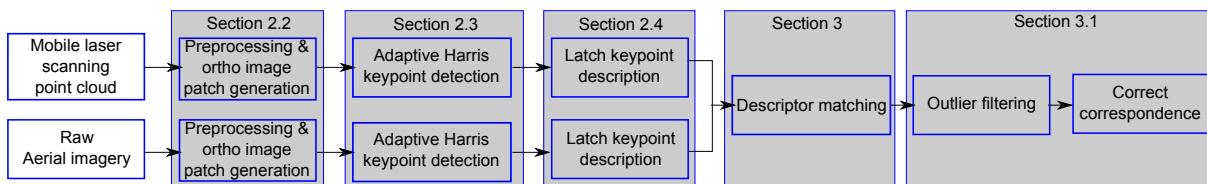


Figure 1. Flow diagram of the developed method for MLSPC to Aerial image registration.

Where MLSPC is the observation and the aerial image is the reference. In this approach, the accuracy is defined by the distance from true orientation of building edge/wall to the estimated orientation of edge/wall, and it is 0.2m for five used test data sets.

The automatic feature extraction and matching from two sensors is a challenging task, due to difference in sensor characteristics. The aerial image is acquired through camera by capturing white light reflection of scene, where, the MLSPC is acquired by lidar sensor and represents the geometrical information of the scene in the form of 3D positions and the reflection intensity of each surface point. Moreover, the optical images have a regular grid of pixels over the image space, whereas the MLSPC's point density is variable and depends on the distance of the lidar sensor to object and on speed of MLS car. Furthermore, the laser reflection intensity of each point is not same as the white light surface reflection intensity.

In this scenario, designing an automatic feature extraction technique which can detect common features, is a difficult but important task. Once achieved automatically, the feature extraction can save lots of extra effort, cost and time. It could be used to improve the accuracy of the already refined or raw data sets. Moreover, the data can be acquired more frequently because no manual adjustment would be required. Furthermore, the extracted features could be used for the automatic data evaluation and quality control.

The purpose of our research work is to develop an automatic pipeline, which can be used to correct the MLSPC using the aerial images of the corresponding area. The technique is developed to automatically extract the accurate low-level (2D) features from MLSPC and aerial imagery.

In the given situation and the data sets, it is apparent that the road surface is a biggest part in the data which is common between the aerial image and the MLSPC. Apart from the different characteristic of the involved information, the road plane area can be used for the 2D features extraction. Mainly, the road markings are objects which are visible in aerial images and also present in the MLSPC as laser reflectance of each 3D point. To generate a 2D image from MLSPC, 3D points are projected to 2D ground plane, and the reflection property of each 3D point is assigned to the corresponding pixel in 2D image. The MLSPC are long consecutive series of points and due to the limitation of the computational power, it is useful to crop the point cloud in to small pieces. Therefore, first the MLSPC is patch-wise cropped to small tiles and then each tile is projected to ground plane to generate a 2D image. Furthermore, each aerial image patch of the same size and covering the same area is also cropped from a large original aerial image. Then for the feature detection from the ortho images, we implemented an adaptive variant of Harris keypoint detection, so, it can automatically detect

common and reliable features from the images of the different sensors. For description of the feature points, we use binary descriptor called Learned Arrangements of Three Patch Codes (LATCH), which is robust to noise in both the MLSPC and aerial ortho images because it utilizes patches of pixels instead of single pixel to establish a binary relation. For descriptor matching, firstly the descriptors are matched based on their binary differences and secondly filtered by exploiting the arrangement of relative Euclidean-distances and angles between both corresponding sets of feature keypoints.

The developed technique is dependent on the road markings in data sets. So, in case the road markings are not presents in the both data sets at all or completely repainted to different positions, then this technique could not yield reliable correspondences. Usually, it is not required to have road marking in each tile because state of the art position filtering methods need accurate correspondences only after every 100 meters.

The organization of sections in this paper is presented according to the developed method's work flow diagram in Figure 1. At the end, section 4 describes results followed by evaluation and conclusion.

2. FEATURE EXTRACTION

As, described earlier that feature extraction will be performed on the 2D images, therefore, the data sets should be cropped and projected on the ground to generate the 2D images. In this section, first we will describe the selection of test area and the preprocessing followed by the feature extraction from both datasets.

2.1 Selection of test area

In this project, we used the data sets of the Rotterdam city. The MLSPC is acquired by the Topcon® IP-S3 mobile mapping system, it has a built-in 360 degree lidar sensor which captures 700,000 pulses per second. First, an arbitrary consecutive part of 527 m in MLSPC is selected, which is visualized in Figure 2. Moreover, the aerial images of the same area are also obtained, where each original aerial image has 20010x13080 pixels and the resolution of 12 cm. The visualization of the aerial imagery is shown in Figure 3. The test area is selected such that the following disturbances are included:

- i) GNSS error (MLSPC), (Figure 22)
- ii) different types of road markings (both)
- iii) occlusions (both)
- iv) traffic (both)
- v) trees (both)
- vi) shadows (aerial)
- vii) different level of noise (MLSPC)

2.2 Preprocessing

Each selected data set is cropped in to 14 tiles, where each tile has a size of 38x38 meters approximately. The size of a tile is influenced by the relative accuracy of the MLSPC and the total computational time required to detect and match features. The 14 tiles of MLSPC and aerial images are shown in Figure 2 and Figure 3 respectively.

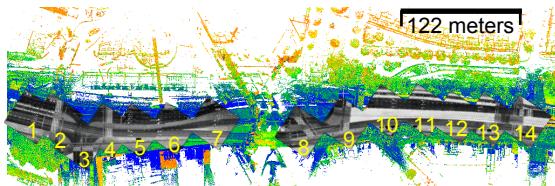


Figure 2. Visualization of the MLSPC of the test area and cropped tiles.



Figure 3. Visualization of the aerial image of the test area and cropped tiles.

2.2.1 MLSPC ortho image generation

The MLSPC is automatically cropped based on the coordinate bounds of X, Y and Z axis. The bounds are originated from the MLS platform's 3D position in the trajectory, and extended outward to crop a tile of size 38x38 meters. Now, each cropped tile is still a small point cloud as shown in Figure 4 (left). Then, to convert it to a 2D image, the obtained point cloud is projected on the ground plane and an ortho image is obtained as shown in Figure 4 (right). Moreover, the laser reflectance property of each projected 3D point is used to calculate the grey values of the corresponding pixels in raster image. The interpolation of the grey values is achieved by the linear interpolation. Furthermore, points above 4 m ground level are also removed before the projection. The 4 m height can easily remove the unnecessary information like trees, poles, building etc., while at the same time preserves the parts of road which are relatively higher.

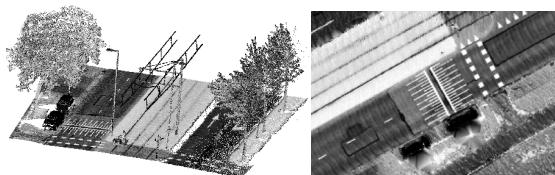


Figure 4. Point cloud patch on the left is converted to an ortho image on the right.

2.2.2 Aerial ortho image generation

As the MLS platform's 3D position is used to generate the MLSPC ortho image, the same position is also back projected to the original nadir aerial image and a patch of 300x300 pixels is cropped as illustrated in Figure 5. At this moment, it is assumed that the interior and exterior

orientations of the aerial image are very accurate and it is considered that there is no image distortion.

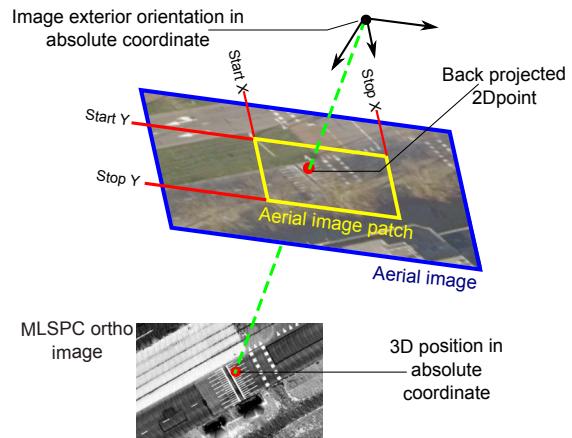


Figure 5. Acquisition of the aerial image patch by projection of 3D point.

2.3 Feature detection

The main focus in this step is to accurately detect all important features in the images. Moreover, the features should be detected reliably and automatically. Corners are the most common 2D features which could be detected from the image gradients. Therefore, a corner feature detection algorithm Harris and Stephens (1988) is used. Harris corner detector is used because it can be adapted automatically according to our feature detection requirements. An automatic parameter adaptation is required because when Harris detector is setup with fixed parameters to perform robustness to weak gradient changes then it becomes too sensitive to the sensor noise and detects many false keypoints. Moreover, the fixed parameters may not even work with the images from the same sensor. To overcome this problem, these parameters can be dynamically computed by an adaptive approach. In order to automatically adopt the parameters according to any image type, we implemented a dynamic threshold computation, which can detect important features reliably. This type of automatic parameterization is also necessary for the automatic feature extraction pipeline and can compensate for cross and inter sensor characteristics and noise.

As shown in Figure 6 (top, middle), the gradient values of the same piece of road marking in both data sets are quite different, where, the adaptive approach detects keypoints reliably as shown in Figure 6 (bottom). Usually, a corner feature can be detected at the intersections of two edges. However, sometimes due to the large difference in descriptor space, an isolated corner point may fail in descriptor matching. That's why instead of relying on a single corner, adaptive approach detects multiple corners, which are more reliable while descriptor matching. Thereby, features are detected as clusters of corner keypoints.

The adaptive approach requires a constant input parameter, which in our case is the number of required keypoints. The number of keypoints required are set to 5000 ± 100 ($100 = \text{tolerance}$), where iterations are incremented by the

gradient threshold on normalized resultant gradient image. The number of required keypoints is calculated from the size of a tile. As, a tile covers an area of $38 \times 38 \text{ m}^2$, which means roughly two keypoints for 1 m^2 . The required number of keypoints is also influenced by the computational costs for matching maximum keypoints reliably. The different steps of detection of 5000 keypoints from a MLSPC ortho image and from an aerial ortho image patch are shown together in Figure 7. An automatic adaptive approach is not only useful for the data from different sensors but also for the images captured with same camera, this type of inter sensor keypoint detection and differences are shown in Figure 8.



Figure 6. Example of adaptive Harris corner keypoint detection of a road marking, multiple keypoints (red dots) are detected over two observable corner.

2.4 Feature description

The Learned Arrangements of Three Patch Codes (LATCH) feature descriptor proposed by Levi and Hassner (2015) is used for keypoint description. LATCH is useful in our problem for two reasons, firstly, it is a binary descriptor, and thereby the descriptor matching is very fast. Where, computational time is an important issue because we would like to process a lot of keypoints. Secondly, the grey values of images from sensors of different characteristic cannot be compared directly but patterns of higher/lower grey values are comparable. Moreover, patch triplet approach is robust to noise and it is better than using a single pixel value as practiced in pixel pairs approach. Due to these properties, LATCH can ignore occluded and cluttered location in the ortho image of the MLSPC. Also, it can compensate for shadows and occlusions in aerial images. The MLSPC has different level of sensor noise than white light camera images, and LATCH descriptor avoid sensitivity to individual sensor noise by not sampling individual pixels.

The arrangement of the all patch triplet is already learned from the data sets provided in Brown, Hua et al. (2011). The data used for the training is of generic type, thereby,

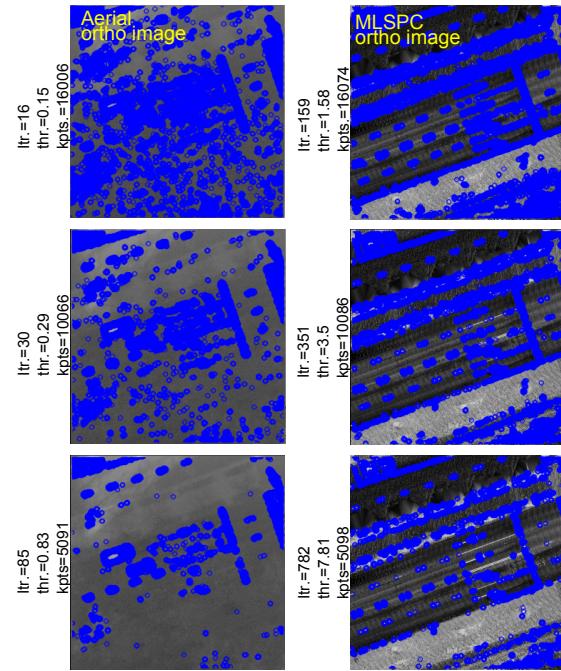


Figure 7. Adaptive Harris keypoint detection of a whole tile, with the total number of keypoints, threshold and required iterations.

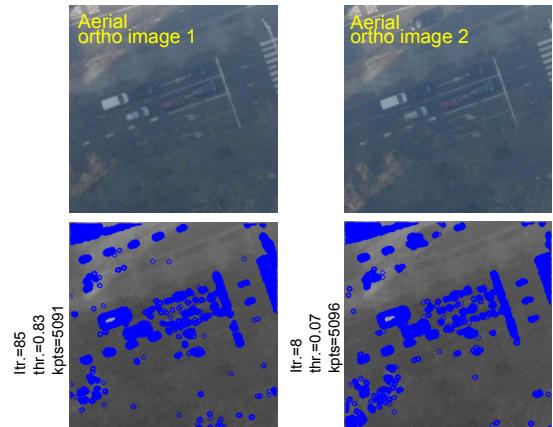


Figure 8. Adaptive approach for different aerial images of the same scene. The underlying image differences can be realized by comparing the threshold, iterations and the obtained keypoints.

it is not required to learn the arrangement from our data. Performance results of the LATCH descriptor over benchmark data set and comparison with other binary descriptors and with float descriptors are also described in Levi and Hassner (2015).

The LATCH descriptor returns a binary string of 32 bytes or 256-bits for each keypoint (Figure 9), so, a total of 256 triplets will be computed for each descriptor. The size of each mini patch is 7×7 pixels. There are total three patches in a triplet, one of the middle one is called ‘Anchor’ and other two are ‘companions’ or C_1 and C_2 Levi and Hassner (2015). A descriptor’s binary string is calculated based on

the comparison between SSD_1 and SSD_2 , here, SSD stands for the sum-of-squared difference of two (7×7 pixels) patches, where SSD_1 is computed between the anchor and C_1 patches and SSD_2 is computed between the anchor and C_2 patches. If SSD_1 is greater than the SSD_2 then the returned binary value is 0 and otherwise it is 1. The descriptors are computed around the all keypoints in both images.

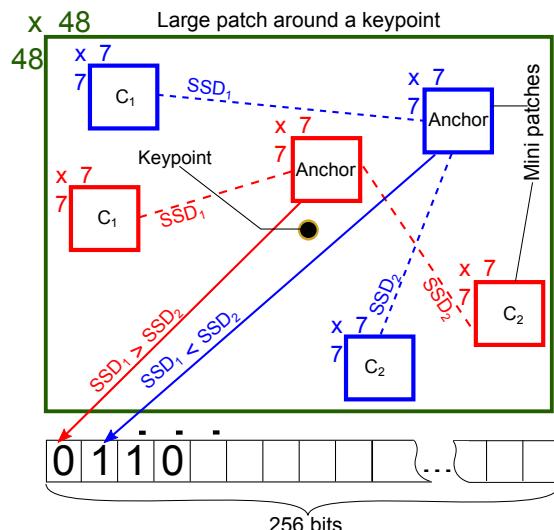


Figure 9. The computation of a particular LATCH descriptor used in this project.

3. DESCRIPTOR MATCHING

Like other binary descriptors, LATCH descriptor matching is also based on similarity of corresponding binary strings. The distance of each descriptor from a list of length n to all descriptor in list of length m is calculated by Hamming distance, which is the number of total different bits occurred in two binary strings. Moreover, for Lowe's ratio test, the threshold of 0.99 is used as proposed by Levi and Hassner (2015), which was proposed as 0.66 in the original setting Lowe (2004).

Sometime in descriptor matching, the correct corresponding descriptors are not near in descriptor space and may be even inaccurate corresponding descriptors are near in descriptor space. For this reason, instead of the one closest descriptor, first five nearest descriptors are obtained by k-NN based descriptor matching. Therefore, for an each query descriptor in first image, up to five nearest neighbours in the descriptor space of the second image are retrieved. Thereby, one query descriptor is linked with up to 5 descriptors as illustrated in Figure 10 and matched keypoints are shown in Figure 11. In the obtained correspondences, the inlier ratio is small, however, more importantly there are all important and correct correspondences. So, during inevitable situations, when there is noise and occlusion, the flexibility of descriptor matching could reliably include correct matches which could be rejected otherwise. Later, the great number of outliers can be removed by a filtering algorithm. In simple case of outlier filtering, the relative Euclidean distance and angle between the keypoints can be exploited.

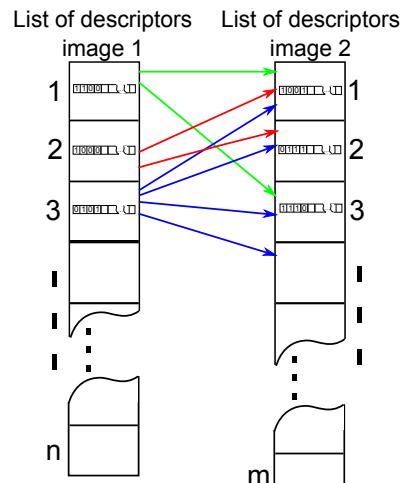


Figure 10. Illustration of the matched descriptors of image 1 and image 2.

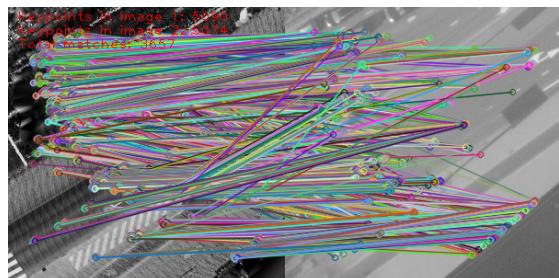


Figure 11. Hamming distance based descriptor matching of an example tile with five nearest neighbours ($k=5$).

3.1 Outliers filtering

First we will discuss the developed outlier filtering approach, followed by homography based approach, homography based technique is only useful when error in rotation is large.

3.1.1 Developed filtering approach

The filtering approach is developed to filter outlier correspondences which are introduced during the descriptor matching. Most of the outliers are included due to the descriptor matching with five nearest neighbours. As, the given data sets have a large error in translation and small error in scale and rotation, we have developed a method to remove the outliers based on their relative Euclidean distances and the angles with a seed feature point. This method is a brute-force method and iterate through each keypoint in any one of the list (in Figure 10), a seed feature point is illustrated in Figure 12. In this illustration, the image 1 has a seed feature point P_1 and its relative correspondence points are P_2, P_3 and P_4 , moreover, in image 2 the seed point P'_1 has relative correspondence points P'_2, P'_3 and P'_4 , where d_n is a relative distance of n^{th} correspondence P_n to a seed feature point and θ_n is a relative angle. Due to the distortion in data sets, these parameters are relaxed by a tolerance, so, $abs(d'_n - d_n) < tolerance_d$ and $abs(\theta'_n - \theta_n) < tolerance_\theta$. The relaxation by tolerance is consider as the maximum accuracy which can be achieved. In other words, a very small value of tolerance could yield very accurate matches but due to the distortion many potential correct correspondences will be also missed. Therefore, it is

important to not neglect these correct matches and increase the tolerance to the maximum value. Consequently, we set $toleranced$ to 0.9 pixel(s) and $tolerance_\theta$ to 0.15 degree, which maintain the final accuracy to pixel level.

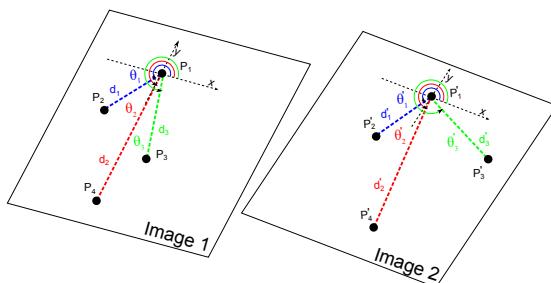


Figure 12. The $(P_2 \leftrightarrow P'_2)$ and $(P_4 \leftrightarrow P'_4)$ are correct matches in image 2 with respect to image 1 but $(P_3 \leftrightarrow P'_3)$ is an incorrect match due to different value of θ_3 .

3.1.2 Homography based filtering approach

When there is large error in rotation between two data sets (which is not the case with the test data sets), then the developed outlier filtering technique could not be used. Moreover, the Homography (computed with RANSAC) based outlier filtering could not be used directly as the inliers ratio is very low after descriptor matching. However, if the error in scale is small, then it is possible to only use the d_n constraint in the developed filtering approach, an example result is shown in Figure 13, which contains a small amount of outliers. From the correspondences in Figure 13, an accurate homography matrix can be computed as the inlier ratio is large. Now, with the obtained homography the outlier correspondences can be removed directly after the descriptor matching. The result of same example after the outlier removal using Homography is shown in Figure 14.

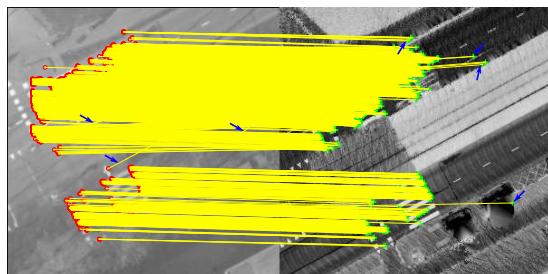


Figure 13. Correspondences computed without θ_n constraint. Blue arrows are pointing toward some visible outlier correspondences.

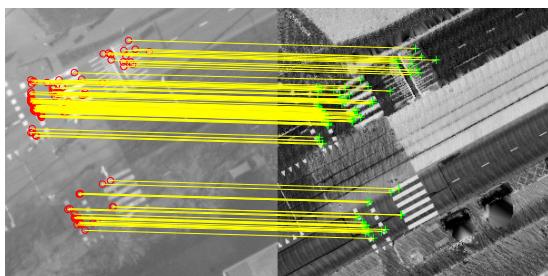


Figure 14. Homography based outliers filtering.

4. RESULTS

In this section we will describe the results according to the feature extraction steps.

4.1 Feature detection

An example result of the keypoint detection is provided in Figure 15. The two different zoom levels are provided to visualize the feature detection, one is a normal view (Figure 15, top) as well as a zoomed view (Figure 15, bottom). Overall, there are lots of clusters of keypoints detected at the corners of road markings. Many keypoints are also detected at the edges of a traffic light passing over the zebra crossing in the aerial image patch.

Comparatively, more keypoints are detected at the corners of the road markings in the aerial images than in the MLSPC ortho image. It can be noticed with a close observation that the correct corresponding keypoints are inside the clusters of keypoints in the both images. For detailed inspection, it is possible to further zoom-in and analyse the grey values and the detected keypoints (in digital copy).

For the keypoint detection from all tiles, the threshold and the total iterations required to obtain the keypoints are provided in Table 1. The extremely different requirements of threshold and the number of iterations required for the tile 2 and tile 6 shows the noise difference in the images of the same sensor.

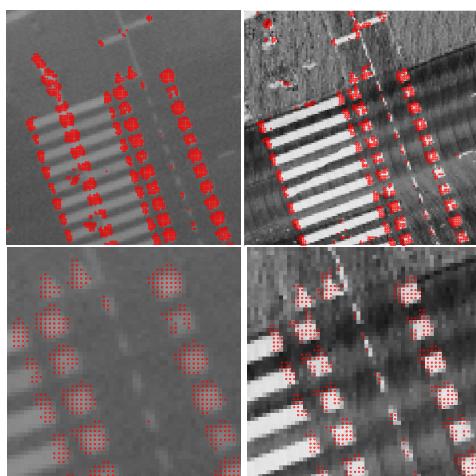


Figure 15. Adaptive Harris corner feature detection from aerial (left) and MLSPC (right) image patch¹.

Title no.	Iterations [MLSPC/Aerial image]	Threshold [MLSPC/Aerial image]	Keypoints [MLSPC/Aerial image]
1	782/85	7.8/0.83	5098/5091
2	1151/179	11.50/1.78	5096/5091
3	1250/346	12.49/3.45	5100/5100

¹ This image can be zoomed to 6400% for a very close observation by using the digital copy of this paper.

4	885/147	8.84/1.46	5095/5099
5	369/25	3.68/0.24	5090/5074
6	160/24	1.59/0.23	5088/5078
7	359/62	3.58/0.61	5092/5060
8	483/121	4.82/1.2	5091/5100
9	256/167	2.55/1.66	5094/5100
10	568/13	5.67/0.12	5097/5100
11	529/7	5.28/0.06	5097/5100
12	310/35	3.09/0.34	5088/5053
13	707/88	7.06/0.86	5098/5098
14	778/273	7.77/2.72	5098/5092

Table 1: Results of the adaptive Harris keypoint detection.

4.2 Feature matching²

In this section we will discuss the results of the developed filtering approach.

In first glance, matched features appear as random clusters of the keypoints detected in the image pair (Figure 16, top). A zoomed visualization of the matched keypoints with connecting lines is also crowded due to many matches (Figure 16, middle). Therefore, more interactive visualization of the results with Lego patterns is provided in (Figure 16, bottom). By a close observation of the Lego patterns, it can be realized that the patterns are exactly same in corresponding images. Which also shows that the features are matched to the pixel level accuracy. Furthermore, the grey values and edges in context of the road markings are also same with respect to each image. We also analysed the normal probability function of the error in translation of the obtained correspondence and the total number of matched correspondences together with mean (μ) and sigma (σ) of X and Y axis for all tiles in Table 2.

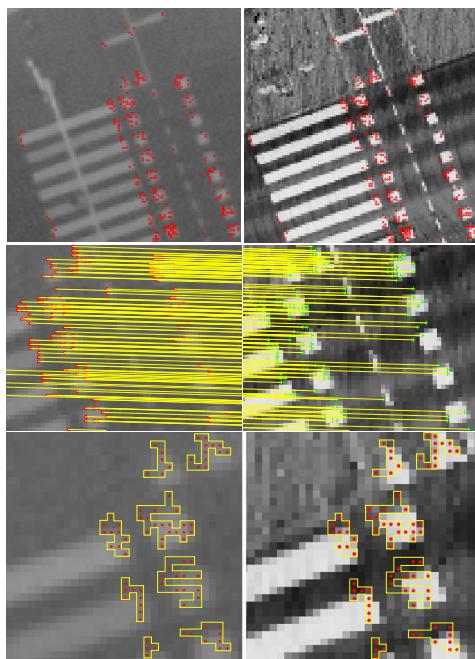


Figure 16. Different visualizations of the feature matching with developed method.¹

² Visualizations of all matching results are provided at:
https://www.researchgate.net/profile/Zille_Hussnain

Tile no.	Matches from descriptor matching	Matches from developed filtering tech.	$\mu_x(m)$	$\mu_y(m)$	$\sigma_x(m)$	$\sigma_y(m)$
1	4393	140	-2.16	0.87	0.03	0.08
2	7016	930	-2.07	0.95	0.06	0.06
3	12961	1703	-2.02	0.81	0.07	0.05
4	8159	702	-1.99	0.93	0.08	0.05
5	3702	275	-2.16	1.01	0.07	0.08
6	5090	316	-1.93	1.00	0.05	0.09
7	5318	456	-1.82	0.87	0.09	0.07
8	9403	495	-1.74	0.85	0.04	0.08
9	6050	438	-1.74	0.90	0.06	0.08
10	4109	65	-1.90	0.87	0.07	0.02
11	6043	748	-1.78	1.00	0.05	0.08
12	4849	185	-1.91	0.93	0.08	0.04
13	6460	561	-1.78	1.05	0.06	0.07
14	8824	750	-1.81	1.02	0.07	0.05

Table 2: Number of Matched keypoints and the normal probability function parameters of translation (all units are in meters) for each tile².

4.3 Discussions

The matching results have shown that the developed method has achieved the pixel level accuracy on the given data sets. The developed filtering approach can work only when the error in rotation and scale is small. When the error in rotation is large, then without the rotation check in the filtering part, only the distance constraint can be imposed. The resultant correspondences together with a small subset of outlier can be used to compute the accurate homography matrix and the outliers can be removed easily from the initial correspondence.

Even though the aerial imagery was captured in winter and the roads are comparatively well visible, the occlusion due to the trees branches was the biggest problem. We have noticed that the low point density and small occlusions in the point cloud do not have much effect on the results. So, the developed method can give better results, when the road surface is clearly visible in the aerial imagery.

Overall, distortion in the aerial imagery due to the atmospheric dispersion, different level of contrast/illumination and variable exposure could hinders the keypoint detection, however, these problems were handled by the adaptive Harris keypoint detection. The false keypoint detection caused by shadows, trees and poles, and matching was handled by the outlier filtering approach.

4.4 Evaluation of estimated shift

The actual error in the MLSPC relative to the aerial imagery could be measured roughly by handpicked correspondences. The evaluation method only considers the error in translation and does not consider the error in rotation. Though not very accurate, still this method can provide a rough evaluation of the error presents in the MLSPC.

An example of manual selection of corner point is shown in Figure 17. In this figure, it has been shown that the top

left corner point of a road marking (bounded by green window) is selected as a same corresponding point in both images. This process was repeated for each tile and few well distributed points were obtained. The distribution of the obtained error is shown in the Figure 18. As it is not possible to show the results from all tiles, only accurate and least accurate results from two different tiles are shown in Figure 18, top and bottom respectively.



Figure 17. Manual selection of road marking's corner point for result evaluation.

5. CONCLUSION

In this paper we have implemented an automatic and reliable feature extraction procedure for the MLSPC and aerial ortho images, which can compute correspondences with up to pixel level accuracy. The feature matching results have shown that the MLSPC can be corrected reliably to the pixel level accuracy.

The filtering technique proposed is feasible for the real world problem of MLSPC correction. The developed filtering technique cannot compensate for the scale variations and large rotations. Moreover, the filtering technique is a brute-force method as it is developed for the post processing purpose and it cannot be used for the real time applications. The advanced searching algorithms will be introduced to decrease the total time required for feature matching and to compensate rotation and scale variations.

Image filtering methods can be used to save the time spent on computation of feature detection parameters, it will avoid the need to adapt the parameters to images from same sensor. Furthermore, the evaluation of the matching result is a rough estimation of the error because manual correspondence selection is error prone. A more accurate evaluation method will be used in future.

ACKNOWLEDGEMENTS

This project is funded and supported by the Dutch Technology Foundation STW, which is part of Netherlands Organisation for Scientific Research (NWO), and which is partly funded by the Ministry of Economic Affairs. We would also like to thank Topcon® for providing us the MLSPC and Cyclomedia® for providing the Aerial imagery.

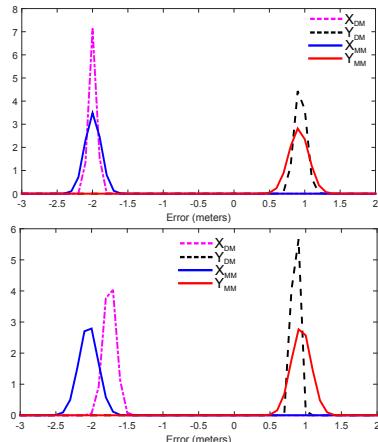


Figure 18. Overlap of PDFs of X and Y coordinates, estimated by the developed method (DM) and manually measured (MM), tile 4 (top) and tile 8 (bottom).

REFERENCES

- Abedini, A., M. Hahn and F. Samadzadegana, 2008. An investigation into the registration of LIDAR intensity data and aerial images using the SIFT approach. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Vol. 2, p. 6.
- Brown, M., G. Hua and S. Winder, 2011. Discriminative learning of local image descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 33(1), pp. 43-57.
- Frueh, C. and A. Zakhor, 2003. Constructing 3d city models by merging ground-based and airborne views. IEEE Conference on Computer Vision and Pattern Recognition, Vol. 2, pp. 562-569.
- Gao, Y., X. Huang, F. Zhang, Z. Fu and C. Yang, 2015. Automatic Geo-referencing Mobile Laser Scanning Data to UAV images. The International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 40(1), p. 41.
- Harris, C. and M. Stephens, 1988. A combined corner and edge detector. Alvey vision conference, Vol. 15, p. 50.
- Jende, P., Z. Hussnain, M. Peter, S. Oude Elberink, M. Gerke and G. Vosselman, 2016. Low-Level Tie Feature Extraction of Mobile Mapping Data (mls/images) and Aerial Imagery. Int. Arch. of Photogramm. and Remote Sens. pp. 19-26.
- Kümmerle, R., B. Steder, C. Dornhege, A. Kleiner, G. Grisetti and W. Burgard, 2011. Large scale graph-based SLAM using aerial images as prior information. Autonomous Robots, 30(1), pp. 25-39.
- Levi, G. and T. Hassner, 2015. LATCH: Learned Arrangements of Three Patch Codes. arXiv preprint :1501.03719.
- Lowe, D. G., 2004. Distinctive image features from scale-invariant keypoints. International journal of computer vision, 60(2), pp. 91-110.