EARTH OBSERVATION AND GEOMATICS ENGINEERING
website: https://eoge.ut.ac.ir

A Comparative Assessment of Selected Algorithms on Multi Domain Registration of Remote Sensing Images

Fatemeh Alidoost, Hossein Arefi*

School of Surveying and Geospatial Engineering, College of Engineering, University of Tehran, Tehran, Iran

Article history:
Received: 10 May 2018, Received in revised form: 20 October 2018, Accepted: 25 October 2018

ABSTRACT

With the development of remotely sensed data acquisition techniques, the integration of complementary data has found a key role in many applications. The prerequisite of the data integration is the data registration. There are many approaches for image registration that they can be categorized based on the nature of data sources as the multi domain and single domain methods. The multi domain methods employ the data sources with different properties of gray values such as irradiance, the return strength of the laser pulse, and the height values. The different aspects of multi domain registration methods include the matching algorithm as feature-based or intensity-based, the level of automation, availability of initial registration parameters, the implementation cost, and so on. In this study, a number of multi domain registration algorithms are selected and compared based on the mentioned aspects and then, an automatic multi domain registration method is proposed to register an intensity image of LiDAR and a satellite image without using initial registration parameters. For this, the combination of the Scale Invariant Feature Transform (SIFT) detector and the mutual information theory is employed to use the strengths of both feature-based and intensity-based matching algorithms to decrease the difficulty of multi domain image registration. The final registration results prove that the combination of the feature-based and the intensity-based matching algorithms could be an efficient solution of multi domain image registration problem, especially for registration between the LiDAR intensity and the satellite images with the RMSE of one pixel.

KEYWORDS

Multi domain registration
Image matching
Mutual information
SIFT
LiDAR

1. Introduction

Image registration is a fundamental task to align different data sources in order to obtain the comprehensive information in many applications such as photo-realistic 3D urban models (Kim et al., 2014), medical imaging (Goshtasby, 2012), 3D reconstruction, localization and navigation, and so on. The laser scanning and optical imaging technologies are two widely used data sources for 3D data acquisition. Aerial or terrestrial laser scanners record 3D coordinates and intensity of objects. Aerial or satellite imageries, on the other hand, provide some gray-scale images in a 2D plane. Each of these technologies has their own advantages and disadvantages (Fei et al., 2008; González Aguilera et al., 2007; Mishra & Zhang, 2012). For example, the optical sensors acquire intensity images with high resolution and high information content but only in a 2D space and therefore, the extraction of 3D data from these images are usually difficult, which is highly time and cost consuming. On the other hand, 3D coordinates of points are recorded very fast and accurate using the laser scanners, but without photometric information of objects. Due to the weaknesses and strengths of each individual data, the combination of 2D and 3D data is utilized as complementary information in many applications. This is often implemented using different image registration techniques. The image

* Corresponding author
E-mail addresses: falidoost@ut.ac.ir (F. Alidoost); hossein.arefi@ut.ac.ir (H. Arefi)
DOI: 10.22059/eoge.2019.249453.1019

matching is the main stage in image registration that principally searches for similarities and common characteristics between datasets.

Many methods have been developed to solve the image registration problem based on matching algorithms, but most of them are applicable to the identical data sources (e.g. optical to optical image data) and do not employ to the multi domain images such as the LiDAR and optical images. This is because of inherent differences between the multi domain images, such as the different radiometric natures that cause both the digital number and even the local gradient vectors to be inappropriate for matching or generating a transformation model between two images, as well as different geometries of acquisition systems in the LiDAR sensors (orthogonal mapping) and the optical sensors (central mapping), different levels of radiometric details, and different temporal resolutions (Mastin et al., 2009; Mishra & Zhang, 2012; Palenichka & Zaremba, 2010; Shorter et al., 2008). Moreover, most of registration algorithms require additional information or initial parameters to solve multi domain registration problems (Lee et al., 2011; Meierhold et al., 2010; Moussa et al., 2012; Parmehr et al., 2013). In this study, first a selected number of known multi domain image registration algorithms are reviewed based on different aspects, such as the selected applicable algorithm, level of automation, availability of initial registration parameters, implementation cost, and so on. Then, an automatic multi domain method using the combination of feature-based and intensity-based matching techniques is proposed to register a satellite image with a LiDAR intensity image as well as the digital elevation models derived from both datasets without using initial and/or additional registration parameters.

2. Image Registration Concept

The image-based registration methods usually consist of the following three steps (Goshtasby, 2012; Mishra & Zhang, 2012; Zitová & Flusser, 2003):

2.1. Pattern extraction

In order to extract and highlight informative and non-redundant information about the scene structures in optical images, some key patterns are extracted as points, corners, lines or edges, as well as areas or polygons. This step is critical for further processing and can affect the final result of registration.

2.2. Pattern matching

In order to align the images spatially, the corresponding patterns should be matched based on matching measures. A general categorization of matching measures includes mathematical, statistical, binary distances and information theory-based matching measures (Cyganek & Siebert, 2009).

2.3. Transformation model estimation

A rigid transformation model is utilized to estimate the geometric relations between the images using the corresponding patterns. Common transformation models are translation, rigid, similarity, affine, projective, cylindrical, and spherical models (Goshtasby, 2012).

3. Image Registration Techniques

The registration techniques can be categorized into two main groups based on the types of extracted patterns as: feature-based and intensity-based methods (Hofmann et al., 2014; Mishra & Zhang, 2012). Each technique can be divided into other groups as shown in Figure 1, which are explained in more details in the forthcoming subsections.

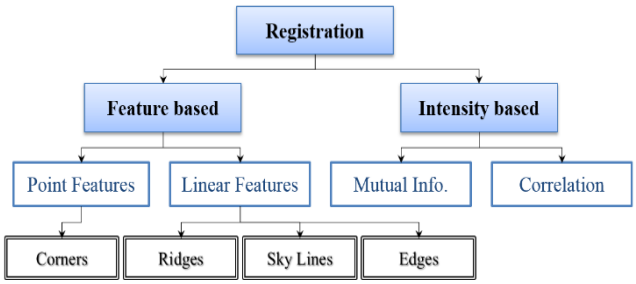


Figure 1. Image registration techniques

3.1. Feature-based registration methods

It is possible to identify some sparse features in image data by local or global descriptors. These features should be distinctive and robust to occlusion (Mikolajczyk & Schmid, 2005). The common local invariant features are linear and point features such as the boundary of building roofs, straight lines, corners, and edges. According to the type of local features, feature-based registration methods can be divided into four sub-groups as: the feature-based registration using corners, edges, skylines, and ridges (Arefi & Hahn, 2005; Rönnholm et al., 2013).

3.1.1. Feature-based registration using corners

The corner points are important features that are widely used in many applications such as the image analysis and registration. In multi domain registration tasks, the detected points should be robust and invariant to color, intensity, noise, blurring, illumination, viewing direction, and geometric changes (Goshtasby, 2012; Kim et al., 2014). The most common of recent corner detectors are Scale Invariant Feature Transform (SIFT) (Lowe, 2004), Speeded Up Robust Feature (SURF) (Bay, Tuytelaars, & Van Gool, 2006) and SIFT derivatives (Mikolajczyk & Schmid, 2005; Schmitt & McCoy, 2011). In addition, a comparison of different detectors has been reported (Kenney et al., 2005; Patel & Panchal, 2014; Rodehorst & Koschan, 2006).

3.1.2. Feature-based registration using edges

The boundary information of image regions can be used as the edge features in the image matching which are robust to the radiometric differences between images (Lee et al., 2011). Some edge detection algorithms extract the points in the image based on the gradient magnitude of points such as Canny, Sobel, SUSAN, and Harris & Stephen edge detectors.

3.1.3. Feature-based registration using skylines

A skyline is a curve or contour between the sky and other man-made objects in urban environments that it is located at the upper parts of the image and is stable to different orientations (Hofmann et al., 2014; Lie et al., 2005; Zhu et al., 2012). The skylines can be extracted much more easily by applying an adaptive threshold on the intensity gradient and using a morphologic filter (Zhu et al., 2012).

3.1.4. Feature-based registration using ridges

In urban environments, the ridges or building roofs can be used as unique linear features to search the similarity between two images. The computational cost of ridge features is usually higher than other features. Most of the ridge detector algorithms extract the ridge lines of a roof by intersecting two non-parallel roof planes (Arefi & Hahn, 2005).

3.2. Intensity-based registration methods

The classical algorithm of intensity-based methods is the Normalized Cross Correlation (NCC) and its modification (Pratt, 2001) such as Zero mean NCC (ZNCC) (Lhuillier & Quan, 2000). Recently, new intensity-based matching algorithms have been developed based on the information theory. The algorithms calculate the statistical relationships between datasets by means of the entropy concept (Cover & Thomas, 2006). The relative entropy can be calculated as a similarity measure between the gray level values of images in image matching. One of the strong measures of similarity is the Mutual Information which is based on the entropy concept (MI) (Cyganek & Siebert, 2009). Based on many researches and studies, the MI-based matching algorithms are stable and more efficient for multi domain image data (LeMoigne et al., 2012; Mishra & Zhang, 2012).

4. Literature Review of Multi Domain Image Registration Algorithms

There are few studies on 2D-3D registration between an optical image and the LiDAR point cloud, which could be categorized into two general groups as feature-based and intensity-based techniques.

In feature-based techniques, different methods are proposed according to the main application (e.g. urban or forestry areas) as well as the type of features (e.g. points, corners, lines and edges) such as the line-based registration method (Cui et al., 2016; Elaksher, 2016; Lee et al., 2011; Liu et al., 2016; Liu et al., 2016), the SIFT-based methods (Meierhold et al., 2010; Moussa et al., 2012; Sattler et al., 2011; Yang et al., 2016), the SURF-based methods (Kim et al., 2014; Kim et al., 2016), morphology-based methods in a forested area (Lee et al., 2016), the geometrical constraints (Zhang et al., 2015), edge-based detectors (Rönnholm et al., 2013), the salient features from the DSMs (Abayowa et al., 2015), the Harris-based method (Budge et al., 2014), and the non-parametric registration techniques (Lee et al., 2015).

The registration problem between an RGB image and an intensity image of LiDAR is commonly solved by employing the SIFT operator as well as its different variations, such as the affine SIFT detector (Moussa et al. 2012) for feature extraction and the RANSAC algorithm for finding the correspondences. The SIFT-based approaches are able to solve the registration of large datasets in fast and easy framework with high accuracy and can handle different characteristics of multi domain images with a high reliability. However, the SIFT algorithm should be initialized manually (Meierhold et al., 2010; Sattler et al., 2011). Instead of SIFT, some researches focused on employing the SURF detector to extract the 2D-3D correspondences and proposed a two-pass RANSAC scheme with the Maximum Likelihood estimation (MLESAC) to register the depth images and optical stereo images (Kim et al., 2014). In order to solve the relative orientation between an aerial image and the LiDAR data, the lines and edges could be extracted by the edge detectors such as the Canny detector (Lee et al., 2011; Rönnholm et al., 2013) or based on the skyline (Hofmann et al., 2014) extracted by thresholding on the optical image histogram by Otsu's method (Otsu, 1979). These methods can demonstrate an automatic and fast registration. However, providing enough initial exterior orientation parameters is necessary for a successful search process. Moreover, the lines can be extracted only in urban environments, and these features are not suitable for matching the acquired images from natural environments.

Most of intensity-based techniques are based on employing the MI as well as different variations of it, such as the Normalized Combined Mutual Information (NCMI) to find the similarity between the intensity and the range images (Parmehr et al., 2013). However, the MI-based techniques are applicable to search through images with only the translation, and a coarse registration is necessary before its application as well. In order to improve the robustness as well as to increase the pixel/sub pixel accuracy of intensity-based techniques, a hybrid registration method could be employed including both the intensity- and feature-based detectors such as the Log-Polar Fast Fourier Transform

(LPFFT), the Harris corner points and the Probability Density Function (PDF) descriptor (Toth et al., 2011).

A comparison between multi domain image registration algorithms can be performed based on different aspects of reviewed methods. One of the most important aspects of the multi domain image registration algorithms is that the whole process can be performed automatically and without using the initial registration parameters such as the sensor model, the pose parameters, the exterior orientation parameters, and the coarse registration parameters. Furthermore, due to the limitation in time and computational volume in many applications, the registration algorithm should be easy and without any requirement to initialize many parameters for having an efficient matching. On the other hand, a developed registration algorithm should be applicable for different environments especially both of the urban and natural areas. Moreover, most of current developed algorithms need two depth maps that one of them is extracted from the stereo images, but in some cases, there is just one image without any additional information. Based on these comparison parameters, an overall score can be considered for each reviewed algorithm as shown in Table 1. As shown in Table 1, the feature-based registration algorithms are more appropriate than the linear feature-based algorithms for the multi domain image registration. It may be also concluded that the Fourier-based methods have better results to register the multi domain images (Toth et al., 2011). In the proposed method by Meierhold et al. (2010), a powerful SIFT algorithm is employed, which need to initialize at least three input parameters (Meierhold et al., 2010). A novel intensity-based registration algorithm was developed by Parmehr et al. (2013) that was efficient in the multi domain image matching, but the number of bins can affect the overall success of the algorithm (Parmehr et al., 2013). Moreover, in multi domain registration problems, the MI technique does not have a good performance without a coarse registration. The generation of a point-based environment model is not problem is solved based on an efficient and fast point matching, but in such methods, the number of utilized images should be high. An easy registration algorithm was developed by Lee et al. (2011) in which used the sensor model of the image to perform a coarse registration. Also, some crucial parameters such as the image chips size should be defined carefully based on the quality and resolution of the images. In the latest two methods proposed by Hofmann et al. (2014) and Rönnholm et al. (2013) the linear features are used to solve the image registration problem in urban area. The main disadvantage of these methods is that the exterior or relative orientation between the point cloud and the image data should be determined manually. Therefore, some known registration parameters should be applied on the point cloud before the matching process.

As a result of the comparison, some developed algorithms use many complicated multi steps processes in matching the features from multi domain images to solve the registration problem. Also, some methods claim to be automatic; however, in the first step, the manual registration is used. Moreover, in order to have enough accuracy and robustness, some reviewed algorithms need an initial information such as the camera calibration parameters, the position and attitude data, or the relative orientation between two sensors.

5. Proposed Method

In this study, an efficient multi domain image registration method is proposed to register the satellite and LiDAR intensity images as well as to register the digital elevation models, which are created from LiDAR and image matching techniques. This method is based on a combination of feature-based (e.g. SIFT) and intensity-based (e.g. MI theory) algorithms to consider the different aspects of registration for multi domain images such as efficiency, robustness, independency of operator or auxiliary information, and applicability in multi sensors application.

Table 1. A comparison of reviewed registration algorithms

Factor / Algorithm	Feature-based – Corner (Toth et al., 2011)	Feature based – Point (Kim et al., 2014)	Feature-based – Corner (Meierhold & Spehr, 2010)	Intensity-based – MI (Parmehr et al., 2013)	Feature-based – Point (Moussa et al., 2012)	Feature-based – Point (Sattler et al., 2011)	Feature-based – Edge (Lee et al., 2011)	Feature-based – Skyline (Hofmann et al., 2014)	Feature-based – Ridge (Rönnholm et al., 2013)
Automatic algorithm?	1	1	1	1	1	1	1	0	0
No need to additional (pose) information or coarse registration?	1	1	1	0	1	1	0	0	0
No need to initializing parameters of algorithm?	1	1	0	0	0	1	0	1	1
Easy implementation?	1	1	1	1	0	0	1	1	0
Work with single frame?	1	0	1	1	1	0	1	1	1
Unlimited application (for both urban & natural areas)?	1	1	1	1	1	1	1	0	0
Sum of scores	6	5	5	4	4	4	4	3	2

The main steps of proposed method is shown in Figure 2.

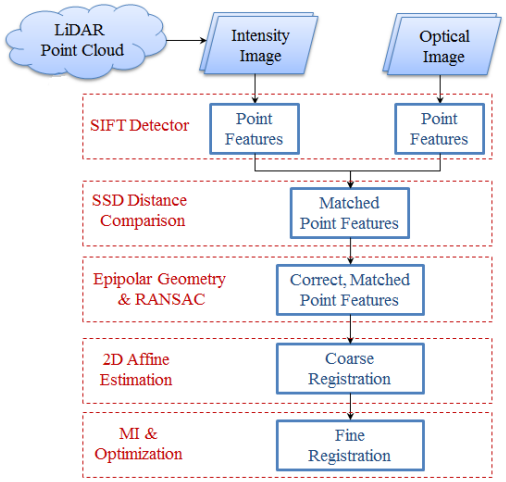


Figure 2. Main flowchart of registration algorithm

First, the SIFT detector (Lowe, 2004) is employed to detect points in each dataset. For the experiments, the SIFT function of the VLFeat library (Vedaldi & Fulkerson, 2010) was used. The SIFT algorithm consists of mainly four steps as scale-space extrema detection, keypoint localization, orientation assignment and keypoint descriptor.

At first step, a scale-space filtering method such as the Difference of Gaussians (DoG) function is used to detect the blobs whose location and scales are identifiable from different views of the same object. Difference of Gaussian, $D(x, y, \sigma)$, is obtained as the difference of Gaussian blurring of an image with two different scaling parameter as σ and $k\sigma$ given by Eq. (1).

$$D(x, y, \sigma) = G(x, y, k\sigma)I(x, y) - G(x, y, \sigma)I(x, y) \quad (1)$$

where [*] is the convolution operator, $G(x, y, \sigma)$ is a variable-scale Gaussian and $I(x, y)$ is the input image. The filter is applied on different octaves of the image in Gaussian pyramid. To detect the exterma, each pixel is compared with its 8 neighbours as well as 9 pixels in the next scale and 9 pixels in the previous scales (Lowe, 2004).

In the second step, any low-contrast keypoints and edge keypoints are removed using two thresholds called the peak selection threshold and the non-edge selection threshold (Lowe, 2004). In the third step, the orientation for each keypoint is assigned by calculating the gradient magnitude and direction in the neighborhood. Then, an orientation histogram is generated and the highest peak of this histogram is considered to calculate the orientation. In the final step, using a set of 16 histograms, aligned in a 4×4 grid, each with 8 orientation bins, a feature vector containing 128 elements is created for a 16×16 pixels in the neighborhood. This feature vector is called the keypoint descriptor (Lowe, 2004). The most important parameters of the SIFT algorithm are the number of levels per octave of the DoG scale space, the peak selection threshold and the non-edge selection threshold. In order to find the matched keypoint between two images, the corresponding SIFT feature vectors between the two images should be matched based on a matching measure. The most common and practical matching measure that can be used to define which areas in different images fit together is the sum of squared differences (SSD) function given by Eq. (2) (Cyganek & Siebert, 2009).

$$D_{SSD} = \sum_{(i,j)\in U} (I_1(x+i, y+j) - I_2(x+d_x+i, y+d_y+j))^2 \quad (2)$$

where, the image region of I_1 is built around a reference point (x, y) in its local coordinate space, and the image region of I_2 is built around a point $(x + dx, y + dy)$ in its local coordinate space. It is assumed that two compatible image regions I_1 and I_2 are compared. For both, the matching regions are defined by a set U of offset values (i, j), measured from their reference points i.e. (x, y) and $(x + dx, y + dy)$, respectively. It is assumed also that all indices defined by U fall into ranges of a valid pixel location for I_1 and I_2, respectively (Cyganek & Siebert, 2009). In the matching problem, the pixel values are the calculated SIFT feature vectors for keypoints located at (x, y) and $(x + dx, y + dy)$. If two SIFT feature vectors are matched, the corresponding SSD measure has a small value. The matching threshold represents a percent of the distance from a perfect match in the range of (0,100]. Two SIFT feature vectors match when the SSD between them is less than the predefined matching threshold. In the third step, an epipolar geometry constraint with RANdom SAmple Consensus (RANSAC) (Fischler & Bolles, 1981) is employed to find the correct correspondences and exclude the outliers. For this, the fundamental matrix (F) is calculated using the Normalized Eight-Point algorithm (Longuet-Higgins, 1981) between corresponding points (p_1 and p_2) as in Eq. (3) (Meierhold et al., 2010).

$$p^T{}_2 F p_1 = 0 \quad (3)$$

Since for calculating the F matrix, the input points should not contain any outliers, the robust RANSAC algorithm (Fischler & Bolles, 1981; González-Aguilera et al., 2009; Meierhold et al., 2010) is used to eliminate outliers. In feature-based registration step, a 2D affine transformation is applied on the corresponding points and six parameters of transformation are calculated by Eq. (4) (Goshtasby, 2012).

$$\begin{aligned} X &= a_1 x + a_2 y + a_3 \\ Y &= a_4 x + a_5 y + a_6 \end{aligned} \quad (4)$$

where (x, y) and (X, Y) are two sets of matched, correct points from the previous step and $a_1 \ldots a_6$ are six parameters of affine transformation. Since the differences between two datasets cannot be described based on only six affine parameters, another the registration step is needed as an intensity-based registration, which is implemented based on the MI theory. MI is a relative entropy between the joint

distribution $P(x, y)$, and the product of distributions $P(x)P(y)$, as Eq. (5) (Cyganek & Siebert, 2009).

$$I(X, Y) = \sum_{x, y \in A} P(x, y) log \frac{P(x, y)}{P(x)P(y)} \quad (5)$$

where the entropy of two random variables X and Y is given by Eq. (6) (Cyganek & Siebert, 2009).

$$H(X \mid Y) = - \sum_{x, y \in A} P(x, y) log P(x, y) \quad (6)$$

The intensity-based registration is an iterative optimization process and the MI metric is used to measure the image similarity for evaluating the accuracy of the registration. The parameters of optimization are the maximum iteration, the number of spatial samples and the number of histogram bins which are used to compute the metric. The affine matrix, from feature-based registration step, is applied on the target image using a bilinear interpolation to transform the target image into the coordinate system of the reference image. Next, the MI metric is calculated between the transformed target image and the reference image. Finally, the optimizer checks if there is a stop condition. The number of maximum iteration should be large enough to reach the acceptable accuracy in registration.

6. Implementation and Results

In this paper, a multi domain image registration is implemented using three datasets (as described in Table 2) over a city area (Terrassa) in Catalonia, Spain (ISPRS, 2015). The satellite images consist of two stereo WorldView-1 images and two stereo Cartosat P5 images with corresponding Rational Parameter Coefficients (RPCs) information. The RPCs have been corrected with ground control points, reference DSM and tie points in a bundle adjustment. They have a good relative and absolute orientation, and should be used for stereo matching and DEM generation without any further correction. Tie point RMSE is below 0.1 pixels for Cartosat-1 and 0.2 pixel for Worldview-1 (ISPRS, 2015). The LiDAR point cloud has a density of 0.5 pt/m^2 and includes an intensity image over the area of interested.

Table 2. Characteristics of datasets

Types	Datasets	Spatial Resolution
Point clouds and intensity image	Airborne laser scanner (LiDAR)	0.5 pt/m^2
Stereo images with corrected RPC	Cartosat P5	2.5 m
	WorldView-1	0.5 m

For DEM registration, three DSMs are generated from LiDAR dataset, P5 and WV1 stereo images based on Kim's method (Kim et al., 2014) with the spatial resolution of 1m, 2.5m and 0.5m, respectively. As shown in Figure 2, in first step, the feature vectors are extracted from each dataset using a SIFT detector. Since the radiometric nature, the image quality, the spatial and temporal resolution of each dataset are different, the different parameters should be chosen for SIFT. The values of the number of levels, peak threshold and edge threshold are about 4, 1 and 25 for intensity image and 6, 4 and 2 for satellite images. These values are 1, 1 and 50 for DEMs. The average number of detected points from each dataset is about 4000 points to ensure to have enough points for the Normalized Eight-Point algorithm in the third step. The matching percent threshold is 100 to keep the points, which are matched 100% based on SDD distance. After the outlier removal in the third step, the 2D affine transformations are calculated between six pairs as LiDAR and P5 images, LiDAR and WV1 images, WV1 and P5 images, LiDAR and P5 DEMs, LiDAR and WV1 DEMs, and WV1 and P5 DEMs. As the final step of implementation, the MI-based registration is applied on each mentioned pairs. The results of registration are shown in Figures 5-10.

Based on the visual assessment, if the radiometric natures of two image datasets are the same, more corresponding points could be extracted by the SIFT algorithm and matched correctly. For example, the number of matched points between the LiDAR intensity image and P5 or WV1 images is much lower than the other datasets. In these cases, the good initializing of SIFT parameters is very important to find enough points for the registration. Another result is that the MI algorithm can just tolerate a few outliers in the point sets and if the number of mismatch points is very large, MI fails in registration. Since the ground control points are not available for this case study, a series of corresponding points were manually extracted from both the images and DEMs and were used for a relative quantitative evaluation. These control points, which were tried to be uniformly distributed through the region, were generally selected in obvious corners for more reliable positioning. The number of control points in images and DEMs are 13 (Figure 3) and 5 (Figure 4), respectively (the lower number of control points is because of difficult positioning in DEMs) and the numerical results of horizontal accuracy between different datasets are reported in Table 3 and Table 4.

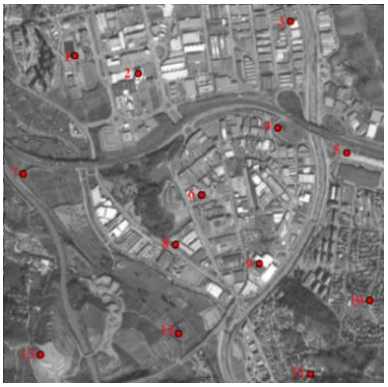


Figure 3. Check points' locations for image registration

Table 3. Errors of image registration in pixel

Datasets	Min.	Max.	Ave.	RMSE
P5 & LiDAR Intensity Images	0.11	3.61	1.43	0.81
WV1 & LiDAR Intensity Images	0.11	2.33	0.98	0.60
WV1 & P5 Images	0.08	1.07	0.66	0.39

Based on Table 3, due to very different characteristics of the keypoints in images with different radiometric natures and spatial resolutions, the MI algorithm was unable to deliver acceptable results and accordingly, a coarse registration is necessary before MI registration. For example, the RMSE of registration between P5 image and LiDAR intensity is higher than that of the others and is about 0.81 pixel, while the registration error between P5 and WV1 images is 0.39 pixel.

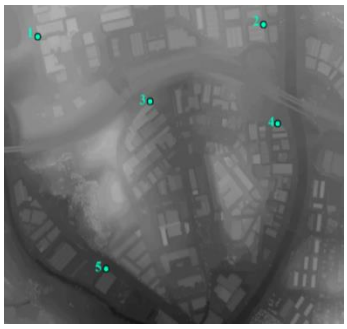


Figure 4. Check points' locations for DEM registration

Table 4. Errors of image registration in pixel

Datasets	Min.	Max.	Ave.	RMSE
P5 & LiDAR DSMs	0.04	5.37	2.02	1.88
WV1 & LiDAR DSMs	0.08	3.05	1.22	0.92
WV1 & P5 DSMs	1.64	8.5	3.4	2.56

The generated DEM from LiDAR point cloud is more accurate than those from satellite stereo images. Therefore, the RMSE of registration between a LiDAR-based DEM and a satellite-based DEM is lower than the average error of registration between two satellite-based DEMs, as shown in Table 4. For instance, the RMSE of registration between generated DEMs from WV1 and P5 stereo images is about 2.56 pixels, while these values are 1.88 pixels and 0.92 pixel between the LiDAR-based DEM and the P5-based DEM and the LiDAR-based DEM and the WV1-based DEM, respectively. As shown in Table 4, the errors of DEM registration is seen to be larger than those of image registration. It is because that the exact pointing on objects in DEMs is not possible and the accuracy of control point selection by the operator to find individual conjugate points between DEMs is decreased significantly. As another result, the overall success of registration is very much dependent to the matched correct points with the enough number and good distribution. On the other hand, the power of SIFT to extract points and the reliability of MI to search the correspondences can be used together to solve the multi domain registration problem with the average RMSE of one pixels.

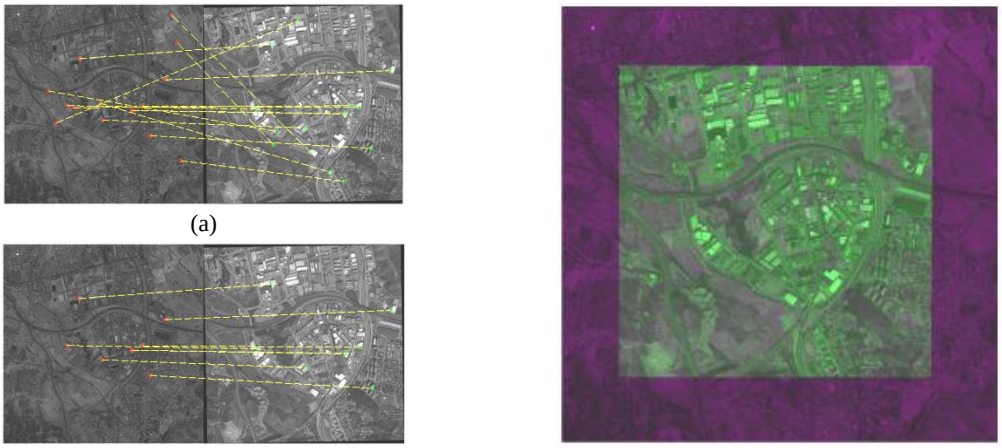
There are a few recent studies on registration between the LiDAR data and satellite/aerial images, whose accuracy are reported in Table 5. Comparing with Table 3, we can find that most of the errors for our experiments are similar or smaller than those of Table 5 in case of registration between the LiDAR data and satellite images. However, since different datasets are used in other studies, the complete comparison is not possible. In addition to accuracy, efficiency is also very important for registration in practice. The proposed method in this study is an automatic technique without needing the auxiliary information about data.

Table 5. Errors, provided by other studies, in pixel

Author	Method	RMSE
(Elaksher, 2016)	linear-based features	1 - 1.5
(Safdarinezhad et al., 2016)	shadow extraction	1
(Shijie Liu et al., 2016)	linear-based features and affine transformation	0.8
(Liu et al., 2018)	linear and planar features	1.66
Ours	SIFT and MI	0.6 - 0.8 0.9 - 1.8

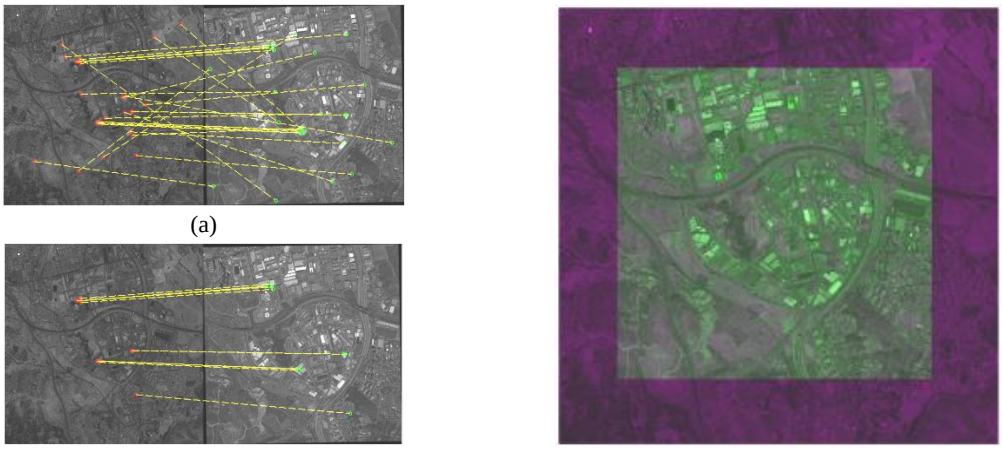
7. Conclusion

In this study, as the first step, some of the multi domain registration algorithms were reviewed in details and compared based on influencing parameters of the registration problem such as the reliability and effectiveness of the algorithm in different conditions such as multi domain and multi sensor data, the low quality data, the existence of the blunder and noise in data, and unavailability of initial parameters. According to the results of this comparison, a hybrid registration technique based on the SIFT and MI was proposed for multi domain images such as the single optical image and the intensity of LiDAR. Because of the different nature of these data, defining the appropriate parameters for the SIFT and an optimization process to find the matched points with maximum values of the MI is very important and an acceptable accuracy is not achievable without the good initializing. Based on the results, the average RMSE of horizontal errors are about 0.6 pixel and 1.7 pixels for image registration and DEM registration, respectively. In the comparison with the recent methods for registration between LiDAR data and satellite/aerial images, the proposed method is automatic without using the initial registration parameters such as the sensor model, the pose parameters, and the exterior orientation parameters. Also, it is applicable for both of urban and natural areas. However, the extraction of some auxiliary images from the original images such as gradient images, orientation images, and frequency-domain images with higher information content would be helpful for the matching algorithms.



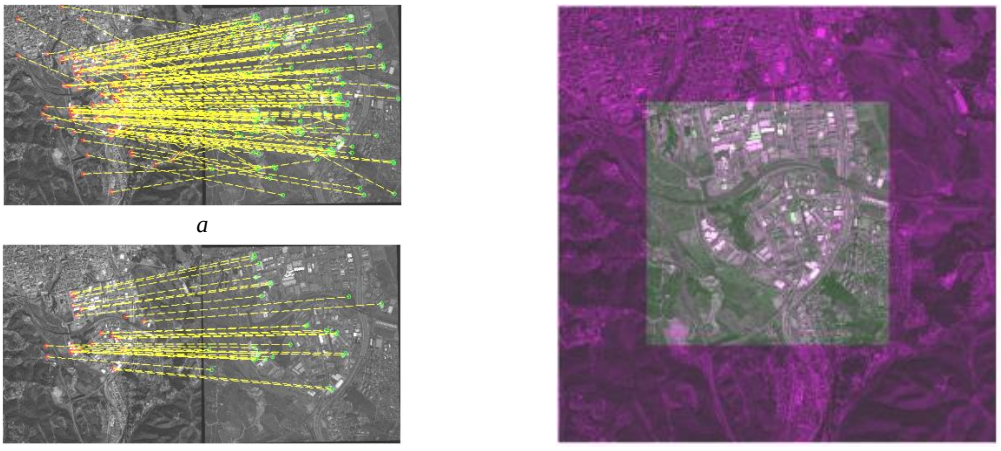
(a) (b) (c)

Figure 5. Results of the P5 image and the LiDAR intensity image registration, (a) the detected points, (b) the matched points, (c) the final result of registration



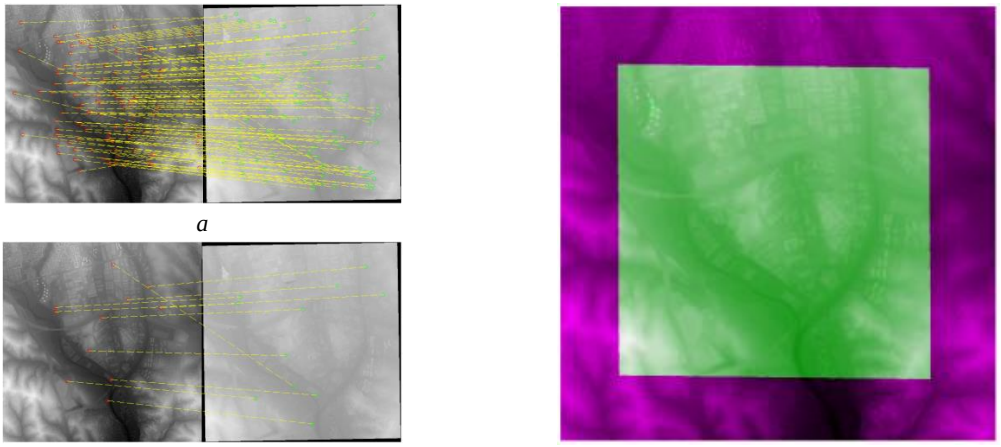
(a) (b) (c)

Figure 6. Results of the WV1 image and the LiDAR intensity image registration, (a) the detected points, (b) the matched points, (c) the final result of registration



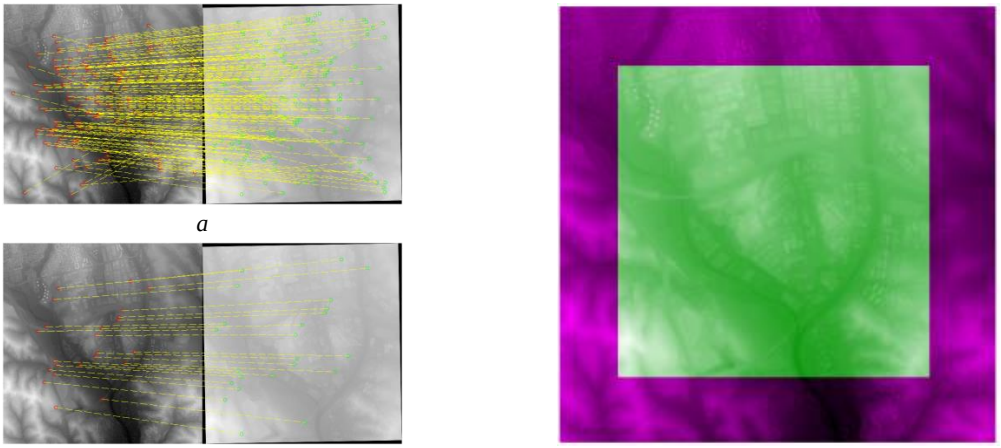
a *b* *c*

Figure 7. Results of the WV1 image and the P5 image registration, (a) the detected points, (b) the matched points, (c) the final result of registration



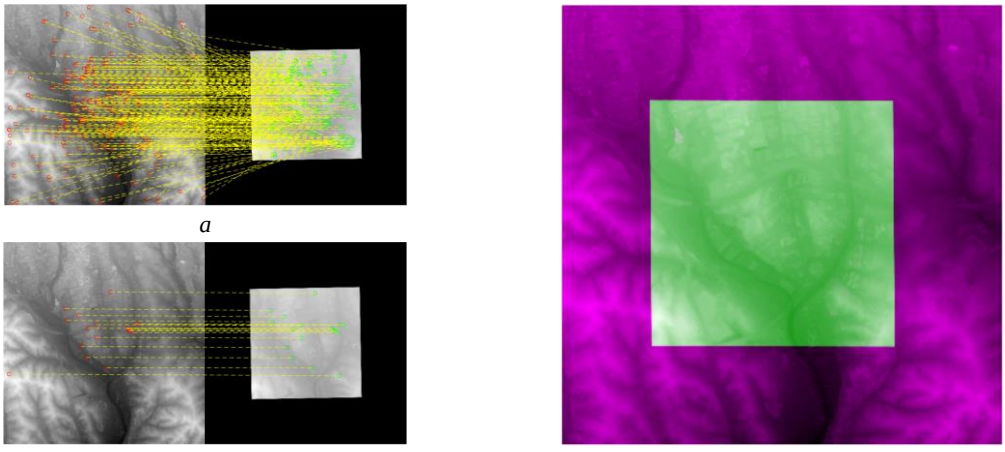
a *b* *c*

Figure 8. Results of the P5 DEM and the LiDAR DEM registration, (a) the detected points, (b) the matched points, (c) the final result of registration



a *b* *c*

Figure 9. Results of the WV1 DEM and the LiDAR DEM registration, (a) the detected points, (b) the matched points, (c) the final result of registration



a *b* *c*

Figure 10. Results of the WV1 DEM and the P5 DEM registration, (a) the detected points, (b) the matched points, (c) the final result of registration

References

Abayowa, B. O., Yilmaz, A., & Hardie, R. C. (2015). Automatic registration of optical aerial imagery to a LiDAR point cloud for generation of city models. *ISPRS Journal of Photogrammetry and Remote Sensing*, *106*, 68-81. doi:10.1016/j.isprsjprs.2015.05.006

Arefi, H., & Hahn, M. (2005). A morphological reconstruction algorithm for separating off-terrain points from terrain points in laser scanning data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36(3/W19), 120-125.

Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In European conference on computer vision (pp. 404-417). Springer, Berlin, Heidelberg.

Budge, S. E., Badamikar, N. S., & Xie, X. (2014). Automatic registration of fused lidar/digital imagery (texel images) for three-dimensional image creation. Optical Engineering, 54(3), 031105.

Cover, T. M., & Thomas, J. A. (2006). Elements of Information Theory. Hoboken, New Jersey: John Wiley & Sons, Inc.

Cui, T., Ji, S., Shan, J., Gong, J., & Liu, K. (2017). Line-based registration of panoramic images and LiDAR point clouds for mobile mapping. Sensors, 17(1), 70.

Cyganek, B., & Siebert, J. P. (2009). An Introduction to 3D Computer Vision Techniques and Algorithms. United Kingdom: John Wiley & Sons, Ltd.

Elaksher, A. F. (2016). Co-registering satellite images and LIDAR DEMs through straight lines. International Journal of Image and Data Fusion, 7(2), 103-118.

Deng, F., Hu, M., & Guan, H. (2008). Automatic registration between LiDAR and digital images. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 487-490.

Fischler, M. A., & Bolles, R. C. (1981). Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Communications of the ACM*, *24*(6), 381-395.

González-Aguilera, D., Rodríguez-Gonzálvez, P., & Gómez-Lahoz, J. (2009). An automatic procedure for co-registration of terrestrial laser scanners and digital cameras. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*(3), 308-316.

Aguilera, D. G., Gonzálvez, P. R., & Lahoz, J. G. (2007). Automatic co-registration of terrestrial laser scanner and digital camera for the generation of hybrids models. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 36, 162-168.

Goshtasby, A. (2012). Image Registration: Principles, Tools and Methods (Vol. 2): Springer-Verlag London.

ISPRS Benchmark Dataset, http://www2.isprs.org/commissions/comm1/wg4/benchmark-test.html, accessed 18 January 2015

Hofmann, S., Eggert, D., & Brenner, C. (2014). Skyline matching based camera orientation from images and mobile mapping point clouds. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 2(5), 181.

Kenney, C. S., Zuliani, M., & Manjunath, B. S. (2005, June). An axiomatic approach to corner detection. In null (pp. 191-197). IEEE.

Kim, H., Correa, C. D., & Max, N. (2014, May). Automatic registration of LiDAR and optical imagery using depth map stereo. In Computational Photography (ICCP), 2014 IEEE International Conference on (pp. 1-8). IEEE.

Kim, P. E., Cho, Y. K., & Chen, J. (2016, December). AUTOMATIC REGISTRATION OF LASER SCANNED COLOR POINT CLOUDS BASED ON COMMON FEATURE EXTRACTION. In Proceedings of the 16th International Conference on Construction Applications of Virtual Reality(Vol. 11, p. 13).

Lee, J.-H., Biging, G.S., & Fisher, J. B. (2016). An Individual Tree-Based Automated Registration of Aerial Images to LiDAR Data in a Forested Area. *Photogrammetric Engineering & Remote Sensing*, *82*(9), 699-710.

Lee, J., Cai, X., Schonlieb, C.-B., & Coomes, D. A. (2015). Nonparametric Image Registration of Airborne LiDAR, Hyperspectral and Photographic Imagery of Wooded Landscapes. *IEEE Transactions on Geoscience and Remote Sensing*, *53*(11), 6073-6084.

Lee, J., Lee, C., & Yu, K. (2011). Autoregistration of high-resolution satellite imagery using LIDAR intensity data. *KSCE Journal of Civil Engineering*, *15*(2), 375-384.

LeMoigne, J., Netanyahu, N. S., & Eastman, R. D. (2012). Image registration for remote sensing. *Geophysical Journal International*, *188* (1), 381.

Lhuillier, M., & Quan, L. (2000). Robust dense matching using local and global geometric constraints. In Pattern Recognition, 2000. Proceedings. 15th International Conference on (Vol. 1, pp. 968-972). IEEE.

Lie, W.-N., Lin, T. C. I., Lin, T.-C., & Hung, K.-S. (2005). A robust dynamic programming algorithm to extract skyline in images for navigation. *Pattern Recognition Letters*, *26*(2), 221-230.

Liu, S., Lv, Y., Tong, X., Xie, H., Liu, J., & Chen, L. (2016). An Alternative Approach for Registration of High-Resolution Satellite Optical Imagery and ICESat Laser Altimetry Data. *Sensors (Basel)*, *16*(12).

Liu, S., Tong, X., Chen, J., Liu, X., Sun, W., Xie, H., . . . Ye, Z. (2016). A Linear Feature-Based Approach for the Registration of Unmanned Aerial Vehicle Remotely-Sensed Images and Airborne LiDAR Data. *Remote Sensing*, *8*(2) 82.

Liu, Z. Q., Li, P. C., Huang, S. H., Ye, C. L., Ma, Q., & Yang, J. J. (2018). AUTOMATIC GLOBAL REGISTRATION BETWEEN AIRBORNE LIDAR DATA AND REMOTE SENSING IMAGE BASED ON STRAIGHT LINE FEATURES. International Archives of the Photogrammetry, Remote Sensing & Spatial Information Sciences, 42(3). 1191-1197.

Longuet-Higgins, H. C. (1981). A computer algorithm for reconstructing a scene from two projections. *Nature*, *293*, 133–135

Lowe, D. (2004). Distinctive Image Features from Scale-Invariant Keypoints. International Journal of Computer Vision, 60(2), 91-110.

Mastin, A., Kepner, J., & Fisher, J. (2009). Automatic registration of LIDAR and optical images of urban scenes. IEEE Conference on Computer Vision and Pattern Recognition, FL, USA, pp. 2639-2646,.

Meierhold, N., Spehr, M., Schilling, A., Gumhold, S., & Maas, H. G. (2010). Automatic feature matching between digital images and 2D representations of a 3D laser scanner point cloud. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci, 38, 446-451.

Mikolajczyk, K., & Schmid, C. (2005). A Performance Evaluation of Local Descriptors. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(10), 1615-1630.

Kumar Mishra, R., & Zhang, Y. (2012). A review of optical imagery and airborne lidar data registration methods. The Open Remote Sensing Journal, 5(1), 54-63.

Moussa, W., Abdel-Wahab, M., & Fritsch, D. (2012). An automatic procedure for combining digital images and laser scanner data. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 39, B5, 229-234.

Otsu, N. (1979). A threshold selection method from gray-level histograms. IEEE transactions on systems, man, and cybernetics, 9(1), 62-66.

Palenichka, R. M., & Zaremba, M. B. (2010). Automatic Extraction of Control Points for the Registration of Optical Satellite and LiDAR Images. IEEE Transactions on Geoscience and Remote Sensing, 48(7), 2864-2879.

Parmehr, E. G., Fraser, C. S., Zhang, C., & Leach, J. (2013). Automatic registration of optical imagery with 3d lidar data using local combined mutual information. ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences, II-5/W2, 229-234.

Patel, T. P., & Panchal, S. R. (2014). Corner Detection Techniques: An Introductory Survey. International Journal of Engineering Development and Research, 2(4), 3680-3686.

Pratt, W. K. (2001). *Digital Image Processing* (Vol. 3). California: John Wiley & Sons, Inc.

Rodehorst, V., & Koschan, A. F. (2006). Comparison and Evaluation of Feature Point Detectors. *Turkish-German Joint Geodetic Days Geodesy Geoinformation Service Daily Life*, Berlin, Germany, 2006, pp. 1–8.

Rönnholm, P., Karjalainen, M., Kaartinen, H., Nurminen, K., & Hyyppä, J. (2013). Relative Orientation between A Single Frame Image and Lidar Point Cloud Using Linear Features. *The Photogrammetric Journal of Finland, 23*(2).

Safdarinezhad, A., Mokhtarzade, M., & Valadan Zoej, M. (2016). Shadow-Based Hierarchical Matching for the Automatic Registration of Airborne LiDAR Data and Space Imagery. *Remote Sensing, 8*(6), 466.

Sattler, T., Leibe, B., & Kobbelt, L. (2011, November). Fast image-based localization using direct 2d-to-3d matching. In Computer Vision (ICCV), 2011 IEEE International Conference on (pp. 667-674). IEEE.

Schmitt, D., & McCoy, N. (2011). Object classification and localization using SURF descriptors. CS, 229, 1-5.

Shorter, N., & Kasparis, T. (2008, July). Autonomous registration of LiDAR data to single aerial image. In Geoscience and Remote Sensing Symposium, 2008. IGARSS 2008. IEEE International (Vol. 5, pp. V-216). IEEE.

Toth, C., Ju, H., & Grejner-Brzezinska, D. (2011). Matching between different image domains. In Photogrammetric Image Analysis (pp. 37-47). Springer, Berlin, Heidelberg.

Yang, Z. H., Zhang, Y. S., Zheng, T., Lai, W. B., Zou, Z. R., & Zou, B. (2016). Co-Registration Airborne Lidar Point Cloud Data and Synchronous Digital Image Registration Based on Combined Adjustment. International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, XLI-B1, 259-264.

Zhang, W., Zhao, J., Chen, M., Chen, Y., Yan, K., Li, L., Chu, Q. (2015). Registration of optical imagery and LiDAR data using an inherent geometrical constraint. *Opt Express*, *23*(6), 7694-7702.

Zhu, S., Pressigout, M., Servières, M., Morin, L., & Moreau, G. (2012). Skyline Matching: A robust registration method between Video and GIS. Usage, Usability, and Utility of 3D City Models–European COST Action TU0801, 03007.

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