Sayyad Shahzad,
Manfred Wiggenhagen

Co-registration of terrestrial laser scans and close range digital images using scale invariant features

ALLGEMEINE VERMESSUNGS-NACHRICHTEN

Dieser Beitrag wurde nach Begutachtung als
PEER REVIEWED PAPER
angenommen.
Tag der Einreichung: 28. Februar 2010

Standard approaches for the co-registration of terrestrial laser scans (TLS) and close range digital images (CRDI) taken separately require often artificial targets. These approaches are reliable but not efficient for larger projects. Our approach applies scale invariant feature detection methods to make the co-registration process more flexible. But the reliability of feature matching based on the design of feature descriptors is sometimes questionable. The accuracy of the applied algorithm can be improved by introducing some additional geometric constraints.

Our approach consists of a three-step procedure. In the first step scale invariant feature detection in the brightness image from the digital camera and the corresponding intensity image from the terrestrial laser scanner is carried out. In the next step, the initial matching values of the corresponding points are corrected by introducing additional constraints. Finally, from each set of match, the affine transformation parameters are calculated so that the 3D point cloud and brightness image can be registered together.

1. Introduction

Terrestrial laser scanning is a relatively new technology in geodesy. It delivers fast and dense 3D information of the object of interest without the mandatory use of targets. Poor intensity image obtained from terrestrial laser scanners can be replaced by multi-channel high resolution images obtained from the digital camera. Thus, 3D point clouds from terrestrial laser scanners with complete texture information obtained from digital images can be used for modelling and visualization of three dimensional objects.

This image- and range-based modelling requires co-registration of both image and range data in one common coordinate system. Using the standard techniques, artificial targets are used to register CRDI and range data of the terrestrial laser scanner (Ullrich et al., 2003). Al-Manasir and Fraser (2006) used coded targets placed on the object to register digital images in the coordinate system of the laser scanner by applying a 3D similarity transformation. However, their proposed method requires a digital camera to be rigidly attached to the laser scanner. These methods are reliable but its quality of the registration depends upon the distribution and choice of targets. Furthermore, the feasibility of these approaches in hazardous areas due to limiting access to the object of interest is always a question.

There exist many co-registration techniques which work automatically without the need of artificial markers or external sensors. (Habib et al., 2002) presented a method of automatic co-registration of point clouds and CCD camera images. They used „tie lines" instead of „tie points" to improve the quality of the co-registration. Dold and Brenner (2006) used planar patches for co-registration of terrestrial laser scans obtained from different positions. Imagery data is used to improve the search algorithm of identifying the corresponding planar patches. Terrestrial laser scanners provide 3D range information along with reflectivity value of each point. This reflectivity value can be used to register range data with digital images. (Haala et al., 2004) oriented digital images from a panoramic camera in a reference system of the terrestrial laser scanner by spatial resection based on control points. Wendt and Heipke (2006) investigated the integration of central perspective images of a pinhole camera and range and intensity data of laser scanners. They used intensity images

along with range data to improve the detection of corresponding points.
Böhm and Becker (2007) used the scale-invariant feature transform (SIFT) developed by Lowe (2003) for co-registration of terrestrial laser scans and images obtained from CCD. They applied SIFT algorithm directly to intensity images of the scanner. They used geometric verification technique (i. e., RANSAC) to filter out falsified matching pairs. The number of correct matches was found only 20 % of total match.
The registration method proposed in this paper is based on the extraction of „tie points" from the intensity image of the terrestrial laser scanner and the corresponding high resolution digital images obtained from a separate hand-held CCD camera. These „tie points" are invariant to scale and rotation, robust to change in illumination conditions and noise, and small change in 3D viewpoint. The approach used to detect these features is referred as „Speeded-Up Robust Features" (SURF) and is developed by (Bay et al. 2006). Unlike SIFT, SURF got different types of descriptors. In this paper we have used the regular version SURF-64 (descriptor length 64) and the up-right version U-SURF (rotation invariance of keypoints is not considered). SURF yields results with accuracy comparable to SIFT with key advantage of less computational cost by making an efficient use of integral images. Furthermore, matching speed of features is enhanced by introducing sign of Laplacian. In Section 2, some of the key differences between the intensity image of terrestrial laser scanner and CRDI are discussed. Keypoints are localized by using two different versions of SURF in Section 3. Nearest neighbour search (using Euclidean distance) was performed for the detection of the corresponding SURF descriptors. The obtained results have been improved by using our proposed algorithm which is based on geometric constraints between two images. In Section 4, conclusion and some topics for further research are outlined.

2. Detail about the data

In this paper, test scans have been acquired using a Z + F Imager 5006 (Z+F Imager, 2009) scanner. It is a phase shift measurement scanner. Its field of view is $(360° \times 310°)$. It has a vertical and horizontal resolution of 0.0018°and a range accuracy of 6.8mm at 50m by scanning an object with 10 % reflectivity. Test images have been taken by using a Samsung 12Mp camera. It has an image size of $5.761mm \times 4.32mm$ and a pixel size of$1.8\mu m \times 1.8\mu m$. A small area with some text and artificial targets in the 3D-laboratory and an open area with natural features has been scanned as well as imaged independently with an arbitrary scale and rotation. This type of data set provides symmetry and similarity in scenes and hence can be used as a measure of reliability of the design of the SURF descriptor.
The standard output of a terrestrial laser scanner is the distance, the horizontal and vertical angles of each point of the object surface with respect to the scanner position. In addition to this range data, the registered intensity corresponding to each range measurement based upon the intensity of the object surface at the wavelength of the implemented laser light is also registered. This intensity value is used to obtain a monochromatic intensity image. This image can be used to register the terrestrial laser scans and the digital images directly together by an image matching algorithm. However, the intensity image differs from a digital camera image in many ways. A summary of the main differences between two types of images is given below.
Light conditions are crucial for *passive sensors* (digital cameras). The shadows on the object caused by ambient light result in an increase in noise and lead to false mapping of grey values in the image. On the other hand, terrestrial laser scanners being *active sensors* are independent from external light and hence shadow effect can be minimized. *Field of view* in case of a laser scanner is larger than for camera views. Multispectral high quality images with complete texture information are obtained through digital cameras while the intensity image obtained from terrestrial laser scanners has inferior quality and it doesn't contain any texture information. Standard digital images have *central perspective geometry* while the intensity images have *spherical geometry.* Whereas straight lines in object space remain *straight* in standard digital images, they may appear *curved* in the intensity images. *The resolution* of intensity images is generally lower as compared to digital images.
The above mentioned differences between the two types of images require an interest point detection and matching technique which is not only invariant to scale and rotation but also robust to minor change in view point and illumination conditions and at the same time robust to detection displacement and geometric deformation.

3. Proposed Algorithm

As mentioned in Section 1, the SURF detector and descriptor was used in the current project for discrete image point correspondences in the intensity image and CRDI. Following pipeline has been adopted to register a first data set together. For the second data set only the results have been shown in Section 3.3.

3.1 Pre-processing

1. Terrestrial laser scanner stores intensity values in 16-bit numbers. The most of the standard image processing tools are designed to work with 8-bit image, therefore intensity values are stored into 8-bit format to work with these standard tools.
2. As image obtained from normal CCD camera deviates from central perspective projection and straight lines in scene doesn't remain straight in image space. *PhotoModeler Scanner* (PhotoModeler Scanner 2009) has been used to calibrate the used digital camera and to find interior orientation parameters of the camera and an idealized central perspective image has been obtained by applying the corrections obtained from interior orientation parameters.
3. Both of images have been resized to reduce computational time.

3.2 SURF based interest point description and matching

The SURF detector is based on the Hessian matrix. Blob-like structures are detected at locations where the determinant of the Hessian matrix has a maximum. Interest points are localized in scale space by performing non maxima suppression. The distribution of the intensity content within the interest point neighbourhood is described by a feature vector. Each interest point is assigned an orientation by calculating Haar wavelet responses in its vicinity. The extracted interest points by using SURF-64 for each image have been shown in Figure 1.
Nearest neighbour search (using the Euclidean distance) has been used to match two sets of descriptors extracted from two images. It has been found that the relationship between the nearest and the second nearest neighbour plays an important role in the quality of results of matching. We have used a value of 0.7 for the nearest neighbour matching. The speed and robustness of matching has been increased by considering the sign of the trace of Hessian matrix for the underlying interest point. The features having same type of contrast are compared together. The matching results can be seen in Figure 2. It is obvious that some matched pairs are false and need to be discarding from correct matches.

3.3 Identification and removal of False Matches

Matching is based on Euclidean distance of the feature vectors without taking into account the geometry of images. Due to the overall descriptor design and different image projections, the matching algorithm based on the nearest-neighbour delivers some degree of false matching. These false matches can be easily identified. A detailed analysis of the falsified matching allows designing an algorithm which helps us to discard these points automatically.
The geometric verification technique applied in this paper makes use of some very simple geometric constraints. For two overlapping images I_1 & I_2, the following criterion should be applicable.

1. If $P_{(i)1}$ & $P_{(i)2}$ for $i \in \{1, ..., n\}$ are the corresponding points in image I_1 & I_2 respectively, n is the number of the matching pairs and dx_i & dy_i is the change in the distance between $P_{(i)1}$ and $P_{(i)2}$ in x & y direction respectively, then the length $d_i = \sqrt{dx_i^2 + dy_i^2}$ and the

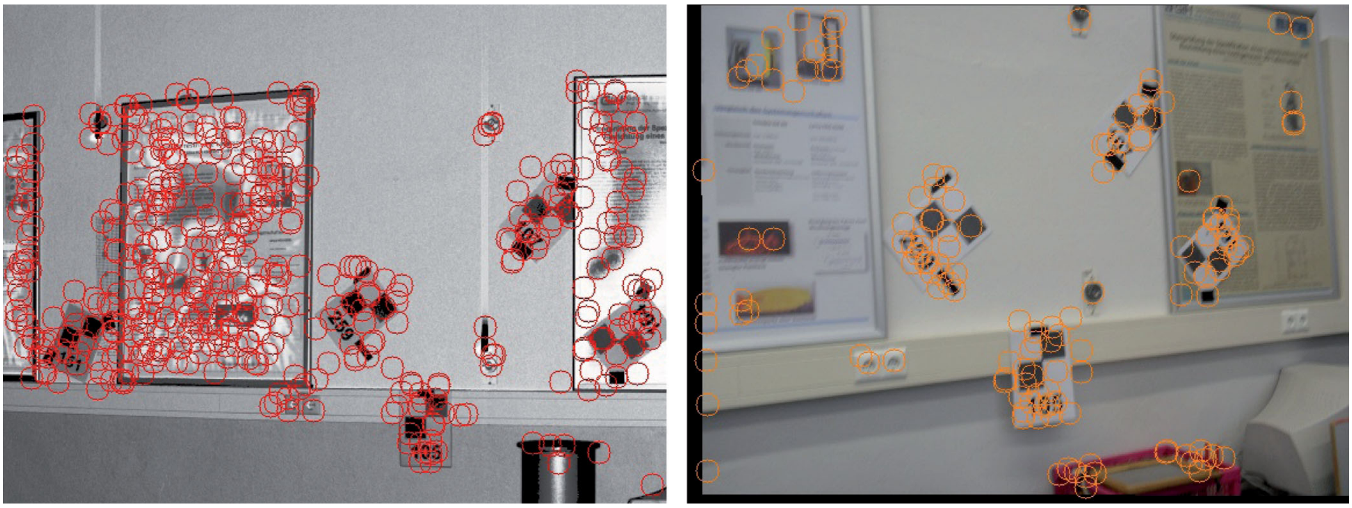

Figure 1: SURF interest points mapped on scanner image (left) and camera image (right)

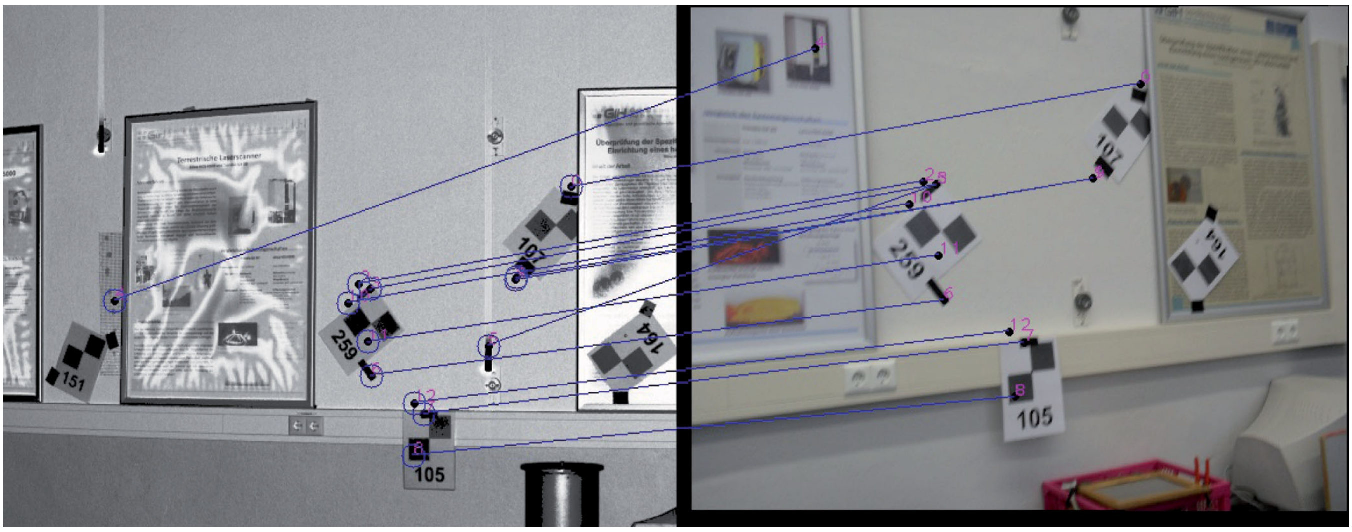

Figure 2: Nearest neighbor search (using Euclidean distance) between scanner image (left) and camera image (right)

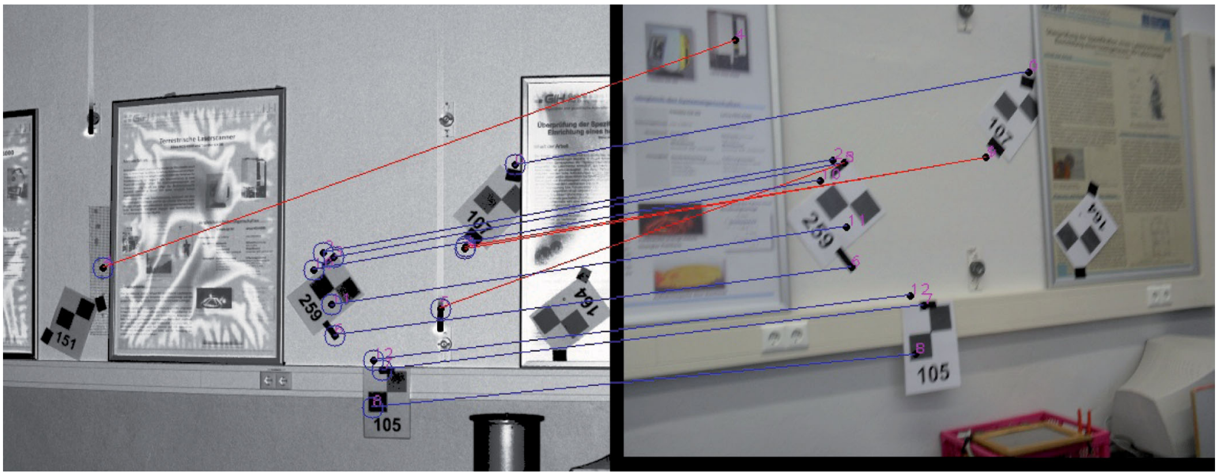

Figure 3: Corrected matching (using our proposed algorithm)

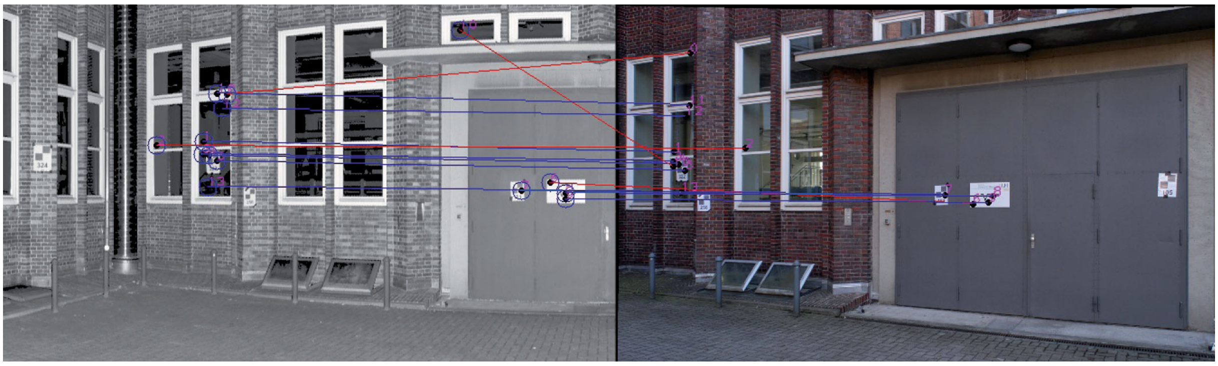

Figure 4: Corrected matching (using our proposed algorithm) for the second data set.

slope $m_i = {dy_i}/{dx_i}$ of a line li joining the points $P_{(i)1}$ & $P_{(i)2}$ should remain constant.

2. If $d_{(ij)1}$ for $j \in \{1, ..., n\}$ is the distance of point i from any other point j in image I_1 and $d_{(ij)2}$ is the distance of point i from any other point j in image I_2 respectively, then $D_i = d_{(ij)1} / d_{(ij)2}$ should be a constant.
3. If $o_{(ij)1}$ is the change in the orientation of point i from any other point j in image I_1 and $o_{(ij)2}$ is the change in the orientation of point i from any other point j in image I_2 respectively, then $O_i = o_{(ij)1} / o_{(ij)2}$ should be a constant.
4. It can be only one corresponding point $P_{(i)2}$ in image I_2 for a point $P_{(i)1}$ of imageI_1.

As our images came from two different sources and have two different image projections, we can apply the geometric constraints defined above with a small change. Instead of defining a constant value for distances and slope, a value with some small standard deviation can be chosen as criteria to verify true match pairs.

A small algorithm has been written by making use of above knowledge and all the false matching pairs have been removed. The result for data set 1 is shown in Figure 3. Wrong matches are shown in red lines.

An other area with some typical features for building, e. g. windows and colums providing symmetry and self-similarity have been scanned as well as imaged. These types of repetitive features cause false matches which need to be identified. The result after applying correction based on our proposed algorithm is shown in Fig.4.

Table 1 shows the number of correct matches after applying our proposed algorithm. It has been found that U-SURF requires less computational time as compared to SURF. Therefore, it is recommended to use U-Surf in case where the change in orientation is small. Finally, affine transformation parameters are computed.

4. Summary and Conclusions

An automatic registration method for terrestrial laser scans and digital images has been represented in this paper. Our proposed method makes use of intensity images directly available as by-product along with the 3D point cloud. Scale- and orientation invariant features using the SURF algorithm have been extracted from the intensity images of terrestrial laser scanner and the corresponding digital images obtained from independent hand-held camera. Two versions of SURF (SURF64 & U-SURF) have been used in this paper. It has been found that with limited change in orientation U-SURF can be used for finding „tie points“ with higher speed. It has been found that this feature matching technique works well for convergent geometry with scale factor (<2) and rotation (<30 degrees), but the exact limitations are under investigation and are part of our future research. Some geometric constraints between two images have used to discard false matching due to the different projection of both images and to improve the reliability of matching algorithm. Our proposed

Table 1: Total correct matches obtained after applying our algorithm to remove false matches.

	Algorithm	Total matches	Correct matches	Percentage correct
Data-set 1	U-SURF	23	18	78 %
	SURF	13	9	69 %
Data-set 2	U-SURF	25	16	64 %
	SURF	14	10	71 %

algorithm works well even only 51 % matches are correct. We haven't considered time consumption explicitly in our experiment. The reason is that this is a well known fact that U-SURF performs faster than SURF. Furthermore, it has been experimentally found that the number of correct matches obtained by U-SURF is higher than SURF. Finally, the result shows that our proposed algorithm is simple, efficient and reliable.

This work is a part of an on-going research. In future some more geometric constraints should be used along with SURF descriptor during the matching stage to get more corresponding points with higher reliability. In addition, this pair wise registration technique must be extended to multi- image registration to enhance robustness and to explore more capabilities of the proposed algorithm.

Reference

[1] Al-Manasir, K., & Fraser, C. S.(2006): Registration of Terrestrial Laser Scanner Data using Imagery. Journal of Photogrammetric Record 21(115), pp. 255–268

[2] Bay, H., Tuytelaars, T., & Van Gool, L. (2008): SURF: Speeded Up Robust Features. Computer Vision and Image Understanding (CVIU), Vol. 110, No. 3, pp. 346-359

[3] Böhm, J. & Becker, S.(2007): Automatic Marker-Free Registration of Terrestrial Laser Scans using Intensity Features. In: 8th Conf. on Optical 3-D Measurement Techniques, Zurich: pp. 338– 344

[4] Dold, C., & Brenner, C. (2006): Registration of Laser Scanning Data using Planar Patches and Image Data. ISPRS Archieves, Dresden: pp. 78-83

[5] Haala, N., Reulke, R., Thies, M. & Aschoff, T. (2004): Combination of Terrestrial Laser Scanning with High Resolution Panoramic Images for Investigations in Forest Applications and Tree Species. Panoramic Photogrammetry Workshop, IAPRS Vol. 34, Part 5/W16

[6] Habib, A., Ghanma, M., Tait, M., & Fox, R. (2002): Using Ground Based Laser Scanners to Establish the Orientation of Terrestrial Imagery. Proceedings of the CIPA WG 6 International Workshop on Scanning for Cultural Heritage Recording, Corfu, Greece

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

[8] PhotoModeler Scanner. (2009): http://www.photomodeler.com

[9] Ullrich, A., Schwarz, R., Kager, H. (2003): Multistationsausgleichung für ein Laserscanner-System mit aufgesetzter Kamera, in: Österreichische Zeitschrift für Vermessung und Geoinformation (VGI), vol. 91,pp. 281 – 287

[10] Wendt, A., Heipke, C. (2006): Simultaneous orientation of brightness, range and intensity images, ISPRS Symposium Comm. V, WG V/3, „Image Engineering and Vision Metrology", Dresden

[11] Z + F Imager. (2009): http://www.zf-laser.com

Addresses of the authors:

M. Sc. Eng. Sayyad Shahzad
Geodätisches Institut
Nienburger Straße 1
30167 Hannover
E-Mail: shahzad@gih.uni-hannover.de

Dr.-Ing. Manfred Wiggenhagen
Institut für Photogrammetrie und GeoInformation
Nienburger Straße 1
30167 Hannover
E-Mail: wiggenhagen@ipi.uni-hannover.de

Zusammenfassung

Standard Ansätze für die Co-Registrierung von terrestrischen Laserscans (TLS) und digitalen Bildern aus dem Nahbereich (CRDI) erfordern oft künstliche Zielmarken. Diese Ansätze sind zwar zuverlässig, aber bei größeren Projekten nicht effizient. Unser Ansatz setzt Methoden zur skaleninvarianten Merkmals-Erkennung ein, um die Registrierungsverfahren flexibler zu gestalten. Da die Zuverlässigkeit der Merkmalszuordnung vom Design des Deskriptors abhängig ist, sind die Ergebnisse allerdings manchmal zweifelhaft. Die Genauigkeit des verwendeten Algorithmus kann durch die Einführung zusätzlicher geometrischer Bedingungen verbessert werden.

Unser Ansatz besteht aus einem Drei-Stufen-Verfahren. Im ersten Schritt wird die skaleninvariante Merkmals-Erkennung im Grauwertbild der Digitalkamera und dem entsprechenden Intensitätsbild des terrestrischen Laserscanners durchgeführt. Im nächsten Schritt werden die Startwerte der korrespondierenden Punkte durch die Einführung zusätzlicher geometrischer Einschränkungen korrigiert. Schließlich werden aus jedem Satz der Zuordnung die affinen Transformationsparameter so berechnet, dass die 3D-Punktwolken und das Grauwertbild ineinander zugeordnet werden können.