

THE COMPARISON OF 2D AND 3D DETECTORS FOR TLS DATA REGISTRATION – PRELIMINARY RESULTS

J. S. Markiewicz^{1*}, Ł. Markiewicz², P. Forys²

¹Department of Photogrammetry, Remote Sensing and Spatial Information Systems, Faculty of Geodesy and Cartography, Warsaw University of Technology; jakub.markiewicz@pw.edu.pl

²Virtual Reality Techniques Division, Institute of Micromechanics and Photonics, Faculty of Mechatronics, Warsaw University of Technology; l.markiewicz@mchtr.pw.edu.pl, p.forys@mchtr.pw.edu.pl

Commission II

KEY WORDS: cultural heritage, TLS data registration, feature-based method, 2D and 3D detectors, accuracy analysis

ABSTRACT:

This paper presents the analysis of possible methods of a terrestrial laser scanning (TLS) data registration using 2D/3D detectors and descriptors. The developed approach, where point clouds are processed in form of panoramic images, orthoimages and 3D data, was described. The accuracy of the registration process was preliminary verified. The two approaches were analysed and compared: the 2D SIFT (Scale-Invariant Feature Transform) detector and descriptor with the rasterized TLS data and the 3D SIFT detector with the 3D FPFH (Fast Point Feature Histograms) descriptor. The feature points were found and preliminary matched using the OpenCV and PCL (Point Cloud Library) libraries. In order to find the best point cloud representation for the registration process, both the percentage and distribution of the correctly detected and matched points were analysed. The materials consisted of the point clouds of two chambers from the Museum of King John III's Palace in Wilanów. They were acquired using the Z+F 5006h and 5003 TLS scanners. The performed analysis showed that the lowest RMSE values were for the 2D detectors and orthoimages. However, in the case of the point number and distribution, better results were obtained for using the 3D detector.

1. INTRODUCTION

Nowadays, generation of a precise, high quality and resolution architectural documentation based on the terrestrial laser scanning (TLS) or dense point clouds from the digital images is still an open and challenging issue. The processing of the TLS data consists of the data acquisition, registration, and generation of 3D models and orthoimages.

Full measurements of large and complex architectural objects are possible using several different TLS scanner stations (positions). Each of the TLS point clouds is acquired in a local reference system of the instrument. Therefore, in the case of multiple sources of data, a transformation to a common global reference system is required. This process is called registration and consists of a determination of exterior orientation parameters, i.e. three rotation angles and three elements of the translation matrix (Lichti and Pfeifer, 2008; Luhmann et al., 2013; Vosselman and Maas, 2010). In order to solve this problem, the affine 3D transformation is performed. However, in case of a scale difference between a geodesic and the scanner coordinate systems, it is also possible to apply the 3D transformation by a similarity. This transformation is calculated using special control points which have to be evenly distributed across the entire overlapping area of the point clouds.

In order to register the TLS point clouds, two main methods are usually used (Vosselman and Maas, 2010). They are based on either the points or the features. The first, target-based method is a conventional approach of the point-based registration and is usually performed using targets placed on the object prior to the scanning. The second commonly applied procedure utilize iterative closest point (ICP) algorithm (Durrant-Whyte and

Bailey, 2006; Liu, 2006; Nuchter et al., 2007; Vosselman and Maas, 2010). The ICP requires a good initial approximation in order to converge to the global minimum. However, even if there is considerable overlap, convergence is not always guaranteed. Furthermore, the ICP algorithm can be computationally intensive and time-consuming in its search for tie points in overlapping scans (Kang et al., 2009).

Effective generation of metric documentation of complex cultural heritage objects is not possible without a high degree of automation.

2. THE STATE OF THE ART

One of the biggest disadvantages of the traditional feature-based orientation methods is the need of placing the targets onto the historical objects. The present research works aim at the full elimination of the targets in order to make the scanning process noninvasive, faster and easier. This can be done for example via detection and matching of corresponding geometrical shapes (e.g. lines, planes etc.) and colour features of the different TLS point clouds. During this process, both 2D/3D detectors and 2D/3D descriptors of keypoints can be utilized. Another advantage of using such methods is the possibility of utilizing the raster data in form of images and transformed TLS point clouds.

The detection and description of the keypoints is a well-studied subject in the field of the 2D images analysis. It allows automatic determination of exterior orientation as in, e.g., the Structure-from-Motion (SfM) approach (Bianco, 2018; Chiabrando et al., 2015; Moussa, 2006; Westoby et al., 2012).

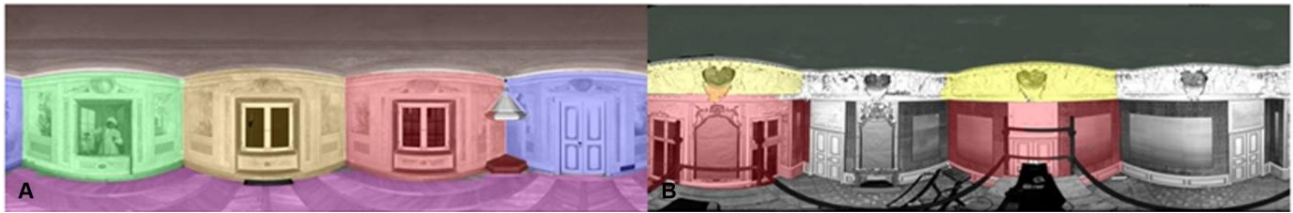


Figure 1. The spherical projection of the point clouds with marked fragments of the used orthoimages: A) Test site I, B) Test site II (Markiewicz, 2017)

Generally, different types of 2D features may be detected on images in order to determine corresponding image contents (Weinmann, 2013). Three types of the detectors are usually used to extract the characteristic points from the rasters: corner detectors (e.g., Harris (Harris and Stephens, 1988), FAST (Features from Accelerated Segment Test; Rosten and Drummond, 2006)), line (Canny, 1986) or blob detectors (e.g., SIFT (Scale-Invariant Feature Transform; Lowe, 1999), SURF (Speeded Up Robust Features; Bay and Ess, 2008)). Nowadays, the SURF and the SIFT algorithms are the most frequently applied detectors of the image features (Urban and Weinmann, 2015).

The 2D detectors were originally designed for the images in central projection. In the case of spherical photographs (e.g., the raw TLS data), image distortions are significant which can make explicit identification and matching keypoints problematic (Markiewicz, 2017). A possible solution involves a transformation of the point cloud into the form of “a virtual photograph in the central projection” (Moussa et al., 2012), i.e. an orthoimage (Markiewicz, 2017). However, in some cases intensity variations between the registered scans can affect the keypoints detection and matching quality. This effect is not present in the 3D feature matching, but this approach is more complex and time-consuming.

A wide variety of 3D keypoint detectors and feature descriptors are described in the literature (Ghorpade et al., 2016; Moreels and Perona, 2006; Salti et al., 2012). The 3D keypoint detectors can be divided into two main categories – fixed-scale and adaptive-scale detectors. The fixed-scale detectors find distinctive keypoints at a specific, constant scale passed as a parameter to the algorithm. On the other hand, the adaptive-scale detectors are based on the iterative 3D shape detection (Salti et al., 2012).

In this paper, a comparison of the typical 2D and 3D feature-based TLS data registration methods was presented. This is a continuation and an extension of the previous ISPRS article (Markiewicz, 2017).

3. TLS DATA REGISTRATION – PERFORMED EXPERIMENTS

The TLS point clouds of two chambers, located at the Museum of King John III’s Palace in Wilanów, were used for the analysis. The data was acquired by two different TLS scanners – Z+F 5006h (Z+F, 2017) and 5003 (Staiger, 2005) (Figure 1) with the resolution of 3 mm at 10 m. The distance between the scanner station and the object ranged from 2 to 10 m. In order to ensure a uniformity of the point cloud resolution, it was resampled to 3 mm. In the first step, the data was filtered by the reflectance intensity threshold. The measurement targets (treated as check points), which are commonly applied in the TLS data registration, were located on the second test site.

The first test site was characterized by a number of adornments such as bas-reliefs, facets or fabrics hanging on the walls. Furthermore, wall paintings contained spatial patterns (Figure 1A). The second test site was even more geometrically complex. It had richer adornments, mirrors in golden frames and a decorative fireplace (Figure 1B). The points clouds were obtained from different positions and angles. This had an influence on the radiometric (intensity) and geometrical quality. The examples of the TLS point clouds are presented in Figure 2.

3.1 Keypoint detection and description

One of the most important step during the registration process is the detection of the tie points in the 3D data. In this article, the two approaches were analysed and compared: the 2D SIFT (Lowe, 2004) detector and descriptor in addition to the 3D SIFT (Scovanner and Ali, 2007) detector with the 3D FPFH (Rusu et al., 2009) descriptor. In the first case, the raw TLS data was converted to the raster form, i.e. spherical projection and orthoimage. The feature points were found and preliminary matched using the OpenCV (OpenCV, 2018) and PCL libraries (PCL, 2018).

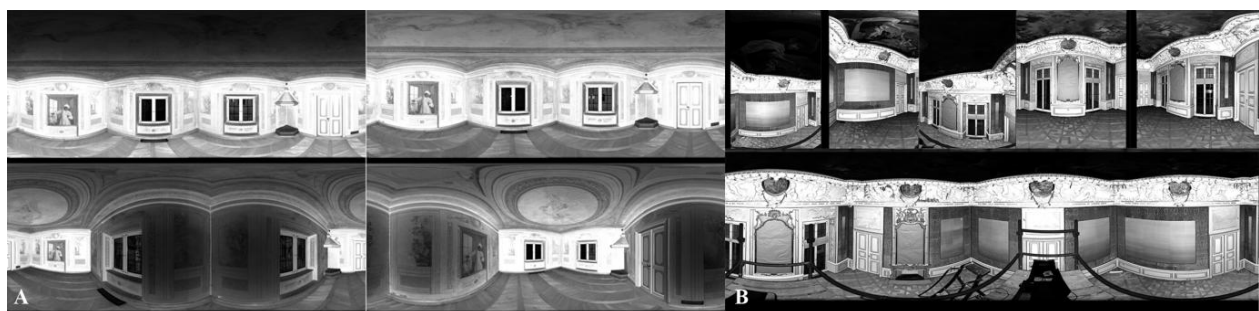


Figure 2. The examples of TLS point clouds (in the spherical form): A) Test site I: I-1, I-2, I-3, I-4, B) Test site II: II-1, II-2, II-3, II-4, II-5, II-6

Id		2D				3D		
		Orthoimage		Spherical image		Matched keypoints	% incorrect points	
		GSD	Matched keypoints	% incorrect points	Matched keypoints			% incorrect points
Test site I	I-1 : I-2	3 mm	1287	23.9	4360	3.8	482	37.4
	I-1 : I-3		1546	11.6	764	9.0	273	27.1
	I-1 : I-4		1631	11.5	278	22.3	255	26.3
Test site II	II-1 : II-6		552	40.6	22397	98.8	49	26.3
	II-2 : II-6		328	57.9	7609	97.5	33	5.5
	II-3 : II-4		581	41.9	350	78.8	47	9.0
	II-4 : II-5		152	26.3	88	81.8	50	9.7

Table 1. The percentage of the correctly detected and matched keypoints

In order to find the best point cloud representation for the registration process, the percentage of the correctly detected and matched points (Table 1) and its distribution (Figure 3 and 4) were analysed. The matching quality was assessed using an incremental bundle adjustment based on the least square method. In the first step, all detected points were treated as the control points and filtered during the iterative process. Similarly to the method presented in the article (Markiewicz, 2017), the difference between the value calculated during the least square adjustment process and the measured value was calculated. If it was smaller than the 5 mm, then was treated as a correct one.

Unfortunately, sometimes it is impossible to place the control points on the object due to its historical nature. In order to solve this problem, an independent quality assessment was developed. The automatically detected tie points were divided into the non-marked control and check points according to the previous article (Markiewicz, 2017).

The first stage of the data analysis consisted of testing the number of the correctly detected and matched tie points (Table 1). It might be noticed that the number of points detected in the 2D rasters is higher than in the 3D point clouds. The spherical images, in comparison to the orthoimages, were characterized by the smaller number of the detected points due to the effect of image "distortion". It was caused by the close proximity of the scanner to the wall. This geometrical deformation was eliminated by using orthoimages. On the other hand, the position of the TLS

scanner did not have an influence on the number of points detected with the 3D detector.

Because of the geometrical differences between the test sites, both of them had to be considered separately. For the Test site I, it appeared that the most reliable pair of the detector and the representation was the 2D detector with the orthoimages, then the 2D detector with the spherical images and finally the 3D detector operating on the point cloud.

In the case of the geometrically more complex Test site II, the number of the 2D detected tie points on the spherical images is lower than on the orthoimages. The 3D approach results were rather stable but still worse than in the 2D detector case.

The number of the tie points is not the only important aspect related to the data registration. Its distribution has also an essential impact on the orientation quality. Most preferably points should be evenly spread across the whole scan but it is not always possible. As an example, the worst cases of check and control point distributions for each of the test site and particular scanner stations are presented in Figures 3 and 4.

Based on the results shown in Figure 3, it might be noticed that the best distribution of the tie points was achieved for the case A (2D detector and orthoimage) and C (3D detector and point cloud). The spherical image approach did not allow to achieve the desired results.

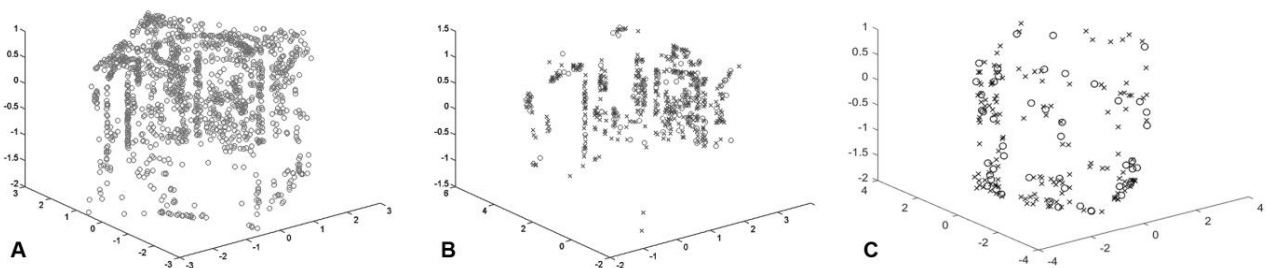


Figure 3. The distribution of the tie points for the worst case (Test site I): A) 2D detector and orthoimage, B) 2D detector and spherical image, C) 3D detector and point cloud

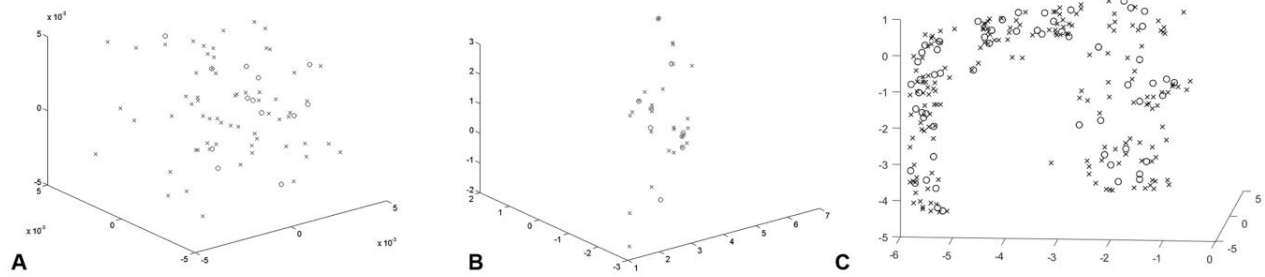


Figure 4. The distribution of the tie points for the worst case (Test site II): A) 2D detector and orthoimage, B) 2D detector and spherical image, C) 3D detector and point cloud

The RMSE of the orientation on the control points											
Id		2D						3D			
		Orthoimage			Spherical image			RMSE [mm]			
		RMSE [mm]									
		GSD	X	Y	Z	X	Y	Z	X	Y	Z
Site I	I-1 : I-2	3 mm	1.4	1.5	2.1	1.6	1.6	2.8	1.6	1.6	1.4
	I-1 : I-3		1.2	1.3	1.9	1.8	2.3	2.1	1.5	1.6	1.4
	I-1 : I-4		1.7	1.9	2.2	1.2	1.3	1.6	1.5	1.9	1.7
The RMSE of the orientation on the check points											
Id		2D						3D			
		Orthoimage			Spherical image			RMSE [mm]			
		RMSE [mm]									
		GSD	X	Y	Z	X	Y	Z	X	Y	Z
Site I	I-1 : I-2	3 mm	1.4	1.4	2.2	1.6	1.5	2.9	1.6	2.5	2.3
	I-1 : I-3		1.2	1.3	1.8	3.3	2.0	2.7	3.8	3.7	4.0
	I-1 : I-4		1.7	2.0	2.0	1.1	1.2	1.7	3.6	3.6	3.2

Table 2. RMSE of the point clouds orientation on the control and check points – Test site I

The similar analysis was done for the Test site II (Figure 4). In this case, the best distribution of the points was acquired using the 3D detectors which outperformed the 2D-based ones in terms of the coverage quality.

To summarize, if the point number and distribution are insufficient, it is recommended to combine the orthoimages with the 2D detector and 3D detector.

3.2 Accuracy analysis

The accuracy analysis was performed using the Matlab software. In order to assess the quality of the data registration, the RMSE (Root Mean Square Error) values on the control and check points were calculated.

Test site I

Table 2 presents the results of the registration based on the control and check points. It may be noticed that the accuracy of the orientation of the different point cloud fragments was very similar but the best results were achieved with the use of the 2D detector and orthoimages.

Test site II

The best orientation results for all of the pairs of scans was achieved using the 2D detector (Table 3). When the spherical

images were used, the orientation accuracy was significantly lower. This was caused by the "pseudo-distortion" effect on the spherical images and the intensity inconsistency between the scans. In the case of the 3D detector, it appeared that its RMSE was two times higher than the one obtained from the 2D detector.

4. FINAL REMARKS AND CONCLUSIONS

The complete and accurate 3D measurement of the cultural heritage object is not a trivial task. It needs special treatment and cannot be scanned using traditional invasive methods. The preliminary results of this research showed that it is possible to automate the registration process using the 2D/3D detectors on different kinds of the TLS data representation (e.g., 3D point cloud, orthoimage).

The following conclusions may be drawn based on the performed experiments:

- In the case of the 2D detectors, the best results were obtained for the orthoimages, in which the "distortion" did not occur.

The RMSE of the orientation on the control points											
Id		2D						3D			
		Orthoimage			Spherical image			RMSE [mm]			
		GSD									
RMSE [mm]						X	Y	Z			
		X	Y	Z	X	Y	Z	X	Y	Z	
Site II	II-1 : II-6	3 mm	2.4	2.2	1.9	2.5	2.5	1.9	4.9	3.9	4.2
	II-2 : II-6	3 mm	2.8	2.4	2.5	2.4	2.7	2.6	5.7	6.6	4.8
	II-3 : II-4	3 mm	2.7	2.2	1.8	2.6	2.3	2.9	6.6	6.7	4.5
	II-4 : II-5	3 mm	2.4	1.6	1.7	2.9	3.0	3.3	5.5	5.6	2.9
The RMSE of the orientation on the check points											
Id		2D						3D			
		Orthoimage			Spherical image			RMSE [mm]			
		GSD									
RMSE [mm]						X	Y	Z			
		X	Y	Z	X	Y	Z	X	Y	Z	
Site II	II-1 : II-6	3 mm	2.4	2.4	1.9	2.5	2.5	1.9	6.6	4.2	2.8
	II-2 : II-6	3 mm	2.6	2.7	3.1	2.4	2.7	2.6	8.8	6.2	6.0
	II-3 : II-4	3 mm	3.4	2.9	2.8	2.6	2.3	2.9	6.8	6.8	5.4
	II-4 : II-5	3 mm	3.4	3.2	1.9	2.9	3.0	3.3	3.8	5.6	5.3

Table 3. RMSE of the point clouds orientation on the control and check points – Test site II

- For the Test site I, the accuracy for both 2D and 3D methods was similar. However, in the case of the Test site II, 2D provided better results than 3D. Probably, it was caused by the lower geometrical quality of the point clouds.
- The 2D TLS registration was heavily affected by the intensity of the laser beam reflectance which differed on the corresponding surfaces. This had a negative influence on the number and distribution of the control and check points. The 3D SIFT detector was invariant to the differences in the intensity as it is based only on the shape analysis.
- The number and distribution of the 3D detected tie points for both test sites are similar for all pair of the scans. The 3D approach was more stable than 2D in this case.
- The 3D feature matching was suitable for finding the corresponding points in the interiors characterized by plenty of geometric details, such as bas-reliefs, facets or fabrics hanging on the walls (see Test site II).
- The 2D detector properly detected the tie points on the orthoimages in the mostly planar interiors which walls were covered by, e.g., paintings, patterns or other fabrics (see Test site I).
- While choosing the right detector (2D or 3D), the several factors have to be considered, such as the computation time, geometrical complicity or TLS scanner position.

To summarize, the 2D detector-based methods were less dependent on the shape features and were able to register the point clouds using mostly the texture analysis. On the other hand, the 3D geometrical approach was more robust to the uncalibrated intensity between the TLS scans.

In the future work, authors plan to focus on the further analysis of the possible registration methods using different 3D detectors

and its parameters and a wider range of the test data of both interiors and exteriors.

REFERENCES

- Bay, H., Ess, A., 2008. Speeded-Up Robust Features (SURF) 110, 346–359. <https://doi.org/10.1016/j.cviu.2007.09.014>
- Bianco, S., 2018. Evaluating the Performance of Structure from Motion Pipelines 1–18. <https://doi.org/10.3390/jimaging4080098>
- Canny, J., 1986. A Computational Approach to Edge Detection. IEEE Trans. Pattern Anal. Mach. Intell. PAMI-8, 679–698. <https://doi.org/10.1109/TPAMI.1986.4767851>
- Chiabrando, F., Donadio, E., Rinaudo, F., 2015. SfM for orthophoto generation: A winning approach for cultural heritage knowledge. Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. - ISPRS Arch. 40, 91–98. <https://doi.org/10.5194/isprsarchives-XL-5-W7-91-2015>
- Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping: Part I. IEEE Robot. Autom. Mag. 13, 99–108. <https://doi.org/10.1109/MRA.2006.1638022>
- Ghorpade, V.K., Checchin, P., Malaterre, L., Trassoudaine, L., Pascal, I., 2016. Performance Evaluation of 3D Keypoint Detectors for Time-Of-Flight Depth Data. <https://doi.org/10.1109/ICARCV.2016.7838686>
- Harris, C., Stephens, M., 1988. A Combined Corner and Edge Detector. Proceedings Alvey Vis. Conf. 1988 23.1-23.6. <https://doi.org/10.5244/C.2.23>
- Kang, Z., Li, J., Zhang, L., Zhao, Q., Zlatanova, S., 2009. Automatic Registration of Terrestrial Laser Scanning Point

- Clouds using Panoramic Reflectance Images 2621–2646.
<https://doi.org/10.3390/s90402621>
- Lichti, D., Pfeifer, N., 2008. Introduction to Terrestrial Laser Scanning 12219.
- Liu, Y., 2006. Automatic registration of overlapping 3D point clouds using closest points. *Image Vis. Comput.* 24, 762–781.
<https://doi.org/10.1016/j.imavis.2006.01.009>
- Lowe, D.G., 2004. Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vis.* 60, 91–110.
<https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- Lowe, D.G., 1999. Object recognition from local scale-invariant features. *Proc. Seventh IEEE Int. Conf. Comput. Vis.* 1150–1157 vol.2. <https://doi.org/10.1109/ICCV.1999.790410>
- Luhmann, T., Robson, S., Kyle, S., Boehm, J., 2013. Close-Range Photogrammetry and 3D Imaging, Close-Range Photogrammetry and 3D Imaging.
<https://doi.org/10.1515/9783110302783>
- Markiewicz, 2017. The example of using intensity orthoimages in tls data registration - A case study, in: *International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences - ISPRS Archives*.
<https://doi.org/10.5194/isprs-archives-XLII-2-W3-467-2017>
- Moreels, P., Perona, P., 2006. Evaluation of Features Detectors and Descriptors based on 3D Objects.
<https://doi.org/10.1007/s11263-006-9967-1>
- Moussa, W., 2006. Integration of Digital Photogrammetry and Terrestrial Laser Scanning for Cultural Heritage Data Recording. University of Stuttgart.
- Moussa, W., Abdel-Wahab, M., Fritsch, D., 2012. an Automatic Procedure for Combining Digital Images and Laser Scanner Data. *ISPRS - Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci. XXXIX-B5*, 229–234. <https://doi.org/10.5194/isprsarchives-XXXIX-B5-229-2012>
- Nuchter, A., Lingemann, K., Hertzberg, J., 2007. 6D SLAM-3D Mapping Outdoor Environments. *J. F. Robot.* 24, 699–722.
<https://doi.org/10.1002/rob.20209>
- OpenCV, 2018. URL <http://opencv.org/>
- PCL, 2018. URL <http://www.pointclouds.org/>
- Rosten, E., Drummond, T., 2006. Machine Learning for High Speed Corner Detection. *Comput. Vis. -- ECCV 2006* 1, 430–443. https://doi.org/10.1007/11744023_34
- Rusu, R.B., Blodow, N., Beetz, M., 2009. Fast Point Feature Histograms (FPFH) for 3D registration, in: *2009 IEEE International Conference on Robotics and Automation*. pp. 3212–3217. <https://doi.org/10.1109/ROBOT.2009.5152473>
- Salti, S., Petrelli, A., Tombari, F., Stefano, L. Di, 2012. On the Affinity between 3D Detectors and Descriptors.
<https://doi.org/10.1109/3DIMPVT.2012.10>
- Scovanner, P., Ali, S., 2007. A 3-Dimensional SIFT Descriptor and its Application to Action Recognition 3–6.
- Sprickerhof, J., Nüchter, A., Lingemann, K., Hertzberg, J., 2009. An Explicit Loop Closing Technique for 6D SLAM. *Computer (Long. Beach. Calif.)*. 1–6.
- Staiger, R., 2005. The Geometrical Quality of Terrestrial Laser Scanner (TLS). *FIG Work. Week 2005 GSIDI-8* 1–11.
- Theiler, P.W., Schindler, K., 2005. Automatic Registration of Terrestrial Laser Scanner Point Clouds Using Natural Planar Surfaces. *ISPRS Tech. Comm. III, WG III/2*.
- Urban, S., Weinmann, M., 2015. Finding a Good Feature Detector-Descriptor Combination for the 2D Keypoint-Based Registration of Tls Point Clouds. *ISPRS Ann. Photogramm. Remote Sens. Spat. Inf. Sci. II-3/W5*, 121–128.
<https://doi.org/10.5194/isprsannals-II-3-W5-121-2015>
- Vosselman, G., Maas, H.-G., 2010. Airborne and Terrestrial Laser Scanning. *Current*. <https://doi.org/10.1080/17538947.2011.553487>
- Weinmann, M., 2013. Visual Features—From Early Concepts to Modern Computer Vision, *Advances in Computer Vision and Pattern Recognition*. <https://doi.org/10.1007/978-1-4471-5520-1>
- Westoby, M.J., Brasington, J., Glasser, N.F., Hambrey, M.J., Reynolds, J.M., 2012. “Structure-from-Motion” photogrammetry: A low-cost, effective tool for geoscience applications. *Geomorphology*. <https://doi.org/10.1016/j.geomorph.2012.08.021>
- Z+F, 2017, <https://www.zf-laser.com>