AUTOMATIC MARKER-FREE REGISTRATION OF TERRESTRIAL LASER SCANS USING REFLECTANCE FEATURES

Jan BÖHM, Susanne BECKER
Institute for Photogrammetry, Universitaet Stuttgart
Jan.Boehm@ifp.uni-stuttgart.de

KEY WORDS: Terrestrial Laser Scanning, Registration, Orientation, Feature Points, SIFT

ABSTRACT
In our paper we explore the application of the SIFT method on terrestrial laser scan data. The energy of the backscattered laser light is represented in a reflectance image. This representation of the data can be directly used with standard SIFT implementations. We show that the extracted features are good candidates for registration. Furthermore, we show that feature matching based on the feature descriptor gives a sufficient initial configuration for registration of terrestrial laser scans. From the resulting set of matches, we compute the rigid body transform for a pair-wise registration. We show that the approach can be extended to not only automatically register terrestrial laser scans, but also automatically register intensity images taken with a separate camera to the laser data.

1. INTRODUCTION

Commonly several separate stations are required to entirely capture the geometry of a scene with a terrestrial laser scanner. A crucial step in the data processing pipeline is the alignment of the stations into a common reference frame, usually referred to as registration. Many approaches have been proposed to address this issue. They differ in the prerequisites on the scene, prior knowledge, extra sensor measurements and the level of automation and robustness.

The standard method for registration is marker-based registration. Either several markers have to be placed in the overlapping areas of the scans and measured with the laser scanner or their coordinates have to be pre-determined with a separate measurement instrument, commonly a total station. The marker-based registration is expected to achieve the highest accuracy. However, this approach requires extra effort for the placement and measurement of the targets. Furthermore, the placement of targets is prohibited for certain sites.

Direct orientation of terrestrial laser scans has been proposed by Schuhmacher and Boehm (2005). They include GPS, compass and pan-tilt sensors to directly measure position and orientation of the laser scanner and thus determine the absolute pose of the data. While this approach not only solves the registration problem but also provides georeferencing of the data, it heavily depends on the quality of the additional sensors. For low-cost sensors, the approach is mainly suited to provide initial approximations for latter refinement.

Several iterative methods are known for point cloud registration. The most important group of algorithms is the Iterative Closest Point algorithms originally developed by Besl and McKay (1992). Several extensions and improvements to the original algorithm have been proposed. 3D least squares matching developed by Akca and Gruen (2005) can also be regarded to be part of this group of algorithms. Another iterative approach is the Normal Distribution Transform originally developed by Biber (2003) for two-dimensional data, and later extended to three-dimensional data. The iterative algorithms obviously require prior knowledge on the pose of the data. Currently no definite comparison exists on the convergence radius of the iterative algorithms.

The Extended Gaussian Image (EGI) for registration of laser scans was proposed by Dold (2005) and Makadia et al. (2006) based on previous ideas. While the EGI originally is only useful for

finding the orientation, this information can be used to reduce the search for the full pose to a search for the translation in a three-dimensional parameter space.
Feature-based registration approaches extract a certain type of features from the measured data and attempt to match features in-between the separate scans. From close-range scanning many curvature-based methods for feature point extraction are known. However, these methods fail in the case of urban scenes where planar surfaces are dominating. Planar segmentation for feature extraction is used by Dold and Brenner (2006) and by von Hansen (2006). The extracted planar patches are matched across the scans to retrieve the pose parameters. Intensity features extracted from a co-registered camera were proposed by Wendt (2004).
In image-based modeling the Scale Invariant Feature Transform (SIFT) proposed by Lowe (2004) is a very popular operator for the extraction of features. The operator works on an image pyramid and computes the difference-of-Gaussian for each level. Local extrema are candidate features. An invariant representation of the area around the feature points is used as a feature descriptor. The features have shown to be largely invariant to scale changes and varying illumination. Obviously, such robust features are also interesting for the registration of terrestrial laser scans. Bendels et al. (2004) have introduced the SIFT method for registration of range data. They use data acquired with a triangulation scanner, which provides a color image for each range data set. Feature points are extracted from the color image and corresponding surface elements are extracted from the range data. The centers of gravity of these surface elements are used as tie points for registration. Here we explore the application of the SIFT method on laser reflectance data. Since scans are taken in an approximately regular polar grid, the collected reflectance data can be represented in the form of an image. Therefore the data can be directly used with standard SIFT implementations.

1.1 Outline

In section 2, we show that the extracted features are good candidates for registration. Furthermore, we show that feature matching based only on the feature descriptor is sufficient for registration of terrestrial laser scans. The pairs of matched feature points obviously contain a high degree of false matches. In our experiments, the number of outliers exceeds 90 percent. Including the geometry of the tie points, we are able to separate the outliers from the correct matches. From the resulting set of matches, we compute the rigid body transform for a pair-wise registration.
In section 3, we show that the approach can be extended to not only automatically register terrestrial laser scans, but also automatically register intensity images taken with a separate camera to the laser data. This offers the possibility to strengthen the network geometry or simply to automate texturing from hand-held cameras.
We demonstrate the results throughout this paper based on the same example dataset. The data was acquired with a Leica HDS 3000 pulsed time-of-flight terrestrial laser scanner. The test scene is an old farm building, a typical application for terrestrial laser scanning. The data was acquired at an approximate sampling distance of 2 cm on the object's surface. The data set is available for download from our web pages.

2. MARKER-FREE REGISTRATION OF TERRESTRIAL LASER SCANS

2.1 Feature extraction

The energy of the backscattered laser light is recorded by almost every terrestrial laser scanner system available. Due to the raster-wise sampling of the point cloud this backscattered energy can be represented as a matrix per point, or in other words as an image. We will refer to this image as the reflectance image. This image can be interpreted similarly to an intensity image.

Figure 1: The laser return image before (left) and after (right) histogram equalization.

However, the backscattered laser light is a signal of high dynamic range. From almost no return due to low reflective, far away surfaces to direct reflection from retro-reflective material, such as license plates, the strength of the return varies over a large range, especially for pulsed time-of-flight systems. For the Leica HDS 3000 this fact is accounted for by storing the reflectance information as 10-bit numbers. Since most of the displays and many standard image processing tools are still designed for 8-bit data, we decide to convert the reflectance image to 8 bit. Instead of simply scaling the data, we apply histogram equalization prior to scaling. This dampens the areas of very low and very high reflectance and makes the image appear more like a typical intensity image. An example of the image pre-processing is shown in Figure 1. For the processing pipeline, this has the advantage that we can rely on a standard implementation for feature extraction and do not have to alter parameters for either reflectance or intensity images. For the feature extraction we chose the Scale Invariant Feature Transform (SIFT) proposed by Lowe (2004). Our application to range data registration is similar to Bendels et al. (2004); however, we apply the feature extraction to backscattered laser light. We us an open implementation of SIFT provided by Vedaldi (2007). Figure 2 shows some examples for the location of the extracted feature points.

2.2 Pair-wise Registration

Besides the location of features, the SIFT method also computes a local feature descriptor, which is based on the local gradient distribution. This feature descriptor allows for a simple scheme of matching feature points across images without taking into account the image geometry. One simply searches for the best match of the feature descriptor by computing the Euclidean distance.

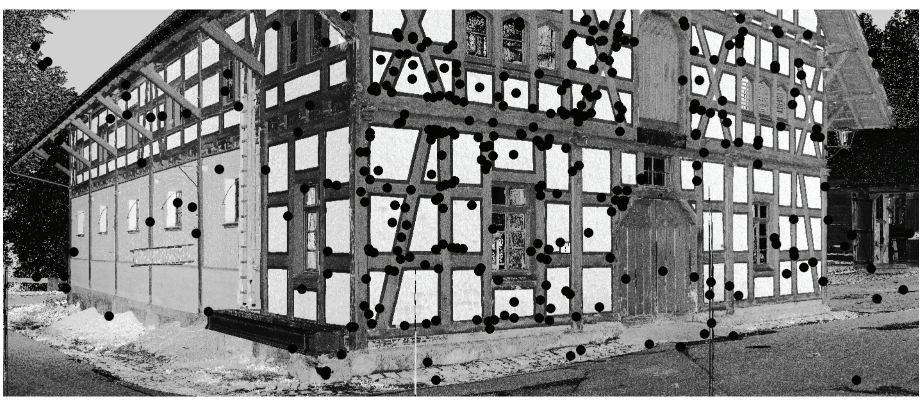

Figure 2: The feature points extracted and matched using the SIFT method.

It is obvious that the pairs of matched feature points contain a high degree of false matches. For one this is due to the nature of the feature descriptor. However, a large part of the false matches also is caused by the symmetry and self-similarity of the façade structure. Repetitive elements such as windows and beams, which are especially dominant in the example scene, cause false matches, when the geometry of the image is ignored. This situation is shown in Figure 3, where reflectance images from two adjacent stations are shown together with two matches based on the feature descriptor. While the matches are locally correct, they violate the global geometry.
To exclude these false matches from the registration we have to include the geometric relationship of the key points. This is done by a RANSAC filtering scheme. Randomly a sample of three point pairs is drawn from all SIFT matches. From the pair of three points a rigid body transformation is computed. All SIFT matches are checked against this transformation for consensus. The sample with the largest consensus is selected for registration. Such a sample of the minimum configuration of points is shown in Figure 4 where the data was imported as tie points to Cyclone and shown together with the point cloud.
In our example 7772 key points are extracted from the first image using the SIFT. Of those 552 are matched to corresponding key points in the second image based on the descriptor. Only 69 are confirmed as valid 3D corresponding tie points using RANSAC. This indicates 87.5 percent outliers in the matched key points.

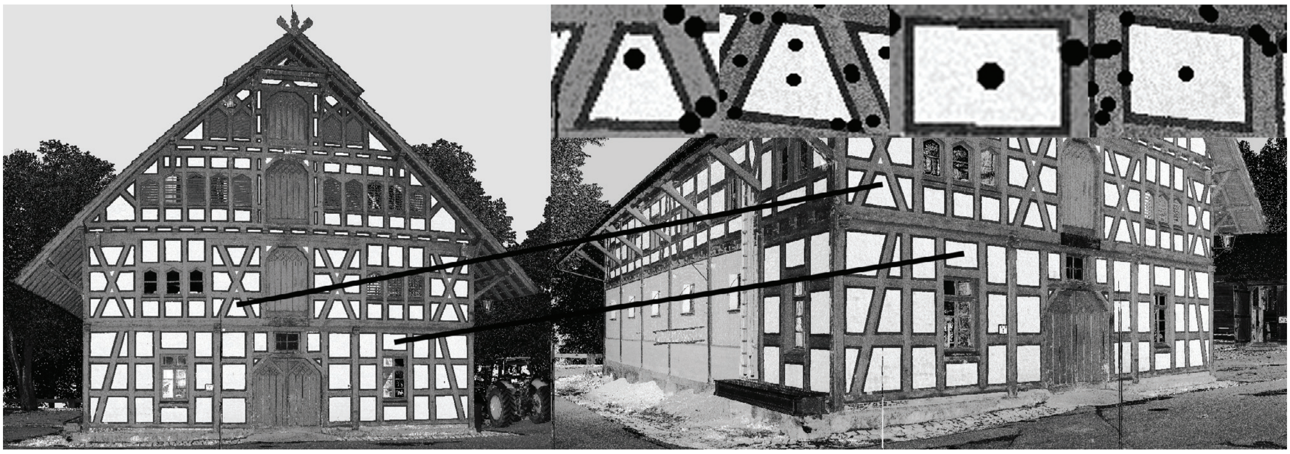

Figure 3: The reflectance images of two adjacent terrestrial laser scan stations and falsely matching key points.

Figure 4: Three corresponding feature points representing the best consensus found through RANSAC imported as tie points in Cyclone for registration.

	Manual Registration	RANSAC Best Sample	RANSAC Consensus
Translation X (m)	-11.765	-11.769	-11.767
Translation Y (m)	7.096	7.100	7.096
Translation Z (m)	-0.298	-0.285	-0.248
Rotation (deg)	28.039	28.048	28.039
Residuals (m)	0.008	0.017	0.054

Table 1: Registration parameters computed with different approaches. The rotation is given with respect to the approximated vertical axis; the residuals indicate the mean absolute error of corresponding points after registration.

From the set of filtered matches, we can compute the rigid body transformation for final registration. In order to assess the accuracy of the automatic registration we have also performed a manual registration, which serves as a benchmark. To obtain comparable results, the final registration was always computed using Cyclone. For the automatic registration, we report two different results. One is computed from only the best sample, consisting of three point pairs. The second result is computed from all 69 points in consensus with this sample. The numbers are given in Table 1. Obviously, the automatic registrations give the same results as manual registration.

2.3 Time Considerations

In order to assess the applicability of the approach we also have to consider the runtime required to complete the registration. The bulk of computation time is spent on the SIFT computation. Extracting the key points needs 15 seconds for each image on average. Matching the SIFT key points takes just under 10 seconds per pairing. Filtering the matches using RANSAC is done in less than one second per matched pair. Since the actual adjustment for the registration is performed in Cyclone, which requires manual import of the data, the exact timing of the adjustment is not possible. However, we can conclude that the pair-wise registration is computed well under one minute for the dataset introduced above. This compares very favorably to manual registration, which requires manual identification and measurement of natural tie points. Furthermore the speed of SIFT is merely an implementation issue. Accelerated versions of SIFT using programmable graphics hardware have been reported for near real-time applications (Sinha, 2007).

3. MARKER-FREE REGISTRATION OF IMAGES

Registration of images to terrestrial laser scans typically involves manual effort if no specially designed targets are used for the control points. The reason is that object points with known 3D coordinates have to be manually identified in the images by a human operator. The idea to automate this process is linking the images to the LIDAR data by a matching process, which is similar to the automatic registration of terrestrial laser, scans as described above.

However, images generated from laser reflectance considerably differ from images that have been captured by photo cameras. On the one hand, the measured laser intensities represent the reflectivity of the measured surface only in a narrow wavelength range (for example 532 nm for the HDS 3000). Furthermore, the viewing direction and the direction of illumination are identical in case of laser scanning. By contrast, photo cameras usually work with ambient light sources, which may cause shadow areas on the object and therefore lead to gray value edges in the photograph. On the other hand, the laser image is not based on central projection but on polar geometry. Thus, straight 3D lines may appear curved in the reflectance image. Another aspect is the sampling distance, which is often much higher for a laser scan compared to the spatial resolution of a photo captured by a camera. For these reasons, the determination of point correspondences between a laser reflectance image and a photograph requires an algorithm, which is insensitive to changes in illumination and scale and uses region descriptors instead of edge detectors.

Figure 5: Key point correspondences for a photograph (left) and a laser reflectance image (right).

Figure 5 depicts the laser reflectance image (right) and a photograph captured by the NIKON camera (left). In order to have similar intensity values in both images, only the green channel of the photograph has been considered for the determination of corresponding points. The resulting key points were extracted and matched by means of the SIFT algorithm. Wrong matches are removed by a RANSAC based computation of a closed form space resection (Zeng and Wang, 1992). For this purpose, the SIFT point correspondences can be used as control point information since each pixel with a valid laser reflectance value refers to the 3D coordinates of the related sample point. The number of inlier matches is about 20%. They are represented by the red dots and lines in Figure 5.

Obtaining the exterior orientation of a calibrated camera relative to the laser scan data allows a fully integrated adjustment of the involved data as proposed by Al-manasir and Fraser (2006), which is expected to strengthen the network geometry. A further application of oriented intensity images is the high quality automated texturing of terrestrial laser data as proposed by Alshawabkeh and Haala (2005).

3. CONCLUSIONS

We have shown a flexible method for the automatic marker-free registration of terrestrial laser scans. The method uses SIFT key points extracted from the reflectance data of the scans. The method can be used to register data at accuracy comparable to that of manual registration using natural tie points. It can also be extended to automate the orientation of images captured from a hand-held camera with respect to the laser data.

REFERENCES

1. D. Akca and A. Gruen. A flexible mathematical model for matching of 3d surfaces and attributes. Proceedings of SPIE - The International Society for Optical Engineering, 5665:184–195, 2005.
2. K. Al-manasir and C. S. Fraser. Registration of terrestrial laser scanner data using imagery. Photogrammetric Record, 21(115):225–268, 2006.
3. Y. Alshawabkeh and N. Haala. Automatic Multi-Image Photo-Texturing of Complex 3D Scenes. CIPA 2005 IAPRS Vol. 34-5/C34, pp. 68-73, 2005.
4. G. Heinrich Bendels, P. Degener, R. Wahl, M. Körtgen, and R. Klein. Image-based registration of 3d-range data using feature surface elements. In N. Silberman F. Niccolucci Y. Chrysanthou,

K. Cain, editor, The 5th International Symposium on Virtual Reality, Archaeology and Cultural Heritage. VAST, The Eurographics Association, 2004.

5. P. J. Besl and N. D. McKay. A method for registration of 3-d shapes. IEEE Transactions on Pattern Analysis and Machine Intelligence, 14(2):239–256, February 1992.
6. P. Biber. The normal distributions transform: A new approach to laser scan matching. In IEEE International Conference on Intelligent Robots and Systems, volume 3, pages 2743–2748, 2003.
7. C. Dold. Extended gaussian images for the registration of terrestrial scan data. In G. Vosselman and C. Brenner, editors, Proceedings of the ISPRS Workshop Laser scanning 2005, Enschede, the Netherlands, 2005.
8. C. Dold and C. Brenner. Registration of terrestrial laserscanning data using planar patches and image data. In IEVM06, 2006.
9. M. A. Fischler and R. C. Bolles. Random sample consensus: A paradigm for model fitting with applications to image analysis and automated cartography. Communications of the ACM, 24(6):381–393, June 1981.
10. D. G. Lowe. Distinctive image features from scale-invariant key points. International Journal of Computer Vision, 60(2):91–110, 2004.
11. A. Makadia, A.I. Patterson, and K. Daniilidis. Fully automatic registration of 3d point clouds. In Computer Vision and Pattern Recognition, volume 1, pages 1297– 1304, 2006.
12. S. Schuhmacher and J. Boehm. Georeferencing of terrestrial laserscanner data for applications in architectural modeling. In 3D-ARCH 2005: Virtual Reconstruction and Visualization of Complex Architectures, volume XXXVI, PART 5/W17 of International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2005.
13. Sudipta N. Sinha, Jan-Michael Frahm, Marc Pollefeys and Yakup Genc. Feature Tracking and Matching in Video Using Programmable Graphics Hardware, Machine Vision and Applications (MVA), 2007.
14. A. Vedaldi. An Implementation of the Scale Invariant Feature Transform, UCLA CSD Tech. Report N. 070012, 2007
15. W. von Hansen. Robust automatic marker-free registration of terrestrial scan data. In PCV06, 2006.
16. A. Wendt. On the automation of the registration of point clouds using the metropolis algorithm. In XXth ISPRS Congress: Geo-Imagery Bridging Continents, 2004.
17. Z. Zeng and X. Wang. A general solution of a closed-form space resection. PE&RS, Vol. 58, No. 3, pp. 327-338, 1992.