Automatic Filtering of Terrestrial Laser Scanner Data from Cylindrical Tunnels

Cornelis STAL, Timothy NUTTENS, Denis CONSTALES, Ken SCHOTTE, Hans DE BACKER, Alain DE WULF, Belgium

Keywords: Terrestrial laser scanning, filtering, tunnel, deformation measurement

SUMMARY

In 2010, the construction of a double tube railway tunnel started under the Scheldt River in the port of Antwerp, Belgium. During the construction process, which will end in 2013, a systematic survey campaign is organized to perform an ovalization monitoring of selected concrete sections in both tunnels. A phase-based terrestrial laser scanner Leica HDS6100 is used to generate dense point sets of each selected section, with a spatial resolution of 4 mm or higher on specific moments in time. Based on every point set, a cross-section is calculated, which is compared with previously calculated cross-sections, to determine the deformation of the tunnel section. The location of the cross-section is defined by a materialized reference target and this cross-section is perpendicular to the axis of a best fit cylinder. This best fit cylinder is defined by a point set of of 7 concrete circular curved elements. After a point filtering has been performed, the cylinder is calculated by a least square fitting algorithm on the point set of each tunnel section.

In this article, a procedure is presented to perform the cylinder fitting and data cleaning steps automatically in an iterative process. This increases the efficiency and the performance of processing the tunnel data. Each step in the iteration starts with the calculation of a best fit cylinder, based on all remaining points in that step. The parameters of this cylinder are estimated using the Levenberg-Marquardt algorithm. This algorithm is an optimizer for least squares problems, based on global non-linearity, and also on the local linearity of the data set, as implemented in the Gauss-Newton parameter estimation method and in the negative descent method. The combination of both methods in the Levenberg-Marquardt algorithm will facilitates the convergence of the final parameters. The resulting cylinder parameters are then used to rotate and translate the point set, and to calculate the distance between each point to the surface of the hypothetical cylinder. A threshold is used to evaluate these distances and to remove points located at a significant distance from the cylinder. A new cylinder is then calculated, based on the remaining points, and outliers are removed again, based on the adjusted cylinder parameters and decreased threshold. The process iterate until this threshold has reached a predefined minimum. Similar to the previously presented manual procedure for the filtering of tunnel data, the final distances are projected as a function of the zenithal angle for cross-section generation. This cross-section will be compared with a manually generated cross-section, in order to validate the results of the proposed procedure and its applicability in tunnel data processing.

Automatic Filtering of Terrestrial Laser Scanner Data from Cylindrical Tunnels

Cornelis STAL, Timothy NUTTENS, Denis CONSTALES, Ken SCHOTTE, Hans DE BACKER, Alain DE WULF, Belgium

1. INTRODUCTION

Terrestrial laser scanning (TLS) is a well know 3D data acquisition technique for various applications, such as archaeology and cultural heritage. The technique is also extraordinarily useful in the field of civil engineering, where it can be applied for tunnel deformation measurements. A terrestrial laser scanner is able to measure millions of points in a limited timeframe, by measuring distances and corresponding horizontal and vertical angles. These distances can be calculated by registering the time of flight of the emission and reception of an electromagnetic signal or phase delays. The performance of pulse based scanners and phase based scanners for tunnel measurements is discussed by Nuttens *et al.* (2010).

In this article, a workflow for processing point sets from newly installed tunnel segments is presented. The main focus of this research is the improvement of the degree of automation of the filtering and cross-section generation process. The development of filtering procedures for terrestrial laser scanner data is more complex in comparison to filtering procedures for airborne laser scanner data. This is mainly due to the increased dimensional complexity of terrestrial laser scanner applications in comparison with airborne applications. Many filtering algorithms were presented for airborne laser scanner data during the last two decennia (Krzystek, 2003; Sithole & Vosselman, 2003). Most of these algorithms use a horizontal reference plane for the entire data set and for the local neighborhood of the processed points, reducing the filtering problem to a 2.5D problem. Besides, terrestrial laser scanner data suffers from the presence of occlusions caused by the perspective geometry of single terrestrial scans (Brodu & Lague, 2012).

In this research, the problem of filtering tunnel data can be simplified by the assumption that the tunnel sections are considered as a cylinder. Even in complex scenes, it is possible to detect these geometrical primitives in point sets, and describe them by a limited number of parameters (Rusu, *et al.*, 2008). Since this cylinder is the only object that needs to be detected and modeled, the presented filtering algorithm is actually a binary estimator of surface- and non-surface data in a 3D space. The preliminary results of this research concern control measurements only, compared to the design of the tunnel, which has a radius of 3.65 m. However, the presented algorithm should also perform well on critical measurements, and can be used for further tunnel data analysis.

2. MEASUREMENTS IN THE RAILWAY TUNNEL OF 'LIEFKENSHOEK'

The 'Liefkenshoek' rail link project is located in the center of the port of Antwerp, north of the city. The project started in 2008 by LocoRail N.V., a consortium of different construction- and engineering companies, and is commissioned by the Belgian rail road authority Infrabel N.V. The total value of the project is around €680 million. By constructing this railway connection under the Scheldt River, the harbor of Antwerp, Belgium, will improve the hinterland connections and decrease the congestion problems of the existing Kennedy railway tunnel. Up to now, the Kennedy tunnel in the south of the city is the only suitable railway connection between the two sides of the river banks and its separate harbor basins. As a result, the new connection will reduce the travelling time between the two banks, and it will anticipate on the predicted increased harbor capacity of 35% of goods by train.
The 16 km long route contains a double tunnel of 6 km, where two separate Tunnel Boring Machines (TBM) were maneuvered under the Scheldt River and the neighboring 'Kanaaldok' (Figure 1). After a given advance of the TBM, a series of eight concrete segments are placed, shaping a cylinder with a width of 1.80 m, an inner diameter of 7.30 m and an outer diameter of 8.10 m. During the entire construction process, a selection of fourteen sections or rings in each tunnel were measured by laser scanning, in order to perform ovalization measurements. Eight of these sections in each tunnel were selected for tension measurements, also using strain gauges in order to monitor deformations of the separate segments (Schotte *et al*., 2010).

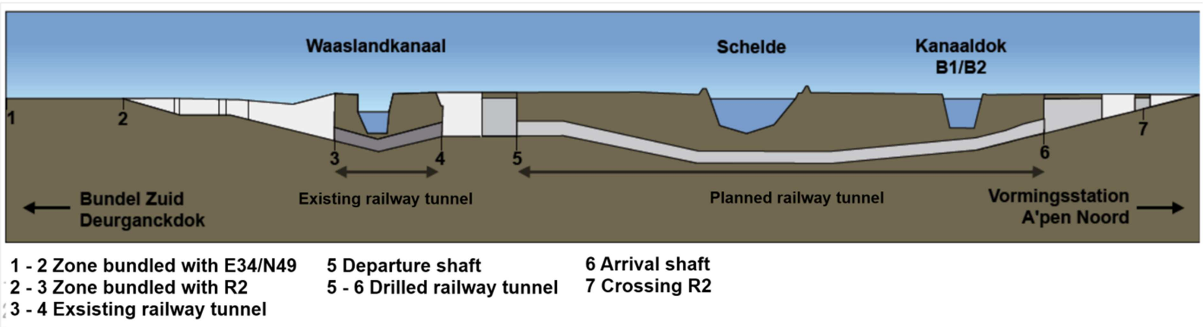

Figure 1: Cross section of the entire 'Liefkenshoek' railway project (source: www.tucrail.be)

Concerning the ovalization measurements, each selected section was measured immediately after placement. During the following four weeks, the same section was re-measured on a weekly basis, and thereafter twice on a monthly basis. During the first measurement, a series of black-and-white targets were mounted on the section surface, to enable the relocation of the position of the cross-section in the point set.

3. TERRESTRIAL LASER SCANNING AND DATA PREPARATION

All point sets, used for ovalization measurements of the different sections in this projects, are measured using a phase based Leica HDS 6100 laser scanner. Based on previous research in this project (Nuttens *et al.*, 2010), this laser scanner, with an experimental standard deviation of 0.4 mm, for very short distances (< 6 m), is very suitable for deformation measurements of tunnel sections. A setup of the terrestrial laser scanner inside the TBM is illustrated in Figure 2.

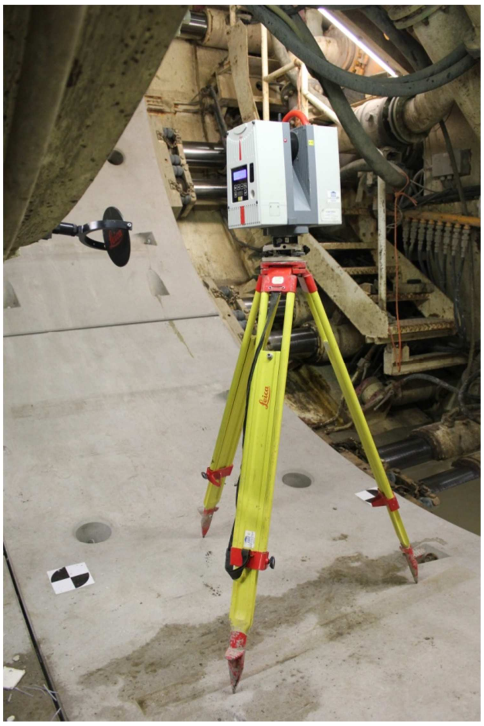

Figure 2: Terrestrial laser scanner in the TBM.

The raw data sets are stored in a zfs-file, which is the native binary file format for Leica and Zoller & Frölich laser scanners. In order to process these data files, the Leica Cyclone point processing software is used. This software package is used for the filtering and processing of the laser scan data and the generation of a cross-section for each measurement, which is then further processed in a CAD package and calculation software. The otherwise time consuming processing in Leica Cyclone is now being replaced by this automated workflow, presented in this paper. After loading the laser scanning data in Leica Cyclone, the coordinates of the master target are determined. This target is marked by a black-and-white label and the software is able to determine the coordinates with a high accuracy, by adjusting the manually indicated initial coordinates of this point. This adjustment is possible by the high local intensity differences and by recognition of the shape of the target (Pfeifer & Briese, 2007). After the selection of the master target, a large number of points, not belonging to the tunnel section, are roughly deleted from the point set. This reduces the number of points in the data set from approximately 40 million for one single scan to just a few million, which significantly decreases the processing time. Finally, the point set is exported as an ascii-file, where each line contains the (x,y,z)-coordinates, as well as intensity values of each single point.

4. AUTOMATIC DATA CLEANING

4.1. The Levenberg-Marquardt algorithm

The calculation of a cross-section through the unambiguously determined center of the master target requires a correct orientation of the point set, as well as the calculation of the parameters of a hypothetical best fit cylinder. At first, the orientation parameters were determined by performing an orthogonal linear regression. The fitted cylinder was based on the rotated point set, on its centroid and on an iterative circle, fitting perpendicular on the regression curve. These parameters were used to perform an iterative removal of points which were not located on the concrete surface. However, the results of the first tests of this procedure were not satisfying. During the acquisition of the point set, the scanner is not exactly positioned at the center of the tunnel, but more or less randomly on the floor of the tunnel without taking into account the orientation or size of the tunnel. The misalignment of the scanner with respect to the hypothetical cylinder resulted in irregular point densities at the concrete surface. As a result, the use of linear cylinder fitting algorithms is not sufficient.

A solid solution was found in the use of the Levenberg-Marquardt algorithm (Marquardt, 1963), as implemented in the Apache Commons Java library (Apache Commons, 2011). In general, the algorithm is used as an optimizer for least squares problems for nonlinear multivariate functions (Lourakis, 2005). The algorithm is known to give good results for parameter estimation of a cylinder based on point sets in a nonlinear way (Rusu *et al.*, 2008).
For a given point set $\mathbf{x}$ with ($\mathbf{x} \in R^3$), the parameter estimation of a cylinder will result in a set of parameters $\boldsymbol{p}$ with $\boldsymbol{p} = (\boldsymbol{x}, \boldsymbol{A}, r)$. Here, $\boldsymbol{x}$ is any point on the cylinder axis, $\boldsymbol{A}$ is the rotation matrix and r is the cylinder radius (Lourakis, 2005).
The algorithm is a combination of the Gauss-Newton method (or Taylor-series method) and the steepest descent method (or gradient method). The Gauss-Newton method on the one hand will estimate the optimal parameters iteratively under the assumption of global nonlinearity, but local linearity by the construction of a Taylor-series. The steepest descent method on the other hand will look for a local minimum of the parameters by iterating as a function of the negative gradient around a data point. Both methods have disadvantages in respectively the possible failure of convergence and very slow convergence (Marquardt, 1963), but their advantages (respectively speed and guarantee of success) are implemented in the Levenberg-Marquardt algorithm (Lourakis, 2005). The algorithm is clearly explained in pseudo-code by Shakarji (1998).

4.2. Implementation in the filter

The ascii-file, generated by the used point processing software, is loaded in a Java-application and each point is stored in a list, represented by an array of doubles for the coordinates, an integer for the intensity values and a separate double for later distance calculations and comparisons. In the developed user interface, the initial cylinder radius r is set to 3.65 m, which is the design radius of the tunnel. Initial parameters for the rotation matrix $\boldsymbol{A}$ are all set to zero during the initialization phase.

Each time the parameters of the temporary cylinder are estimated, and the rotation and translation parameters are used to calculate the new coordinates of each point in the entire point set. During the parameter estimation, the shortest distance between each point and the surface of the cylinder is calculated and is assigned to that given point. Since the z-axis of the original point set is fixed with the zenith, the cylinder can be described by only six parameters and no rotation around the axis of the cylinder is required. The squared distances are used in a removal procedure, where each point is evaluated against a given threshold. This threshold is based on the perpendicular maximal allowed squared distance between a point and the surface of the cylinder. The resulting points are used in the next step of the iterative process, where the parameters of the cylinder are re-estimated based on the remaining points. After each step, the threshold is reduced until a given value is reached. The result of this procedure is a filtered point set, stored in a separate file, where each point represents the concrete surface of the section.

The filtered point set can be used for further tunnel deformation analysis, with respect to the calculated cylinder axis and design value for the tunnel radius. In order to perform this analysis, the filtered points are inserted in a kD-Tree for neighborhood analysis. During the initialization of the filtering process, the original coordinates of the master target are inserted for the definition of a cross-section. These coordinates are rotated and translated into the new coordinate system and are finally projected on the axis of the cylinder, resulting in the base point of the cross-section. A set of 4000 hypothetical points are calculated, with a distance of 3.65 m to this base point and a zenithal increment of 0.1 gon, perpendicular to the axis of the cylinder. These points represent the local tunnel surface if no ovalization would have taken place. Thereafter, an adjusted distance between this point and the axis of the cylinder is calculated by the mean of the distances of the 5 nearest points from the filtered data set. The final distance per zenithal angle is then adjusted by the mean of 5 distances left and right from this angular value (e.g. [6.3 → 6.8 → 7.3] gon). This will result in a two dimensional representation of distances as a function of the zenithal angle.

5. RESULTS

The results of the proposed procedure are illustrated by the automatic filtering of three randomly selected measurements. From these point sets, the coordinates of the master targets were determined in the original local coordinate system. From the approximately 40 million points in the entire point set, only a few million points were retained, located in a roughly selected cross-section. These points are processed by the developed Java application. As demonstrated in Figure 3 and Figure 4, the algorithm converges properly to a sudden asymptote after 13 iterations (approximately 3.0 mm RMS (Figure 3) and 2.5 µm error on the radius calculation (Figure 4)).

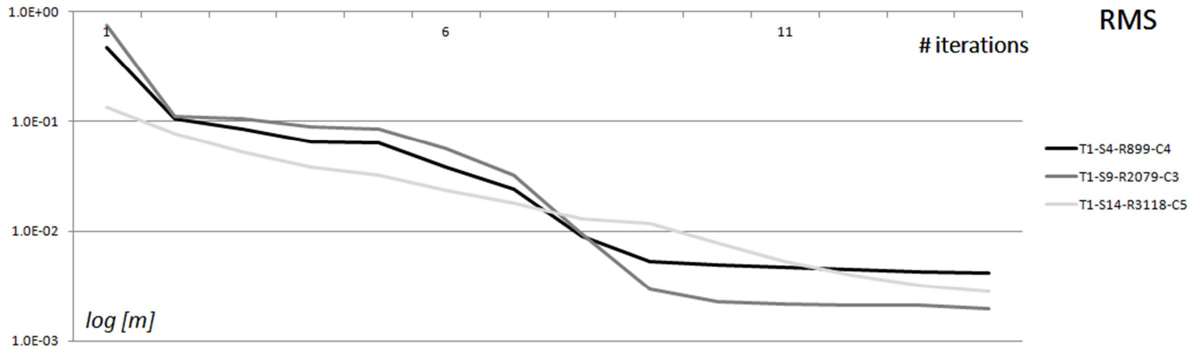

Figure 3: Evolution of the RMS of the cylinder over different iterations

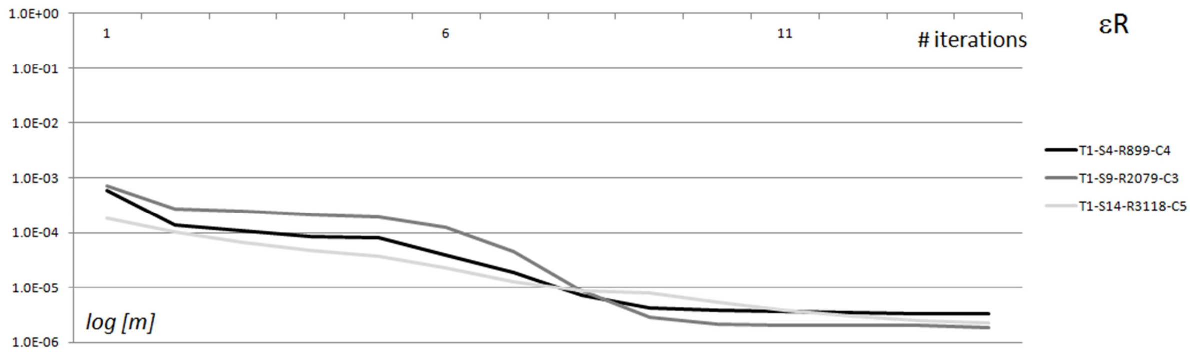

Figure 4: Evolution of the error of the calculated radius over different iterations

As mentioned above, the filtered point set is used to generate a cross-section through the master target, and perpendicular to the axis of the cylinder. The difference between the true radius of the tunnel and the design is calculated for each 0.1 gon along this cross-section. These values are also compared with previously generated cross-sections of the same measurement, following the 'manual' processing workflow. The results of this comparative study are presented in table 1. The given values are related to the deviations from the design radius of the tunnel (3.65 m). Based on this table, it becomes clear that the results of the manual and automatic filtering seem to be comparable: the mean difference between the automatically and manually generated data sets is limited to millimeter level. The MAE (Mean Absolute Error) is defined as

$$\frac{1}{n}\sum_{i=1}^{n}|d_i - \bar{d}|$$

and the RMSE (Root Mean Square Error) is defined as

$$\sqrt{\frac{1}{n}\sum_{i=1}^{n}(d_i - \bar{d})^2}$$

with d_i as the difference d at a given angular index i and $\bar{d}$ as the mean differences over the entire set. The MAE is defined by the deviations from the mean difference, where the direction of the deviation is eliminated. The fact that the MAE and RMSE of the differences between the two data sets are comparable to the other data sets, indicates that all data sets have the same distribution parameters, assuming normal distributions.

Table 1: Overview of the filtering results for 3 cross-sections (values in m)

		Min	Max	Mean	StDev	MAE	RMSE
T1-S4-R899-C4	Automatic	-0.0113	0.0039	-0.0031	0.0047	0.0038	1.680E-05
	Manual	-0.0100	0.0089	-0.0021	0.0047	0.0041	1.717E-05
	Difference	-0.0080	0.0067	-0.0010	0.0040	0.0034	1.235E-05
T1-S9-R2079-C3	Automatic	-0.0104	0.0018	-0.0018	0.0026	0.0018	4.280E-06
	Manual	-0.0019	0.0010	-0.0001	0.0005	0.0004	1.851E-07
	Difference	-0.0114	0.0024	-0.0017	0.0027	0.0019	4.496E-06
T1-S14-R3118-C5	Automatic	-0.0066	0.0080	-0.0010	0.0025	0.0018	4.394E-06
	Manual	-0.0068	0.0032	-0.0013	0.0027	0.0022	5.138E-06
	Difference	-0.0019	0.0127	0.0003	0.0020	0.0012	4.289E-06

In order to statistically evaluate the results, the significance of the difference between the automatically and manually filtered data sets is tested by performing a two sided t-test (Moore *et al.*, 2009). The null hypothesis (H_0) and alternative hypothesis (H_A) are set up to assess whether the differences significantly differ from 0.

$$H_0\colon \bar{d} = 0 \qquad H_A\colon \bar{d} \neq 0$$

In words, the null hypothesis states that the average difference between the automatically and manually filtered dataset is equal to zero, in contrast to the alternative hypothesis, which states that the average difference between the two datasets is unequal to zero. The statistic

$$H_0\colon t_{\left(v,\frac{\alpha}{2}\right)} \leq \frac{\bar{d} - 0}{s_{\bar{d}}}\sqrt{n} \leq t_{\left(v,1-\frac{\alpha}{2}\right)}$$

is used to reformulate the null hypothesis. The results of the t-test are demonstrated in table 2.

Table 2: One sample t-test with evaluation value 0

	t	df	Sig. (2-tailed)	Mean Difference	95% Confidence Interval of the Difference	
					Lower	Upper
T1-S4-R899-C4	-1.908	62	0.061	-0.0010	-0.0020	0.0000
T1-S9-R2079-C3	-4.393	50	0.000	-0.0016	-0.0024	-0.0009
T1-S14-R3118-C5	1.042	58	0.302	0.0003	-0.0003	0.0008

Based on these values, the null hypothesis can be stated for the first and the last data sets (respectively T1-S4-R899-C4 and T1-S14-R3118-C5) with a significance level of 0.05. For these sets, at least 95% of the possible difference between the two sets is caused by a random effect. However, the null hypothesis is rejected for data set T1-S9-R2079-C3 with a 95% level of confidence, and a significant difference between the two sets is noticed. Although the algorithm converged to low RMS values, the final results of the algorithm are not statistically identical to the manually filtered cross-sections.

6. DISCUSSION AND FURTHER RESEACH

The use of the Levenberg-Marquardt algorithm has proven to be very useful for the automatic filtering of cylindrical tunnel point sets. Although most of the results were satisfactory, the implemented procedure will not necessarily result in cross-sections that are statistically identical to the manually processed point sets. Next to possible misclassification in the manually processed point set, this is mainly caused by the fact that only a very small amount of points from the original point set is taken into account. As a result, the best fit cylinders have a much bigger radius than the height of the cylinder. This may result in the incorrect calculation of rotation parameters, and this effect will increase after each iteration.

Solving this issue will require further improvement of the Java application. Currently, the procedure takes each point from the point set for the cylinder fitting, and requires much computation time and computer capacity. Better cylinder fitting will be performed by using a bigger cylinder radius / height ratio, thus a better spatial variation of the points. In order to reduce the processing time, the best fit cylinder will be computed by taking random samples from the entire point set.

7. CONCLUSION

Terrestrial laser scanning is intensively used for deformation measurements during the ‘Liefkenshoek’ rail link project. Earlier research has demonstrated that the systematic and random error of the used terrestrial laser scanner data are less than 1 mm. Processed point sets from these measurements were used for the generation of cross-sections through a master target. Besides, these cross-sections should be perpendicular on the axis of a best fit cylinder. In this paper, a procedure is presented to filter these point sets automatically in an iterative process, using the Levenberg-Marquardt algorithm. The results of this procedure are promising, since the mean difference between automatically and manually filtered point sets are limited to millimeter level. However, based on a t-test for the analysis of the results, further optimization of the procedure is still possible.

ACKNOWLEGDEMENT

This research is part of the research project “3D CAD modeling of spatial architectural volumes, using terrestrial laser scanning and airborne laser scanning”. The authors would like to express their gratitude to the Research Foundation Flanders (FWO) for funding the work presented in this paper.

REFERENCES

Apache Commons (2011). Commons-Math: The Apache Commons mathematics library [Online]. http://commons.apache.org/math/ (accessed January 27 2012).

Bodru, N. & Lague, D. (2012). 3D Terrestrial LiDAR Data Classification of Complex Natural Scenes using a Multi-scale Dimensionality Criterion: Applications in Geomorphology. ISPRS Journal of Photogrammetry and Remote Sensing, 68 (1), pp. 121-134.

Crosilla, F., Macorig, D., Sebastianutti, I. & Visintini, D. (2011). Point Classification by a Sequential Higher-order Moments Statistical Analysis of LiDAR Data. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 38 (5), p. 6 (on CD-ROM).

Delaloye, D., Hutchinson, J. & Diederichs, M. (2011). Accuracy Issues associated with LiDAR Scanning for Tunnel Deformation Monitoring. Proceedings of the 14th Pan-Am CGS Geotechnical Conference, Toronto Canada, October 2-6, p. 6 (on CD-ROM).

De Wulf, A. (2010). Rapport Ovalisatiemetingen Liefkenshoekspoorverbinding (in Dutch: *Report Ovalization Measurements Liefkenshoek Railway Connection)*. Internal report.

Eberly, D. (2008). Distance Between Point and Circle or Disk in 3D / Least Square Fitting of Data / Fitting 3D Data with a Cylinder / Rotation Representations and Performance Issues, Geometric Tools, LCC, http://www.geometrictools.com

Krzystek, P. (2003). Filtering of Laser Scanning Data in Forest Areas Using Finite Elements. International Archives of Photogrammetry and Remote Sensing, 34 (3), p. 6 (on CD-ROM).

Levenberg, K. (1944). A Method for the Solution of Certain Non-linear Problems in Least Squares. Quarterly of Applied Mathematics, 2 (2), pp. 164-168.

Lourakis, M. (2005). A Brief Description of the Levenberg-Marquardt Algorithm Implemented by Levmar. Foundation of Research & Technology – Hellas (Forth), Greece, pp. 1-6.

Marquardt, D. (1963). An Algorithm for Least-Square Estimation of Nonlinear Parameters. Journal of the Society for Industrial and Applied Mathematics, 11 (2), pp. 431-441.

Moore, D., McCabe, G. & Craig, B. (2009). Introduction to the Pactice of Statistics. New York, NY, USA: W. H. Freeman.

Nuttens, T., De Wulf, A., Bral, L. De Wit, B., Carlier, L., De Ryck, M., Stal, C., Constales, D. & De Backer, H. (2010). High Resolution Terrestrial Laser Scanning for Tunnel Deformation Measurements. Proceedings of the 24th FIG International Congress. Sidney, Australia, April 11-16, p. 15 (on CD-ROM).

Pfeifer, N. & Briese, C. (2007). Geometrical Aspects of Airborne Laser Scanning and Terrestrial Laser Scanning. International Archives of Photogrammetry and Remote Sensing, 36 (3), pp. 311-319.

Rusu, R., Marton, Z., Blodow, N. & Beets, M. (2008). Learning Informative Point Classes from Acquisition of Object Model Maps. Proceedings of the 10th International Conference on Control, Automation, Robotics and Vision (ICARCV), Hanoi, Vietnam, December 17-20, p. 6 (on CD-ROM).

Schotte, K., De Backer, H., Nuttens, T., De Wulf, A. & Van Bogaert, P. (2010). Strain Gauge Measurements during the Assembly of the Diabolo Tunnel. Proceedings of the 34th IABSE Symposium, Venice, Italy, 24 September 22-24, p. 8 (on CD-ROM).

Shakarji, C. (1998). Least-Squares Fitting Algorithms of the NIST Algorithm Testing System. Journal of Research of the National Institute of Standards and Technology, 103 (6), pp. 633-641.

Sithole, G. & Vosselman, G. (2003). Comparison of Filtering Algorithms. International Archives of Photogrammetry and Remote Sensing, 34 (3), p. 8 (on CD-ROM).

BILBIOGRAPGICAL NOTES

Cornelis Stal (°1985, Waalre, the Netherlands) is a PhD student working on the combination of airborne and terrestrial laser scanning for 3D city modeling. His special interest is in the (automatic) generation of geometric, radiometric and semantic rich 3D models, derived from irregular point sets and other spatial datasets. This means that both laser scanning as a discipline in the land survey and geo-IT (GI-systems, GI-programming, GI-management,...) are important pillars of his research.

CONTACTS

Cornelis Stal
Ghent University – Department of Geography
3D Data Acquisition Cluster
Krijgslaan 281 (Building S8)
9000 Ghent
Belgium
Tel. +32 9(0) 264 46 56
Fax. +32 9(0) 264 49 85

Email: Cornelis.Stal@UGent.be
Web site: http://geoweb.ugent.be/data-acquisition-3d