EVALUATION OF PHOTOGRAMMETRIC SUPPORT FOR PAVESCAN

Stijn Verlaar[a], **Roderik Lindenbergh**[b], **Ben Gorte**[b], **and Massimo Menenti**[b]

[a] Breijn B.V., Graafsebaan 65, Rosmalen, The Netherlands
sverlaar@breijn.nl

[b] Chair of Optical and Laser Remote Sensing, Delft University of Technology, Kluyverweg 1, Delft, The Netherlands
(r.c.lindenbergh, b.g.h.gorte, m.menenti)@tudelft.nl

KEY WORDS: Pavescan, photogrammetry, control points, road modelling, laser profiles

ABSTRACT:

Pavescan is a low cost mobile system for road modelling survey. Because of the absence of navigation sensors it has several practical drawbacks compared to most of the other mobile mapping systems, but those sensors are very expensive and do not fulfil most of the accuracy requirements. Pavescan will be more attractive if some of the practical drawbacks are reduced, while at the same time the absolute accuracy of its laser scans will reach the sub-centimetre level. This paper shows how close range photogrammetry can support Pavescan. The approach is able to achieve sub-centimetre absolute accuracy, but unfortunately not in combination with reducing practical drawbacks.

1 INTRODUCTION

Road modelling is important in the road construction industry for either making a new road design or maintenance of the existing road. New roads are often made on top of an old road. For the construction of such a road, it is of major importance how much and where there need to be milled from the old road. Milling is the removal of material (in this case asphalt) with a rotating cutter device. Measurements on the road's surface are needed to obtain a model of the road and to decide for the optimal milling amounts. Especially the accuracy in height is important, since this directly has an effect on how much and where there need to be milled. One can imagine that milling more than necessary, leads to unnecessary costs because more new asphalt is needed. Once the new road has been built, road models can check whether the road fulfils the requirements.

In the last decade, different types of mobile mapping systems emerged, (Toth, 2009, Haala et al., 2008). Most common are the systems consisting of one or more laser profilers or digital cameras for the actual measurements, combined with GPS and IMU systems for the positioning of the measurements, (Barber et al., 2008). In (Yu et al., 2007) it is demonstrated how a laser mobile mapping system can detect cracks of a few millimeters in the road surface. (Jaakkola et al., 2008) describe how laser mobile mapping data is used to classify road markings and kerbstone points and to model the surface as a triangulated irregular network. Disadvantage of these type of mobile systems is that they are relative expensive to build, while it is difficult to obtain vertical elevation measurements with a relative accuracy at the millimeter level.

Pavescan measures profiles across the road by laser scanning at a series of positions, (Mennink, 2008). The Pavescan system with a schematic laser bundle is shown in Figure 1. The separate scans are linked via control points, which have to be measured in an additional survey. This is the main practical drawback of Pavescan and therefore it is desirable to reduce the required number of control points. It has already been demonstrated, (Kertész et al., 2008, Vilaca et al., 2010), that digital cameras are very useful in analyzing road surfaces. Therefore integrating close range photogrammetry was chosen as the approach to fulfil these objectives for being low cost and having high accuracy potential. Only if a high degree of automation is achieved, the approach will be interesting to use.

Figure 1: The Pavescan system with a schematic laser bundle (Illustration from (Mennink, 2008))

The main goal of integrating close range images into Pavescan is to achieve millimeter to sub-centimetre absolute accuracy in height of the road profile measured by the laser scanner, in combination with reducing the practical drawbacks of the system. This implies that the laser scans stay the main product with the images as support to strengthen the positioning of the scan. Via the approach of (close range) photogrammetry, (Kraus, 2007), the position and orientation of the camera can be retrieved for every image. Besides, for each point measured in two or more images, its 3D coordinates can be calculated in the same coordinate system as the control points. In this way, a DTM can be constructed of the area covered by the images. Regarding the control point reduction of the traditional survey, there are basically two approaches:

1. Measure the current control points in the images and calculate their 3D coordinates
2. Use the camera positions as control points

The first approach replaces the traditional measurements of the control points by image measurements. This may even be automated if the control points can be detected automatically in the images. Some control points still need to be measured with a traditional survey, because also photogrammetry requires control points to link the model to local coordinates. The second approach uses the camera positions as a reference for the scan. If a stable construction is realized, the vector between the camera and laser scanner remains constant. A drawback of this approach is that the range accuracy of the laser scanner should be taken into account. An advantage of this approach is that the orientation of the camera will also be known, and consequently the orientation of the scanner too. Thus, the position of all scan points can be calculated. Therefore the second approach is preferred. However, the first approach is used in this research to determine the accuracy, since no 'true' camera positions and orientations can be obtained as comparison.

An important question is whether the number of control points needed for photogrammetry could be smaller than the number of control points used for Pavescan. If so, one likes to know how the number of control points relates to the accuracy of the road profile to be obtained.

2 METHODOLOGY

2.1 Workflow

The images should be taken with the camera rotated such that only the road is captured (and not the trailer, compare Figure 1). A constant baseline is needed to regulate the total number of images. Often, photogrammetric images are captured based on an equal time interval. However, Pavescan is moving with a variable velocity, so images need to be captured based on an equal distance instead in order to maintain a constant baseline.

The next step is to automatically find corresponding points in the overlapping parts of the images. A set of corresponding points is referred to as a tie point. The approach used in this research is first to find per image characteristic points with SIFT (Scale Invariant Feature Transform) and then to match these characteristic points to obtain a tie point. The result from this step is a list with image coordinates of corresponding points per image pair. For this research, the matching of characteristic points is adapted from the matching algorithm provided by SIFT in order to regulate and optimize the matching procedure.

The widely accepted approach to obtain camera positions and orientations and object coordinates of all tie points is called bundle adjustment, (Kraus, 2007). It is standard to use additional software for the bundle adjustment due to its complexity. BINGO is specialized (close range) photogrammetric software that offers large control on the bundle adjustment, (Kruck, 1984). A diagram of the workflow is illustrated in Figure 2.

2.2 Error sources

During the process to the end result, several errors and other factors that influence the quality of the bundle adjustment are introduced. These errors and other factors are subdivided into those that arise before bundle adjustment and those that arise during the bundle adjustment. Before the bundle adjustment, one deals with image measurements and several control points while during the bundle adjustment one has to deal factors that are more difficult to quantify, such as the configuration of tie points and control points, as well as possible lens distortions. The possible error sources and other influential factors are shown in Figure 3.

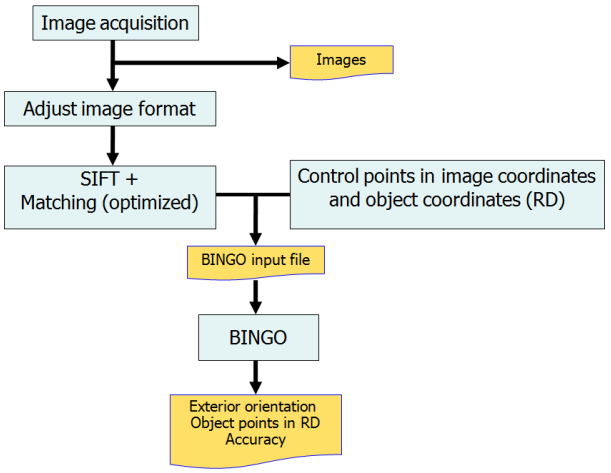

Figure 2: Workflow of working with BINGO

In general, one can say that the observations contain errors that are introduced before bundle adjustment. The other factors that have an influence on the quality of the bundle adjustment arise during the bundle adjustment. Because it is not known how to determine the influence of for instance the configuration of the control points, we could not set a theoretical accuracy of the end result.

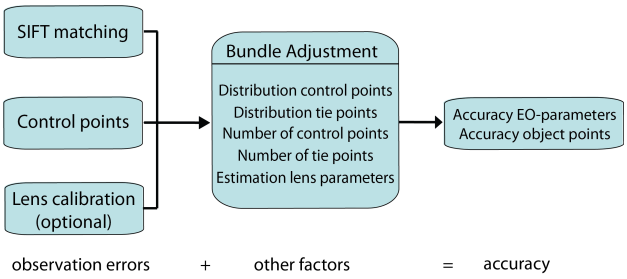

Figure 3: Error sources

2.3 Test case

A test area was chosen to examine the capabilities of combining photogrammetry within Pavescan. For this acquisition, a Canon EOS 350D with 20 mm lens is mounted on Pavescan at a height of approximately 3.6 metres. The camera is rotated manually to the back to capture just the road surface. This pitch angle is estimated by eye at 20-30 degrees. A schematic setup of the test survey is illustrated in Figure 4.

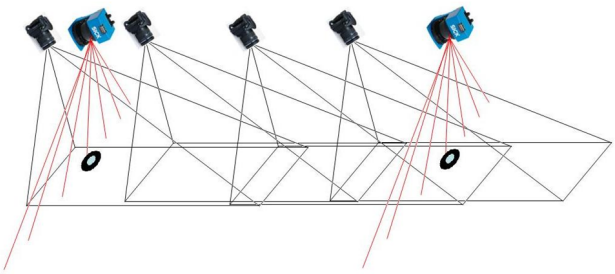

Figure 4: Schematic setup of the test survey

The test site is a parking lot with a clear crossfall of the asphalt. The images were taken in full colour and with manual focus. In

total 66 images were taken with an aimed baseline of 1.6 metres, i.e. one long strip of around 100 metres. At the beginning, halfway and at the end of the strip three nails were hit into the asphalt and measured in RD coordinates (the standard Dutch coordinate system) by tachymetry. These points, provided that they are visible in the images, can be used as control points to link the image coordinates to local coordinates. An extra control point is made every ten metres and also a scan is made by Pavescan for additional laser data. Additionally, three sets of check points are measured on the road surface, which are not marked by paint or a nail, i.e. these are not visible in the images. Each set of check points consists of six points and are approximately located on a line perpendicular to the driving direction. These points can be used for an independent check on a future result. A selection of control points is visualized in Figure 5. A least squares line was fitted through the control points to emphasize that the control points are almost on a line. Control points that are not used in the bundle adjustment are used as check points, which allows for comparison of their calculated 3D coordinates with coordinates surveyed using tachymetry. The difference is a measure of the accuracies of the object points and the bundle adjustment.

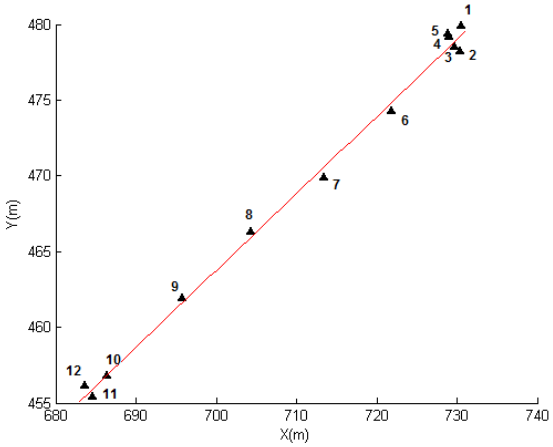

Figure 5: Configuration of control points

2.4 SIFT

Scale Invariant Feature Transform (SIFT) is a technique, (Lowe, 2004), that detects distinctive features in an image. These features are referred to as keypoints and will be used as candidate correspondences. A noteworthy property of SIFT is that the keypoints are invariant to image scale and rotation. For each keypoint a highly distinctive descriptor is computed based on its local (4 by 4 pixels) image region. Each descriptor is built from $4 \cdot 4 = 16$ small descriptors each containing gradient magnitudes of eight directions. Therefore, the descriptor vector contains $16 \cdot 8 = 128$ elements.

Furthermore, in (Lowe, 2004) a robust matching of the keypoint's descriptors is provided across a substantial range of distortion, change in viewpoint and change in illumination. This matching program is available in Matlab code. The best candidate match for each keypoint is found by identifying its nearest neighbour in the descriptor space. The nearest neighbour is defined as the keypoint with minimum Euclidean distance for the invariant descriptor vector. The Euclidean distance between point $\mathbf{p}$ and $\mathbf{q}$ is defined as:

$$d = \sqrt{\sum_{n=1}^{k} (p_n - q_n)^2} \tag{1}$$

where p_n and q_n denote the n-th element of the descriptor vector of $\mathbf{p}$, resp. $\mathbf{q}$, for $n = 1$ to $k = 128$. However, in the SIFT matching algorithm, the distance is expressed in a dot product. As explained in (Wan and Carmichael, 2005), the distance can also be calculated by taking the arc cosine of the dot product of the descriptor vectors:

$$d = \cos^{-1}(\mathbf{p} \cdot \mathbf{q}) \tag{2}$$

For one keypoint, the distance to all keypoints in another image is calculated in order to find its nearest neighbour. For instance,

$$D = \begin{bmatrix} \cos^{-1}(\mathrm{des}_{11} \cdot \mathrm{des}_{21}^T) \\ \cos^{-1}(\mathrm{des}_{11} \cdot \mathrm{des}_{22}^T) \\ \vdots \\ \cos^{-1}(\mathrm{des}_{11} \cdot \mathrm{des}_{2j}^T) \end{bmatrix}, \tag{3}$$

is the distance vector from keypoint 1 in image 1 to all keypoints in image 2, with des_{ij} the descriptor vector of keypoint j in image i. The smallest value in this distance vector corresponds to a keypoint j that is the nearest neighbour. Not every feature from an image should however be matched with a feature in the other image because some features are not sufficiently alike or arise from background clutter. Therefore, these features are discarded by comparing the Euclidean distance of the closest neighbour to that of the second-closest neighbour. The distance of the closest neighbour divided by the distance of the second-closest neighbour is called the distance ratio. Matches above a certain maximum distance ratio should be rejected to eliminate false matches. A maximum distance ratio of 0.6 has been used.

An image at full resolution has almost 8 million pixels that all need to be examined for being a candidate keypoint. A computer with a 1.83 GHz processor did not succeed in finding keypoints due to a lack of memory. Matching all these keypoints would even ask for more memory. Therefore we are forced to reduce the image size from 3456×2304 pixels to 1500×1000 pixels. Table 1 shows the processing times of finding keypoints and the matching. However, we have more knowledge than is used in the algorithm, because we can estimate the location of a potential matching keypoint since we know the baseline and camera properties (and we assume that the images are captured in one straight line). So, instead of comparing all keypoints in one image with all the other keypoints in the next image, a keypoint is only compared with keypoints from a small area in the other image where the match should be found. This is illustrated in Figure 6. This matching strategy is around six times faster than the original matching as can be seen in Table 1.

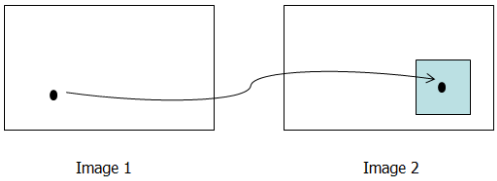

Figure 6: Search only in a small area for a match.

The SIFT matching algorithm also plots the matches between two images by drawing lines between the locations of the matching keypoints from one image to the other. A selection of matches, visualized by this program, is illustrated in Figure 7. The figure presents two images where the overlapping area (of image 1) is on the left hand side from the black dashed line. A match is represented by a cyan coloured line running from the image coordinates of a point to the image coordinates of its corresponding

Table 1: Computation times for finding and matching keypoints for the original algorithm and the optimized algorithm.

	keypoints	matching i+1	matching i+2	total
Original				
resized	13s	80s	77s	173s
full	n.a.	n.a.	n.a.	n.a.
Optimized				
resized	13s	12s	2s	27s
full	120s	> 20m	> 5m	> 27m

point. However, in the red rectangle matches are drawn starting from the right hand side of the black dashed line, so these are obvious incorrect matches. It is remarkable that all incorrect matches in this example are at or close to road markings. Apparently repetitive patterns are sensitive for wrong matching. Wrong matches have a negative influence on the accuracy of the bundle adjustment and should therefore be eliminated, which is done by RANSAC, (Fischler and Bolles, 1981).

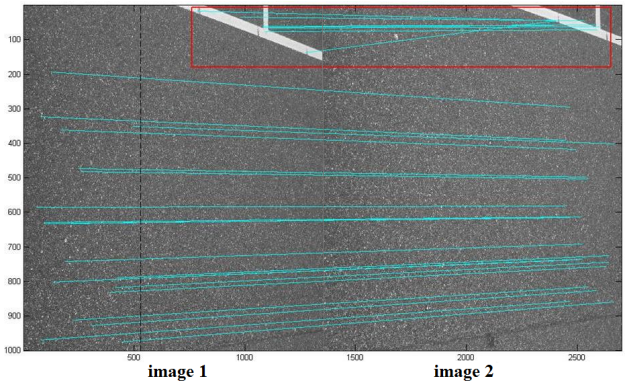

Figure 7: Failure of SIFT

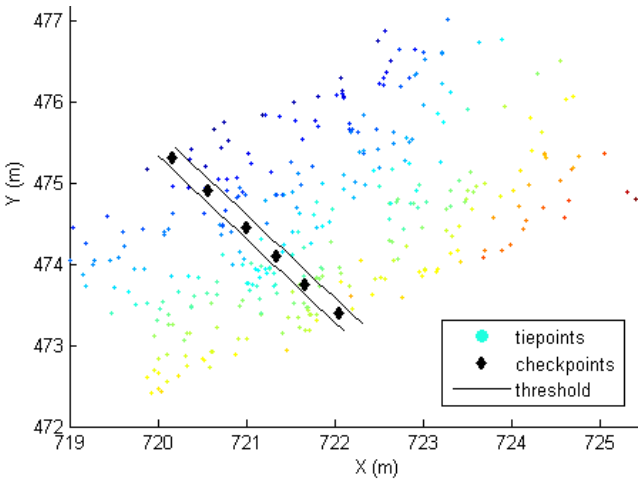

Figure 8: The selection of tie points within a small distance from the check points.

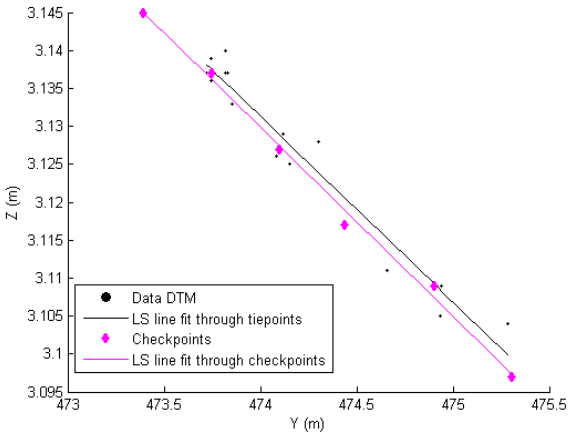

Figure 9: Validation of method

3 VALIDATION AND RESULTS

First, the method is validated with a set of independent check points. These check points are measurements on the road's surface and are not visible in the images. Hence, these points can not participate in the bundle adjustment, which makes them fully independent and therefore ideal check points. The validation uses six check points that are located approximately on a line perpendicular to the driving direction. The height of the check points is compared with the height of the obtained 3D coordinates of the tie points (object points). All object points within a small (horizontal) distance threshold from this line are selected as can be seen in Figure 8.

In Figure 9, the y-coordinates of the selected object points and the check points are plotted against their height. In this way, the slope of the road (of the part that was covered by the images) becomes visible. The least squares line fit through the check points shows very little difference with the least squares line fit through the object points. Absolute differences are 1-2 mm while the difference in slope is 0.022 degree. Now that the method is validated, the results can be judged properly.

It is expected that if more tie point observations are used in the bundle adjustment, the accuracy of the result of the bundle adjustment will improve. The effect of increasing the number of tie point observations to the quality of the bundle adjustment is evaluated next. A sequence of 28 images was processed with almost all available control points. The only variable is the number of tie point observations, obtained with variable distance ratios. There are two measures of quality, namely internal precision and external accuracy. The effect of varying the number of tie point observations to the internal precision is shown in Table 2 and the effect on the external accuracy is shown in Table 3. It is observed that the more tie point observations are used, the better the internal precision of the bundle adjustment. The external accuracy does however not improve if more tie point observations are used. The contrary seems even true. Best accuracies are observed when using the least number of tie point observations. This is not in line with the expectations. For the remainder of this research, only 28 tie point observations per image pair are used, since this yields best results and saves processing time.

Due to the long and narrow shape of the strip, the configuration of the control points will always be weak. This could also be seen in Figure 5. If more than 36 images or fewer than four control points were used, the bundle adjustment diverged (i.e. no solution could be obtained). A reduction of control points is therefore not possible. Those control points with the largest distance to the line, as visualized in Figure 5, are considered to contribute most to the quality of the solution. At the beginning of the strip, control point 2 in combination with control point 5 is considered

Table 3: Check point residuals (mm) for different number of tie point observations.

Check	**28 obs.**			**83 obs.**			**223 obs.**		
point #	**ΔX**	**ΔY**	**ΔZ**	**ΔX**	**ΔY**	**ΔZ**	**ΔX**	**ΔY**	**ΔZ**
3	3	2	-1	4	2	-1	4	2	0
4	2	1	1	2	1	1	2	0	3
34			0			4			1
35			1			1			2
36			-1			1			5
37			-2			1			8
38			1			-4			7
39			-1			-3			10
RMS	3	2	1	3	2	2	3	1	6

Table 2: The RMS precision of the bundle adjustment for camera position and orientation for different number of tie point observations per image pair.

Average number	(mm)			(°)		
of observations	**X**	**Y**	**Z**	φ	ω	κ
28	7	12	5	0.52	0.09	0.57
83	4	7	3	0.29	0.04	0.31
223	2	4	2	0.15	0.02	0.16

the best and control point 1 in combination with control point 3 is considered the weakest. Both combinations are executed in a bundle adjustment complemented with control points 7 and 9. The residuals on the check points are shown in Table 4. The table shows that a slight improvement of the configuration of the control points yields a significant better accuracy in height. Using the best configuration of control points makes it possible to process more images within one bundle adjustment.

Table 4: Check point residuals (mm) for for the weakest and strongest control point configuration.

Check	**Weakest configuration**			**Strongest configuration**		
point #	**Δ X**	**Δ Y**	**Δ Z**	**Δ X**	**Δ Y**	**Δ Z**
1				2	0	1
2	-5	0	-10			
3				3	2	-1
4	1	-2	14	2	1	0
5	-1	-4	16			
6	9	8	4	0	4	-5
8	14	10	39	18	12	32
34			8			6
35			3			7
36			-5			5
37			-12			3
38			-17			6
39			-26			3
RMS	8	6	17	8	6	11

4 DISCUSSION AND CONCLUSION

Regarding DTM construction, the angle of view of a standard lens is not sufficient to capture the entire width of the road, provided that the height of the camera is fixed at a height of 3.6 metres. A fisheye lens could be capable of covering the entire width of the road, but it is expected that its DTM will have a much lower accuracy since the ground resolution of the images will be much lower at the edges.

The main objective of this research was, however, to improve the accuracy of the laser profiles of Pavescan and to reduce the number of control points. A maximum sequence of 36 images could be handled successfully with BINGO, independent on the number of available control points or tie points. The bundle adjustment could not converge to a solution when using more than 36 images. It was shown that control points are of highest concern. The accuracy of the bundle adjustment strongly depends on the configuration of the control points. For a sequence of 28 images ($\approx$ 43 metres), about five control points are needed to achieve sub-centimetre accuracy of the object points. The practical feasibility for integrating close range photogrammetry into Pavescan is doubtful, since too many control points are needed that should spatially be well distributed and measured with tachymetry (or with similar accuracy). Even if more than 36 images could have been processed, the accuracy is not expected to get better unless more and better distributed control points are used. Cutting the bundle adjustment in small pieces may work. The configuration of the control points is much better in this way. Each strip would however still require at least five control points.

The original setup of Pavescan works with one control point every ten metres and the absolute accuracy of the laser scans depends mainly on the accuracy of the control points. It is concluded that close range photogrammetry does not improve the accuracy of Pavescan measurements in combination with decreasing the number of control points. An improvement of accuracy is only possible if at least the same amount of control points are used that are highly accurate and spatially well distributed. The practical feasibility differs therefore from the theoretical feasibility. Therefore it is not recommended to integrate photogrammetry in Pavescan.

The use of other software is not expected to give a better result, since BINGO is designed for (close range) photogrammetry and all settings were optimized. There are, however, other possibilities to overcome the problems with processing a strip of images. Although it is known that the scan provides a precise profile of the road's surface, such a measurement cannot be incorporated in photogrammetric software. A solution could therefore be to code a new bundle adjustment program with additional constraints that force the object points to follow the laser scan measurements.

ACKNOWLEDGEMENTS

The authors would like to thank Pieter van Dueren den Hollander and Edwin van Osch from Breijn B.V. for their supervision of the research. This research was also successful because Breijn allowed us to make modifications to the Pavescan system.

REFERENCES

Barber, D. M., Mills, J. P. and Smith-Voysey, S., 2008. Geometric validation of a ground-based mobile laser scanning system.

ISPRS Journal of Photogrammetry and Remote Sensing 63(1), pp. 128–141.

Fischler, M. and Bolles, R., 1981. Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM* 24(6), pp. 381–395.

Haala, N., Peter, M., Cefalu, A. and Kremer, J., 2008. Mobile lidar mapping for urban data capture. In: Proceedings 14th International Conference on Virtual Systems and Multimedia, pp. 95–100.

Jaakkola, A., Hyyppä, J., Hyyppä, H. and Kukko, A., 2008. Retrieval algorithms for road surface modelling using laser-based mobile mapping. *Sensors* 8, pp. 5238–5249.

Kertész, I., Lovas, T. and Barsi, A., 2008. Photogrammetric pavement detection system. *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* XXXVII(B5), pp. 897–902.

Kraus, K., 2007. *Photogrammetry, Geometry from Images and Laser Scans*. Second edn, De Gruyter, Berlin, New York.

Kruck, E., 1984. BINGO, Ein Bündelprogramm zur Simultanausgleichung für Ingenieuranwendungen, Möglichkeiten und praktische Ergebnisse. *International Archive for Photogrammetry and Remote Sensing*.

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

Mennink, J., 2008. Vastleggen dwarsprofiel in het verkeer. *Asfalt* 2, pp. 10–13.

Toth, C., 2009. R&D of Mobile Mapping and future trends. In: Proceedings of the ASPRS Annual Conference, Baltimore, United States.

Vilaca, J., Fonseca, J. C., Pinho, A. and Freitas, E., 2010. 3D surface profile equipment for the characterization of the pavement texture TexScan. *Mechatronics* 20, pp. 674–685.

Wan, V. and Carmichael, J., 2005. Polynomial Dynamic Time Warping Kernel Support Vector Machines for Dysarthric Speech Recognition with Sparse Training Data. In: Proceedings, Interspeech 2005, pp. 3321–3324.

Yu, S., Sukumar, S. R., Koschan, A. F., Page, D. L. and Abidi, M. A., 2007. 3D reconstruction of road surfaces using an integrated multi-sensory approach. *Optics and Lasers in Engineering* 45, pp. 808–818.