Fusion of point clouds from TLS and MVS for the generation of a 3D ship model

Daniel WUJANZ, Sven WEISBRICH, Frank NEITZEL,
Charalampos GAITANIS, Reza MOTMAEN ARANI, Svetozar PETROVIC

Abstract

One aim of this project was the determination of sensor positions by applying terrestrial laser scanning on a ship. Furthermore the outer shell of the vessel has been surveyed by using multi view stereo software namely Visual SfM and 123Dcatch. Finally the captured interior and exterior data had to be transformed into a common coordinate system where lucent targets have been applied.

1 Introduction

In 2011 the German Research Centre for Geosciences (GFZ Potsdam) purchased new equipment for gravimetric measurements usable on air- and shipborne platforms. The prevailingly planned application is the airborne gravimetry. Nevertheless, a first test of the new instrumentation has been performed on a vessel and focussed on the basic functionality of the system. The campaign took place at lake Müritz, Germany where the Klink, a liner with a capacity for 180 passengers, of the "Weisse Flotte - Müritz" served as a mobile platform.

The goal of airborne or shipborne gravimetry is to determine variations of the Earth gravity field along the aircraft's (or ship's) trajectory. Hence the gravimeter can be considered as the major sensor of the measurement system. In order to assign gravity measurements to points in space correctly, a determination of the trajectory is necessary, which has been carried out by Global Navigation Satellite System (GNSS) instrumentation consisting of four receivers and antennas. By carrying out GNSS measurements the possibility to calculate accelerations originating from the movement of the airplane or ship arise, which is a necessary prerequisite for gravimetric measurements on a mobile platform. Using this supplementary information together with the gravimeter recordings, which represent the total acceleration including both these accelerations and the accelerations of the Earth's gravity field, the latter can be determined. Finally, positions, velocities and accelerations determined by GNSS refer to respective GNSS antennas and have to be transformed to the actual origin of the gravimeter inside the instrument's housing. Therefore, the so called attitudes (three angles: roll, pitch and yaw) as well as relative positions of the components of the measurement system are needed. These three angles can be determined by using an Inertial Measurement Unit (IMU), which includes fibre-optic gyros, or by processing the recordings of at least three appropriately distributed GNSS antennas. The integration of all components of the measurement system into a fixed vehicle coordinate system by applying a terrestrial laser scanner (TLS) is the major ambition of this project. Topical developments in multi view stereo systems arouse interest if in addition the outer shell of the ship could be captured while the vessel was off dock by applying consumer cameras.

2 Related Work

Modern multi view stereo (MVS) solutions such as Bundler / PMVS2 (SNAVELY *et al.* 2008; FURUKAWA & PONCE 2007) are composed of many different segments fulfilling various tasks. An overview about the structure of such systems is given by SNAVELY *et al.* (2006) where the most important steps are:

- Detection of keypoints within the imagery e.g. LOWE's (2004) scale-invariant feature transform (SIFT),
- matching of the keypoints for instance described by SCHAFFALITZKY & ZISSERMAN (2002),
- determination of the camera parameters within a bundle adjustment (LOURAKIS & ARGYROS 2009) based on EXIF-tags for the initialisation of the focal length,
- computation of the camera locations via structure from motion (SfM) as e.g. proposed by BROWN & LOWE (2005).

Numerous practical applications of MVS have been undertaken in the field of computer vision (AGARWAL 2009) over course of the last years while lately the attention of the geodesy community has been caught. NEITZEL & KLONOWSKI (2011) apply an unmanned aerial vehicle (UAV) as a data acquisition platform in order to capture imagery and generate a 3D model of a landfill based on it. Furthermore a complete workflow, a comparison and accuracy analysis is undertaken. KERSTEN *et al.* (2012) conducted experiments on objects from architecture, cultural heritage and archaeology generated with different MVS packages. For sake of quality assessment a comparison with point clouds from TLS has been carried out. The civil engineers NASSAR *et al.* (2011) apply an available SfM solution for the volume detection of several types of excavated soils. As TLS have already been established as a surveying instrument no related contributions are outlined in this report.

3 Data Acquisition

The survey of the Klink was carried out in the homonymous town located about 150 km north-west of Berlin. A Leica C10 terrestrial laser scanner (TLS) and a digital Nikon D2X camera have been applied in order to capture the 25 m long, 5.10 m wide and 3.55 m high vessel. A tight time frame of 5 hours demanded to capture the imagery for the 3D model of the outer shell at first. The reason for that was the need to cast off in order to rotate the boat around its vertical axis so that all sides were visible, which has been carried out 3 times at a distance of roughly 30 m from the camera standpoints at the quay to the object. During the rotation of the boat 50 images have been captured from two standpoints about 10 m apart. After docking further 100 pictures have been taken from the quay facing starboard. A nearby roughly 8 m high tower has been used to capture additional 10 images from the upper deck. The Klink can be characterised as a challenging object in terms of its radiometric and geometric properties. Its white surface features low-contrast where keypoints are probably hard to detect while the symmetric shape of the boat can cause further problems during determination of homologous points. Additional problems are favoured by adverse lighting conditions and reflections of the sun from the water's surface.

For the survey via TLS of the upper and lower deck, which were equipped with the sensor systems mentioned above, the compensator of the instrument was switched off to avoid falsification by the rocking motion of the boat. Hence a fixed coordinate system has been established that is related to the vessel whereas no torsion of its hull had to be expected. Eight scans have been captured with a point spacing of 5 mm at 10 m object distance that have been connected with tilt- and turn targets that were attached by switching magnets onto the vessel and were distributed in such a way that a possibly large volume was described. A critical area was the staircase that connects the upper with lower deck which measured only 1.90 m by 1.10 m at the entrance and drops by 1.70 m at 45° slope. The survey was carried out as an open traverse starting on the upper deck and finally ending in the front part of the lower deck.

As the final aim is to connect both point clouds derived from TLS and MVS 24 targets have been applied distributed over all windows of the lower deck. The used Leica black and white targets have been printed on lucent foil while a sheet of paper has been placed between window and target, in case that the target has been captured by the laser scanner, or the paper has been attached over the target onto the window so that it ensured sufficient contrast for the pictures that have been taken from outside the ship. Figure 1 illustrates the moored vessel where lucent targets are ready for capturing images of the lower deck while the TLS is placed on the upper deck on its first standpoint.

Figure 1: Front part of the Klink with lucent targets prepared for image acquisition

4 Data Processing

Processing of the point clouds captured by the TLS has been carried out in Leica Cyclone 7.3 for registration of the single datasets as well as approximation of geometric primitives. Raindrop Geomagic Studio 12 was used for transforming the CAD-models into the registered point clouds. The generation of point clouds based on the captured imagery has been done via Visual SfM and Autodesk’s free yet “commercial” 123Dcatch software.

4.1 Terrestrial Laser Scanning

After registration of eight standpoints the misclosures added up to 1.9 mm on average based on 10 targets that led to 54 tie point connections with a maximum deviation of 5.5 mm related to the longest sighting of 30 m on the lower deck. The next step after registration was modelling the gravimeter, the INS unit as well as the mounting plate. The gravimeter was described by an approximating cylinder and a planar patch with a standard deviation of 2 mm respectively 0.9 mm while the mounting plate was again described by a planar segment leading to a standard deviation of 1.7 mm. Based on a given blueprint the cuboid-shaped IMU housing was modelled in AutoCAD and then transformed into the according point cloud by applying Geomagic where the average deviation summed up to 6.0 mm. For the two types of GNSS-antennas CAD-models have been derived based on their geometric description within the technical spec-sheets and registered into the point cloud leading to an average deviation of 7.3 mm. The next step which is mandatory for the combination of TLS and MVS was the automatic extraction of target centres that has been achieved through Cyclone. Figure 2 depicts the registered dataset where the sensor positions are highlighted by rectangular shapes: the smaller four indicate GNSS-antenna positions while the larger one describes the sensor platform including gravimeter and IMU on the lower deck.

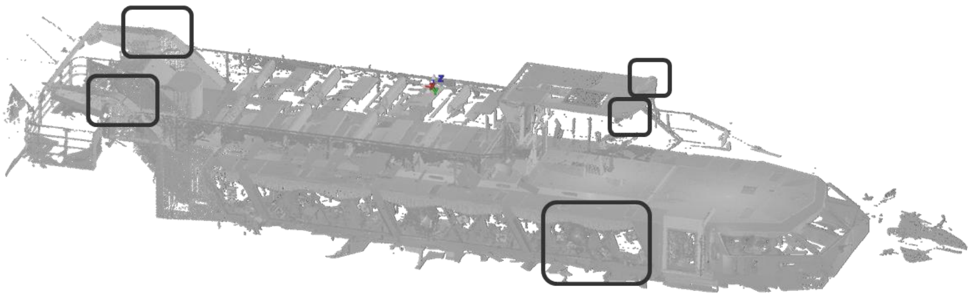

Figure 2: Registered TLS point cloud with highlighted sensor positions

A close look at the outcome of the produced data reveals various aspects that disregard established quality measures from surveying and adjustment calculation as mentioned by DOLD & BRENNER (2008) as well as WUJANZ (2012). It has to be mentioned that the computed misclosures only describe relations between adjacent datasets but do not allow drawing conclusions about the absolute accuracy of this open traverse. Another limiting factor is that the standard deviations computed from Cyclone refer to the distance between approximated surface and respective points but not how well the adjusted parameters satisfy the functional relationship of the geometric primitive. Another issue is caused by the transformation into a common coordinate system between CAD-models into point clouds or even between overlapping point clouds. This processing step is mostly carried out by solutions that are based on the iterative closest point algorithm (BESL & MCKAY 1992). The quality of such a registration is usually represented by an average deviation between the two datasets and again doesn't fulfil the necessary expressiveness that is needed to evaluate the quality and reliability of the transformation. Thus far there is only one approach (GRÜN & AKCA 2005) that is based on adjustment calculation and hence fulfils the previously mentioned demands of quality assurance.

4.2 Multiple View Stereo

The model of the outer shell of the vessel has been processed by different software packages namely Visual SfM and 123Dcatch. Based on several sets of images in order to clarify the following questions:

- Can MVS be applied for the acquisition of dynamic scenes?
- How robust are the solutions against outliers (e.g. caused by waves within the imagery)?

The first image bundle consists of roughly 50 images that have been captured during the rotation of the ship off dock, referred to as “rotation” in the following, while the second set features 100 images while the ship moored on the dock (subsequently called “dock”). In order to avoid the impact of outliers caused by the movement of the ship within the first set of images masked versions have been prepared where parts featuring areas of no interest have been covered in white. Masked image bundles are highlighted with the extension “masked”.

4.2.1 Visual SFM

Bundler is an academic SfM solution that is capable of handling unordered sets of images (SNAVELY 2006, 2008) and computes camera location in 3D space as well as a sparse point cloud based on the imagery. In order to increase the achieved point density from Bundler’s output FURUKAWA & PONCE’s (2007) patch-based multi view stereo software (PMVS2) can be applied. The parameters of the interior orientation that are computed by Bundler are the focal length and two coefficients that describe the radial distortion of the optical system. Furthermore a 3 x 3 matrix containing the camera’s rotation and a vector containing three translative components represent the relative orientation. Each computed point contains information about its location in space, its assigned colour as well as a list of views from where it has been captured (SNAVELY 2009).

Visual SfM is a collection of the solutions mentioned above which integrates additional functionality such as for instance a multicore bundle adjustment (WU *et al.* 2011) or computing SIFT keypoints on the graphics processing unit (GPU). A benefit of this solution is, apart from its graphical user interface (GUI), its high-performance even on large datasets (KERSTEN *et al.* 2012). In contrast to Bundler Visual SfM only models a single parameter for the radial distortion while the principal point is assumed to be located in the centre of an image. By introducing fixed calibrated camera parameters the principal point can be altered as well as the focal length which is otherwise initialised by applying the embedded EXIF-header.

The first test was conducted on the “rotation / masked” image bundle which resulted in a point cloud composed by 365 886 points. Interestingly the point cloud mostly features points on the starboard side while a closer analysis reveals that images from both sides have been matched to this side hence the result is not usable. A point cloud consisting of 864 449 points has been processed based on the original “dock” imagery, as illustrated in Figure 3, while the “masked” version did not produce satisfactory results. In summary no applicable dataset could have been produced with Visual SfM.

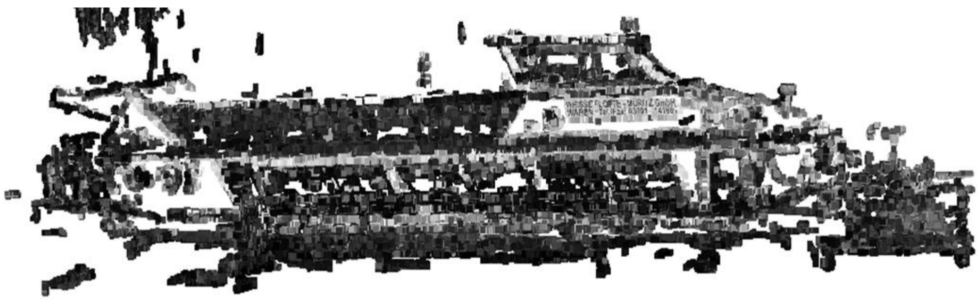

Figure 3: Outcome of original "dock" image bundle processed with Visual SfM

4.2.2 123Dcatch

123Dcatch is professional software by AutoCAD's developer Autodesk and has formerly been known as project Photofly in its previous version which is based on a programme called smart3Dcapture implemented by the French firm acute3D. The software uses a locally installed editor to load the imagery and furthermore serves as a communication platform to a server where textured meshes are computed via cloud computing. The capabilities of this editor are straightforward and do not allow much modification of the outcome. In case that the automatic matching process fails a manual stitching process can be started where unfortunately no more than four corresponding points can be determined which allows the connection to already connected images. No detailed information is given on which algorithms are applied yet it can be assumed that expertise from its chief technology officer Renaud KERIVEN and his PhD students (COURCHAY 2009, VU 2011) has been used.

The first tests by applying 123Dcatch have been conducted by processing the "masked / rotation" and "masked / dock" image bundles where no results have been achieved due to the fact that no images have been matched. Processing of the original "rotation" imagery led to 361 441 triangles. Due to the fact that the outcome is represented by a photo textured mesh the result appears subjectively to be better. However, a close visual analysis of the surface reveals several problematic areas. Two major parts of the ship showed erroneous effects: firstly on the front part of the boat and secondly in lower areas of the hull plating where reflections of the water occurred, see Figure 6. Another problem becomes visible at second glance and is caused by the triangulation process on example on the original "dock" set represented by 44 305 triangles. As depicted in Figure 4 objects that are close to the boat's hull are connected to it, which is a pole in this case. Similar problems occurred on the command bridge where points within the bridge have been triangulated to points from the outer shell. Unfortunately no possibilities to alter the computed mesh are implemented in order to correct these effects. The results computed with 123Dcatch have been used to generate a combination of MVS and TLS as the outcome appeared to be usable for this sake.

Figure 4: Outcome of original "dock" image bundle. Rectangular shapes highlight errors caused by the triangulation

4.3 Combination of the generated point clouds

The problem of connecting interior and exterior data has been solved by applying lucent print-outs of Leica's black and white targets as depicted in Figure 5. Two scenarios can be solved by this approach. On one hand scans from the interior of the boat can be carried out where a sheet of paper has to be placed between target and window while on the other hand images from the outside can be acquired by putting the piece of paper on top of the target. A number of 24 targets have been distributed over the six windows on each side of the Klink bearing a stable configuration in mind where the volume described by the targets should be as large as possible. The target centres have been automatically determined by applying Leica Cyclone respectively by hand in case of the MVS dataset derived by 123Dcatch.

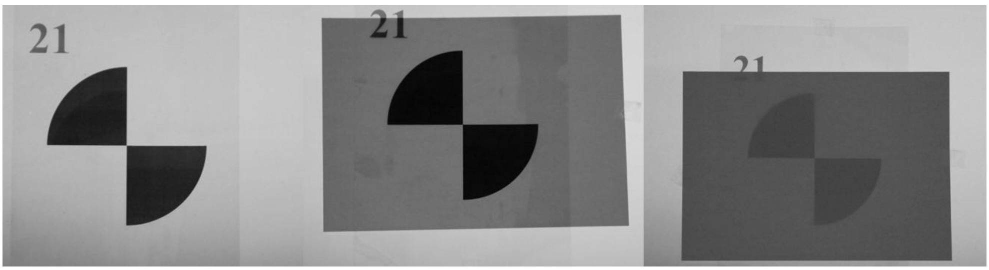

Figure 5: Lucent target (left), target prepared for interior data acquisition (middle) and for capturing imagery from outside (right)

In order to combine the datasets an adjustment has been conducted where corresponding points served as input data and three translations, namely t_x, t_y and t_z, three rotations (r_x, r_y, r_z) and a scaling factor m have been estimated. Therefore LOESLER's (2012) JAG3D software package has been used. The average misclosures summed up to 3.21cm in x-direction, 3.49 cm in y-direction and 0.16 cm in vertical direction. Table 1 displays the computed results where reasonable standard deviations for all parameters have been achieved.

Table 1: Outcome of the seven parameter adjustment

	t_x [m]	t_y [m]	t_z [m]	R_x [°]	R_y [°]	R_z [°]	m
Adjusted value	0.7543	0.8108	-3.8662	0.0114	0.3323	84.5199	0.5708
Standard dev. 1σ	0.0167	0.0165	0.0163	0.1891	0.2292	0.1948	0.0019

Figure 6 illustrates the outcome of the combination. The geometry of the windows close to the tie points can be rated as similar between TLS and MVS dataset showing only smaller deviations. However are the datasets starting to drift apart the further the area of interest is located from the tie points. Apart from the corrupt area, as accentuated by the rectangle in the figure below, the outcome can be seen as a promising state of the art which couldn't have been achieved with established geodetic techniques.

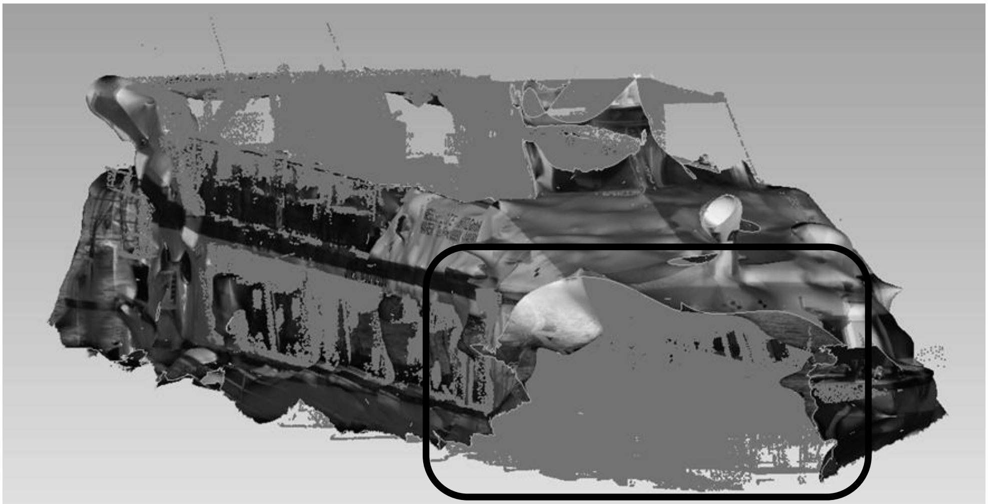

Figure 6: Combined model of both datasets derived from TLS and MVS

5 Conclusion and Outlook

This contribution describes first tests that have been conducted by applying TLS for static data acquisition on board of a ship while MVS came in to use for the survey of the outer shell. Furthermore a possibility to connect interior and exterior data is presented by applying lucent targets. Concerning the apparent quality of the MVS dataset improvements can be achieved by capturing more images, by shooting the imagery under better lighting conditions and by attaching pseudo random patterns onto the vessel's hull. The results of the combination of TLS and MVS data could have been improved by additional targets on the upper deck's windows. Unfortunately no sound conclusions concerning the achieved geometric quality can be drawn from a geodetic point view as no interpretable quality measures are implemented neither for the MVS or TLS software packages. Both MVS solutions had not implemented quality measures while Cyclone, the TLS processing software, displays only misclosures between standpoints but no standard deviations how well the functional relationship is satisfied. The same issue occurred for the matching process where given CAD models had to be transformed into the registered point cloud by using Geomagic. The reason for this generalised point of view can be explained by the origin of these algorithms from the field of computer vision or computer science in general. FOERSTNER (2002) tellingly compares the differences in scientific perspectives between computer vision and photogrammetry in his contribution. Hence approaches from the large field of computer vision or related fields shouldn't be seen as competitors but rather as an extension of the toolbox of surveyors. Nevertheless these algorithms have to be altered in such a way that demands from the field of engineering geodesy are satisfied yet one could say that they have to be "geodesised ".

Acknowledgements

The authors express their gratitude to the Weisse Flotte-Müritz and especially to the crew of the Klink for a pleasant co-operation during the survey of the vessel described in this study. We also appreciate the pleasant collaboration with Hartmut Pflug and Franz Barthelmes from the German Research Centre for Geosciences (GFZ). Furthermore we have to thank Przemyslaw Bielecki, Jacob Collstrup and Marcus Lachmann for their support during data acquisition.

References

Agarwal, S., Snavely, N., Simon, I., Seitz, S., Szeliski, R. (2009): *Building Rome in a Day*. International Conference on Computer Vision, Kyoto, Japan.

Brown, M. & Lowe, D (2005): *Unsupervised 3d object recognition and reconstruction in unordered datasets*. International Conference on 3D Digital Imaging and Modelling, pp. 56-63.

Besl, P. & McKay, N. (1992): *A method for registration of 3D shapes*. IEEE Trans. Pattern Anal. Mach. Intell. 14, pp. 239–256.

Courchay, J., Pons, J., Keriven, R., Monasse, P. (2009): *Dense and accurate spatio-temporal multi-view stereovision*. 9th Asian Conference on Computer Vision, Xi'an, China.

Dold, C., & Brenner, C. (2008): *Analysis of Score Functions for the Automatic Registration of Terrestrial Laser Scans*, International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. XXXVII, Beijing, China.

Foerstner, W. (2002): *Computer Vision and Photogrammetry – mutual questions: geometry, statistics and cognitions*. Bildteknik/Image Science, Swedish Society for Photogrammetry and Remote Sensing, pp.151-164.

Furukawa, Y. & Ponce, J. (2007): *Accurate, dense, and robust multiview stereopsis*. In Proc. International Conference on Computer Vision and Pattern Recognition , pp. 1-8.

Gruen, A., Akca, D. (2005): *Least squares 3D surface and curve matching*. In: ISPRS Journal of Photogrammetry and Remote Sensing, 59 (3), pp. 151-174.

Kersten, T., Lindstaedt, M., Mechelke, K., Zobel, K. (2012): *Automatische 3D-Objektrekonstruktion aus unstrukturierten digitalen Bilddaten für Anwendungen in Architektur, Denkmalpflege und Archäologie*. 32. Wissenschaftlich-technische Jahrestagung der DGPF, 14.-17. März, Potsdam, Germany.

Loesler, M. (2012): *Freie Netzausgleichungssoftware Java Graticule 3D*. http://derletztekick.com/software/netzausgleichung. (accessed 25.01.2012)

Lourakis, M. & Argyros, A. (2009): *Sba: A software package for generic sparse bundle adjustment*. ACM TOMS, 36(1), pp. 1–30.

Lowe, D. (2004): *Distinctive image features from scale-invariant keypoints*. International Journal of Computer Vision 60, 2, pp. 91-110.

Nassar, K., Aly, E., Jung, Y. (2011): Structure-from-motion for earthwork planning. Proceedings of the 28th ISARC, Seoul, Korea.

Neitzel, F. & Klonowski, J. (2011): *Mobile 3D Mapping With a Low-cost UAV System on Example of a Landfill Survey*. In: *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Proceedings of the UAV-g conference, Zurich, Switzerland.

Schaffalitzky, F. & Zisserman, A. (2002): *Multi-view matching for unordered image sets, or 'How do I organize my holiday snaps?'*. European Conference on Computer Vision, Copenhagen, Denmark, pp. 414–431.

Snavely, N., Seitz, S. M., & Szeliski, R. (2006). *Photo tourism: Exploring photo collections in 3d*. ACM Transactions on Graphics, 25(3), 137–154.

Snavely, N., Seitz S., Szeliski R. (2008): *Modeling the World from Internet Photo Collections*. International Journal of Computer Vision 80(2), pp. 189-210.

Snavely, N. (2009): *Bundler v0.4 User's Manual*. Retrieved from http://phototour.cs.washington.edu/bundler/bundler-v0.4-manual.html#S6 on the 21.01.2012.

Vu, H., Labatut, P., Keriven, R., Pons, J.-P. (2011): *High accuracy and visibility-consistent dense multi-view stereo*. IEEE Transactions on Pattern Analysis and Machine Intelligence.

Wu, C., Agarwal,S., Curless,B., Seitz, S. (2011): *Multicore Bundle Adjustment*, 24th IEEE Conference on Computer Vision and Pattern Recognition. Colorado Springs, United States of America.

Wujanz, D. (2012): *Towards transparent quality measures in surface based registration processes: Effects of deformation onto commercial and scientific implementations*. To be published in: Proceedings of the XXII Congress of the International Society of Photogrammetry and Remote Sensing, Melbourne, Australia.