FULLY AUTOMATIC FEATURE-BASED REGISTRATION OF MOBILE MAPPING AND AERIAL NADIR IMAGES FOR ENABLING THE ADJUSTMENT OF MOBILE PLATFORM LOCATIONS IN GNSS-DENIED URBAN ENVIRONMENTS

P. Jende^a, F. Nex^a, M. Gerke^b, G. Vosselman^a

^a Dept. of Earth Observation Science, Faculty ITC, Universiteit Twente, The Netherlands - (p.l.h.jende, f.nex, george.vosselman)@utwente.nl

^b Institute of Geodesy and Photogrammetry, Technische Universität Braunschweig, Germany m.gerke@tu-bs.de

KEY WORDS: Mobile Mapping, Feature Matching, Image Orientation, Adjustment

ABSTRACT:

Mobile Mapping (MM) has gained significant importance in the realm of high-resolution data acquisition techniques. MM is able to record georeferenced street-level data in a continuous (laser scanners) and/or discrete (cameras) fashion. MM's georeferencing relies on a conjunction of Global Navigation Satellite Systems (GNSS), Inertial Measurement Units (IMU) and optionally on odometry sensors. While this technique does not pose a problem for absolute positioning in open areas, its reliability and accuracy may be diminished in urban areas where high-rise buildings and other tall objects can obstruct the direct line-of-sight between the satellite and the receiver unit. Consequently, multipath measurements or complete signal outages impede the MM platform's localisation and may affect the accurate georeferencing of collected data. This paper presents a technique to recover correct orientation parameters for MM imaging platforms by utilising aerial images as an external georeferencing source. This is achieved by a fully automatic registration strategy which takes into account the overall differences between aerial and MM data, such as scale, illumination, perspective and content. Based on these correspondences, MM data can be verified and/or corrected by using an adjustment solution. The registration strategy is discussed and results in a success rate of about 95%.

1. INTRODUCTION

As a mobile, but terrestrial acquisition method, Mobile Mapping has become an important supplement to traditional geo-data acquisition techniques. A terrestrial perspective allows sensors to collect data from close vicinity and a different angle than aerial data. Resulting data products can be used in various fields, such as urban and infrastructure planning, real estate and transportation management. Hence, MM is not exclusively, but predominantly useful in urban areas. As MM requires a positioning solution based on satellite navigation to georeference its data postings, urban areas are challenging environments for accurate and continuous localisation. Especially so called urban canyons shaped by multi-storey buildings hamper reliable position fixes. The direct line-of-sight between the GNSS receiver and the satellite may be obstructed and the signal from the satellite might not reach the receiver, or the signal is reflected at façades or other objects and is received delayed in time. Although both scenarios have a different impact on the position estimation, the resulting accuracy and reliability are potentially decreased.

In this paper the absolute georeferencing component, i.e. GNSS, is provided by aerial images, whereas GNSS is still used to approximate MM positions. Aerial images are not affected by GNSS outages or multipath effects, and rely on high quality positioning equipment, accurate image positions after aerial triangulation as well as accurately calibrated camera systems. Hence, aerial images' orientation parameters can be employed to correct any overlapping data set as long as the registration with the target data set is accurate and reliable enough. This paper proposes a fully automatic pipeline to register MM and aerial nadir images in an efficient and accurate manner. The main contribution is the introduction of a two-step registration mechanism based on an approximated transformation between the MM and the aerial data set to enable a reliable feature

matching procedure robust against repeated patterns, illumination changes, and other differing image properties, such as original perspective or to some extent image content. In a future work, the adjustment procedure of MM images will be presented.

2. RELATED WORK

Many authors from different fields work on understanding GNSS error behaviour and develop solutions to mitigate these effects or restore the data's correct orientation information. GNSS multipath effects, for instance, can be discarded or filtered by using shadow matching approaches which utilise 3D building models to detect GNSS signals with unlikely incident angles (Gu et al. (2016); Strode and Groves (2016); Groves et al. (2013)). In the research field of autonomous driving, many methods rely on lane marking detection in conjunction with a digital map to support the localisation task (Gruyer et al. (2014), Schindler (2013), Roh et al. (2016)). A related method is visual odometry where features are tracked across multiple frames of the sensor system installed on the platform to allow for a better relative positioning (Badino et al. (2013), Zhang and Singh (2015)). These approaches are all designed to work in a real-time fashion to reduce GNSS-induced localisation issues in difficult scenarios, but cannot reach decimetre-grade accuracy, since external reference information is missing.

To retrieve highly accurate sensor orientations of MM platforms, many authors rely on surveyed Ground Control Points (GCPs) which can be integrated into an adjustment solution (Cavegn et al. (2016), Han and Lo (2016), Hofmann and Brenner (2016)). Using GCPs is, however, labour-intensive, difficult to automate, and therefore relatively expensive compared to the approach presented in this paper.

3. METHODOLOGY

3.1 Overview

The MM images which are used in this procedure have been recorded with an omnidirectional camera system. The resulting panoramic image is encoded in an equirectangular projection with a 360 * 180 degrees field of view. The aerial nadir images have been acquired with traditional photogrammetric techniques and have a classic central perspective projection. Obviously, scale and perspective differ dramatically between the aerial and MM data set rendering a direct registration difficult if not impossible.

In our previous work (Jende et al. (2016), Jende et al. (2016)), various techniques had been presented which increase the resemblance of the data sets. A crucial step is the projection of MM data onto an artificial ground plane to attain a pseudo-aerial perspective of a scene. In conjunction with resampling and blurring the MM image, the optical transfer function and the resolution of the aerial image can be simulated.

Moreover, MM orientation parameters may be inaccurate, but in the low-metre range at most, and can be thus used to constrain the search for correspondences. This assumption holds in particular for the registration of adjacent MM images, since their relative orientation is accurate.

These steps which have been discussed in our previous works are essential to simplify the registration process while making it more reliable. This contribution extends these methods towards necessary triangulation procedures, techniques to cope with repeated patterns, an adjustment of illumination differences, a multitude of quality measures, a full automatisation of the entire pipeline, and a thorough performance evaluation. In particular, enabling a reliable registration is our major contribution. Computing a transformation between the MM and the aerial data set for each image pair individually is based on a set of techniques to increase the resemblance of both data sets, mitigating wrong correspondences by e.g. repeated patterns, and by checking the plausibility of the result.

The final result is a set of 3D to 3D correspondences between both data sets which can be used to correct or to verify the MM images' orientation, respectively. In the following sections, the algorithm's details are discussed as well as conducted registration results are shown.

3.2 Ortho-projection of Mobile Mapping images

As mentioned earlier, a reprojection of the MM data increases the similarity to the aerial images and makes the registration procedure easier and less prone to mismatches. This step is, however, also useful as a prerequisite to match MM images with each other in order to retrieve the same feature points in at least two images, and consequently to obtain their coordinate in object space. Raw MM images are encoded in the equirectangular projection which maps a sphere onto a rectangular plane, thus every pixel in the image corresponds to an angular measurement. Hence, this projection entails significant distortions which makes a registration between MM images difficult and with aerial images infeasible.

There are different possibilities to reproject equirectangular images, be it a projective or homography transformation, exploiting the angular relation between the ground (defined by the height of the camera) and the panoramic image's geometry (Inverse Perspective Mapping), or by defining a ground plane in absolute coordinates around the platform's position and utilising a back-projection mechanic. Our approach is based on the latter idea, as the advantage is the intrinsic information of the defined plane in space which enables the restriction of search for correspondences in other images.

For the registration of MM images, images are projected with a ground resolution of 3 cm. For the registration with aerial nadir images, the resolution is set to 12 cm to match the aerial images' ground sampling distance.



Figure 1. Complete workflow for the registration of Mobile Mapping and aerial nadir images



Figure 2. Two adjacent Mobile Mapping images

3.3 Computation and verification of a transformation between adjacent MM images

In order to retrieve reliable correspondences between neighbouring MM images, a transformation is computed. To this end, the KAZE (Alcantarilla et al. 2012) detector and descriptor are used for feature matching. KAZE uses a similar detector and descriptor combination in common with SURF (Bay et al. 2008). Due to the superior scale-space computation and robust feature detection and description of KAZE, this algorithm has been selected for the registration task. Although both images share the same scale after the MM image has been ortho-projected and resampled to 3 cm, keypoints at different scaled instances of both images can still be retrieved to support the registration task. The default parametrisation had been changed to better fulfil MM ortho-images' requirements, thus a full scale (only two octaves are used) or rotational invariance is not needed, and the detection threshold for keypoints had been adjusted accordingly. Wrong correspondences are filtered, and the correct transformation is determined by RANSAC (Fischler and Bolles 1981).

Even though the images' relative orientation is accurate and could be used to derive a mapping function, there are two advantages to compute a transformation independently from MM orientation parameters. Firstly, a transformation needs to be found between ortho-projected MM images which may contain distortions due to an uneven ground surface. If a transformation is defined by image correspondences, distortions can be accounted for. Secondly, MM orientation information can be used to verify a computed transformation by image correspondences. The latter is realised by a projection of the image centre's image coordinate from the first into the second image by using the computed transformation. After a conversion to metric coordinates the distance to the second image's centre is computed, which is then compared to the relative orientation between the MM images using MM orientation information. If no transformation could be found or appears to be wrong, a transformation based on orientation parameters is computed instead.

3.4 Registration of MM images

Although MM images have been registered with KAZE features in the previous step, MM images are registered again by using corner correspondences and the computed transform. These corner correspondences are re-used for the registration with the aerial images, and KAZE correspondences - being blob features - are inapt for this task. Corner features can be identified at distinct locations in the images, such as road markings, kerbstones or manholes which are likely visible in the aerial images as well. Blob features, however, comprise image regions which are less distinct and more prone to individual image properties. This quality is useful to identify arbitrary correspondences between images, as it is used in this procedure to compute an initial transformation between two MM images, or between one MM and one aerial image later. If the emphasis is placed on recognisable, trackable and exactly localisable features in multiple MM and aerial images, corner features are more

useful. Corners are detected in one of the MM images only at subpixel accuracy by using the Förstner operator (Förstner and Gülch 1987). Using the previously computed transformation, these corners are projected into the other image. A template is extracted from the first image and since the transformation, be it based on KAZE features or orientation parameters, is accurate, a corresponding search window in the second image can be kept small. The registration relies on normalised cross correlation and least squares matching. The latter allows the corners to be defined at subpixel accuracy as well. However, since the resolution of the MM ortho-images is 3 cm and the registration with the aerial images is performed at 12 cm resolution, even a computationally cheaper normalised cross correlation will lead to a high corner accuracy.

3.5 Triangulation of 3D points from MM images

Since the corner features have been identified on the artificial ground plane defined earlier, their approximate 3D coordinate is already known. Due to an uneven ground surface, however, distortions in the projection may lead to an inaccurate object coordinate. Hence, all corner features are converted into spherical coordinates using the inverse of the ortho-projection to coincide with the equirectangular projection.

As the location of all MM cameras and their respective yaw deviation¹ is known and correspondences have been identified, a triangulation is performed to obtain object coordinates for all corner features.

3.6 Resampling and blurring the MM patch

Prior to a registration between the MM and the aerial data set, a couple of pre-processing steps have to be performed. As mentioned earlier, the MM data has been projected onto the ground with a resolution of 3 cm. To enable a registration with the aerial image at the same scale, the MM data is resampled to 12 cm accordingly (holds for the aerial images used in the experiments). Moreover, the MM images are rotated to be aligned with the aerial images in order to avoid accounting for rotational invariance.

Since the distance between the object and the camera and the optical instruments differ significantly between both recording systems, their optical transfer functions deviate. To simulate this difference, the MM ortho-image is blurred with a Gaussian filter.

3.7 Initial transformation between MM and aerial image

Prior to the actual registration of both data sets, a transformation between a MM patch and all overlapping aerial image patches is sought. To retrieve all aerial images which overlap with the respective MM image, its recording location is projected into the entire aerial block. An initial transformation, similar to the one depicted in 3.3, is computed to support the registration of corner features. As mentioned earlier, corner features are more suitable to be identified precisely, and tracked across multiple images. KAZE features are more applicable to register an individual image data set due to the scalability and robustness of the approach.

However, another equalisation procedure – the Wallis filter – is applied to both images (Wallis 1974). Especially contrast and illumination differences between the aerial and MM image may hamper a successful registration. Experiments could show that Wallis filtering equalises the images and improves the quality of the registration. As the orientation parameters of the MM image may be inaccurate and are subject to be adjusted, the computed transformation between the aerial and MM image cannot be verified unlike earlier in 3.3. Thus, only the plausibility of the transformation can be checked based on thresholding elements of the computed transformation matrix, i.e. translation component. If a transformation cannot pass this check with at least two overlapping aerial images, the MM image is discarded.





3.8 Registration of MM and aerial image

The registration of both data sets relies on corner features which have been detected in a previous step (see Figure 3 for scheme). As an individual transformation is known between the MM image and every aerial image patch, these corner features can be projected into the corresponding aerial image. For each corner feature, a template and a search window are defined; in the MM and the aerial image, respectively.

Again, the registration is based on normalised cross correlation and least squares matching. This step entails that there are individual correspondences between one MM image and an arbitrary number of aerial images. Once the same point is matched between a MM and multiple aerial images, the aerial images do not have be matched with each other again to derive correspondences, as they are already registered. Due to this parallel matching process and slightly differing image properties across the aerial images, minor matching offsets of the aerial image coordinates for the same point may occur. To retrieve the exact same point across all aerial images, a check is performed, and if necessary the point in the aerial image is adjusted.

A subsequent outlier removal based on RANSAC is performed. An important requirement for this step is a minimum number of ten correspondences in order to have a stable estimation.

3.9 Triangulation of aerial images' correspondences

The triangulation procedure accounts for an arbitrary number of aerial images as well by using a least squares approach. Within the triangulation framework there are multiple checks for blunders. Because the workflow processes MM images sequentially, but at least two at a time with the same corner features, the matching procedure for each MM image is individual. Therefore, the same object coordinate derived from aerial images may be returned more than once. This is used for outlier removal as well. The result is an observation in object

¹ Roll and pitch deviation are pre-corrected and thus 0 degrees.

space in at least two MM images and at least two aerial images. These correspondences can be used within an adjustment solution to correct the MM platform's trajectory.

4. RESULTS AND DISCUSSION

As mentioned earlier, Mobile Mapping is primarily useful in urban areas. Road markings and other distinct features on the ground are usually abundant and well visible from both platforms. In these areas, our method performs well and returns reliable results for data adjustment. Narrow alleys or areas with no, sparse, or small road markings, however, are rather difficult to process. First of all, due to sensor differences and a different acquisition date, alleyways are prone to dramatic illumination changes. Especially aerial images may have a low contrast in these areas which reduces the salience of potential features. If only a few road markings are available a reliable transformation between the MM and aerial image is difficult to find. Since the relative orientation of the MM platform is accurate, though, not every MM image in the trajectory has to be registered with the reference data set. Therefore, an adjustment solution has to be designed to account for this case. This section comprises the registration result of three different MM trajectories in the city of Rotterdam, NL. The selected test areas have different characteristics which are described in the three subsections.

Test	No	Images	Number of	Correct
area	images	matched	correspondences	matches ²
1	45	7	129	93,5 %
2	107	25	1079	96,8 %
3	136	14	1381	94,8 %

Table 1 Registration results for three urban test areas in Rotterdam, NL

4.1 Test area 1

The first and smallest test site with 45 MM images in total includes a major junction on the one hand and a road without any road markings on the other hand (Figure 4). This example shows a limitation of our approach. The MM images, however, which could be matched with the aerial reference returned a high number of inliers (see Table 1).



Figure 4. Test area 1; green: recording locations with correspondences to the aerial images; blue: recording locations without correspondences

A major challenge in this area are the different shades of the road surface in conjunction with illumination differences. As an example, see Figure 5. All correspondences are correct, but may be slightly displaced due to aforementioned image differences. This makes a registration rather difficult, as the image properties differ locally and hamper the exact localisation of the correlation's peak. Mismatches were caused by a similar behaviour. This area is quite a negative example also regarding a potential adjustment of the data set. Correspondences with the aerial reference can just be found in the areas of junctions. A solution to this problem is to extend the test area to include other road markings on the other side of the test area in order to allow for a more stable adjustment obviating the risk of error propagation.





4.2 Test area 2

The second test area (see Figure 6) consists of a multi-lane road with a couple of crossing. In total, there are 107 MM images, 25 of them could be registered with over 1000 correspondences to the aerial data set. Interestingly, most of the images which were matched are at crossings where there is an abundance of road markings. In the actual registration procedure, however, more images with correspondences could be identified. Due to the threshold of at least ten correspondences per image pair, especially images comprising only lane markings and thus fewer corner points, were discarded. In particular, changing the threshold to accept fewer valid correspondences would have likely led to more listed recording locations (green dots in Figure 6) in test area 2. In other words, the procedure can be adapted in order to process MM images with fewer correspondences which may, however, entail the risk of a less stable transformation and thus an inferior result.



Figure 6. Test area 2; green: recording locations with correspondences to the aerial images; blue: recording locations without correspondences

² Average of inlier percentage per image pair



Figure 7. Test area 2; Although the images differ due to some occlusions, our approach could yield only correct correspondences

A challenging task in the registration between aerial and MM images is the inherent original perspective. Even though the MM image can be projected to better adapt it to the aerial image, elements not visible in the other image can hinder a successful registration. This problem is well visible in Figure 7 where a car is folded to the side in MM image, or branches of a tree slightly occlude the view in the aerial image. Although a procedure based on some image understanding, i.e. connecting road markings to create graphs for matching or identifying unique anchor points in both images seem intuitive, our approach based on an initial transformation and a subsequent corner matching is more robust in that regard to account for unforeseen differences between the images.



Figure 8. Test area 2; Registration may be hampered due to time difference

A related issue is a potential time difference between the data set which causes differences in image content. As depicted in Figure 8, the zebra crossing in the MM is worn off, and thus some image correspondences are incorrect. In order to use aerial images for the correction of MM data, it is crucial that the acquisition date does not differ too much. This holds in particular for data acquired in or over cities, respectively, as these areas may change relatively fast.

The last example for test area 2 (see Figure 9) represents a typical road marking where the illumination and contrast properties between both images differ to a great extent. In order to obtain correct correspondences under such conditions, the procedure relies on feature detection in only one image. A separate feature detection in the aerial image would have potentially led to a different keypoint response and thus wrong correspondences. Moreover, a Wallis-filter which equalises the images is used in a prior step to support determining the transformation between the MM and aerial image. Once the transformation is known, corner points can be projected into the other image, and the search for a correspondence can be limited to a small search window. Since this test area has well distributed and a high number of

correspondences across its trajectories, an adjustment can be reliably performed.



Figure 9. Test area 2; Successful registration also in areas of contrast and illumination differences

4.3 Test area 3

Test area 3 comprises the greatest number of recording locations and correspondences with the aerial image data set (see Table 1). However, only 14 out of 136 MM images could be registered, and their distribution is unequal (see Figure 10). Although, the potential number of salient features in this area is comparably high in total, the defined criterion of a minimum number of ten correspondences per image pair had a strong impact on several cases. Moreover, only two to three aerial images overlap this area while the southern trajectory is mostly shaded in all aerial images making a registration difficult. Hence, the combination of circumstances led to a mediocre result regarding the distribution of correspondences, whereas the total number of correspondences per registered image is relatively high.



Figure 10. Test area 3; green: recording locations with correspondences to the aerial images; blue: recording locations without correspondences

Similarly, once a valid transformation could be identified and enough salient features are present, a great number of correspondences can be identified (see Figure 11). A challenge for feature-based registration techniques relying on template or descriptor approaches is repeated patterns which introduce ambiguities into the matching procedure. Our approach is able to cope with these patterns in most of the cases, as a transformation is computed prior to the actual registration procedure. Within this step, the transformation is estimated in multiple iterations and checked for its plausibility.



Figure 11. Test area 3; 69 correct correspondences between MM and aerial image

5. CONCLUSION

This paper presented a fully automatic workflow to register MM and aerial images. Techniques to overcome their overall differences had been presented and results had been discussed. It could be shown that the procedure works in a reliable fashion with a success rate of about 95%. As a result, correspondences between both data sets enable an adjustment of the MM data. To this end, future developments will concentrate to extend the pipeline towards this goal. However, a registration cannot be performed under certain conditions, e.g. low textured areas, not enough salient features, great differences in image content, and occlusions. Thus, oblique aerial images will be introduced as a supplementary source of orientation information.

ACKNOWLEDGEMENTS

This work is part of the TTW-Open Technology Programme with project number 13589, which is (partly) financed by the Netherlands Organisation for Scientific Research (NWO).

REFERENCES

Alcantarilla, P. F., A. Bartoli and A. J. Davison, 2012. KAZE features. Computer Vision–ECCV 2012: pp. 214-227

Badino, H., A. Yamamoto and T. Kanade, 2013. Visual odometry by multi-frame feature integration. *Proceedings of the IEEE International Conference on Computer Vision Workshops*

Bay, H., A. Ess, T. Tuytelaars and L. Van Gool, 2008. Speededup robust features (SURF). *Computer vision and image understanding* 110(3), pp. 346-359.

Cavegn, S., S. Nebiker and N. Haala, 2016. A systematic comparison of direct and image-based georeferencing in challenging urban areas. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1, pp. 529-536.

Fischler, M. A. and R. C. Bolles, 1981. Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. *Commun. ACM* 24(6): 381-395.

Förstner, W. and E. Gülch, 1987. A fast operator for detection and precise location of distinct points, corners and circular features. *Proc. ISPRS intercommission conference on fast processing of photogrammetric data*. 1987. Groves, P. D., Z. Jiang, M. Rudi and P. Strode, 2013. A portfolio approach to NLOS and multipath mitigation in dense urban areas. *The Institute of Navigation*

Gruyer, D., R. Belaroussi and M. Revilloud, 2014. Map-aided localization with lateral perception. *Intelligent Vehicles Symposium Proceedings*, 2014 IEEE

Gu, Y. L., L. T. Hsu and S. Kamijo, 2016. GNSS/Onboard Inertial Sensor Integration With the Aid of 3-D Building Map for Lane-Level Vehicle Self-Localization in Urban Canyon. *IEEE Transactions on Vehicular Technology* 65(6), pp. 4274-4287.

Han, J. and C. Lo 2016. Adaptive Time-Variant Adjustment for the Positioning Errors of a Mobile Mapping Platform in GNSS-Hostile Areas. *Survey Review*: 1-6.

Hofmann, S. and C. Brenner, 2016. Accuracy assessment of mobile mapping point clouds using the existing environment as terrestrial reference. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1, pp. 601-608.

Jende, P., Z. Hussnain, M. Peter, S. Oude Elberink, M. Gerke and G. Vosselman, 2016. Low-Level Tie Feature Extraction of Mobile Mapping Data (MLS/Images) and Aerial Imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XL-3/W4, pp. 19-26.

Jende, P., M. Peter, M. Gerke and G. Vosselman, 2016. "Advanced Tie Feature Matching For The Registration Of Mobile Mapping Imaging Data And Aerial Imagery." *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B1, pp. 617-623

Roh, H., J. Jeong, Y. Cho and A. Kim, 2016. Accurate Mobile Urban Mapping via Digital Map-Based SLAM. *Sensors* 16(8): p. 1315.

Schindler, A., 2013. Vehicle self-localization with high-precision digital maps. 2013 IEEE Intelligent Vehicles Symposium (IV).

Strode, P. R. R. and P. D. Groves, 2016. GNSS multipath detection using three-frequency signal-to-noise measurements. *GPS Solutions* 20(3), pp. 399-412.

Wallis, K. F., 1974. Seasonal Adjustment and Relations Between Variables. *Journal of the American Statistical Association* 69(345), pp. 18-31.

Zhang, J. and S. Singh, 2015. Visual-lidar odometry and mapping: low-drift, robust, and fast. 2015 IEEE International Conference on Robotics and Automation (ICRA).