IMPROVED 3D SPARSE MAPS FOR HIGH-PERFORMANCE SFM WITH LOW-COST OMNIDIRECTIONAL ROBOTS

Pedro Cavestany[⋆†], *Antonio L. Rodríguez*[⋄], *Humberto Martínez-Barberá*[†], *Toby P. Breckon*[‡]

University of {⋆Cranfield, UK †Murcia, Spain ‡Durham, UK}, ⋄Inst. Transuranium Elem. Joint R. C., Italy

ABSTRACT

We consider the use of low-budget omnidirectional platforms for 3D mapping and self-localisation. These robots specifically permit rotational motion in the plane around a central axis, with negligible displacement. In addition, low resolution and compressed imagery, typical of the platform used, results in high level of image noise ($\sigma \backsim 10$). We observe highly sparse image feature matches over narrow inter-image baselines. This particular configuration poses a challenge for epipolar geometry extraction and accurate 3D point triangulation, upon which a standard structure from motion formulation is based. We propose a novel technique for both feature filtering and tracking that solves these problems, via a novel approach to the management of feature bundles. Noisy matches are efficiently trimmed, and the scarcity of the remaining image features is adequately overcome, generating densely populated maps of highly accurate and robust 3D image features. The effectiveness of the approach is demonstrated under a variety of scenarios in experiments conducted with low-budget commercial robots.

Index Terms— structure from motion, mobile robot, omnidirectional, noise, feature filtering

1. INTRODUCTION

Research on mobile navigation in complex environments has been significantly boosted with the manoeuvrability of holonomic robots [1, 2]. A holonomic robot platform has as many actuators as degrees of freedom. In the case of a wheeled robot, which has three degrees of freedom (two normal directions and rotation angle), a robot needs three actuators to be holonomic. Specifically in our work this configuration is achieved by three independently commanded wheels, which are able to move almost friction-less along the perpendicular direction to their axis of displacement. This paradigm is represented by the omni wheels (Fig. 1a). The manoeuvrability offered by this design allows an omnidirectional robot to turn on the spot and move sideways or diagonally while keeping its orientation (Fig. 1b). Such omnidirectional platform often offers key manoeuvrability characteristics with a wide range of application domains [3, 4, 5].

Sequential Structure from Motion (SfM) techniques have been applied to obtaining robust 3D mapping and self-localisation on mobile robots [6, 7, 8]. When these methods are applied to low-cost computing platforms, the scene map usually consists on a set of sparse 3D scene features. Moreover, rapid changes in the camera viewpoint due to the abrupt rotations that characterize omnidirectional platforms (Fig. 1), the low camera resolution and the image compression required for wireless video streaming reduce the quantity and accuracy of 3D features that can be retrieved from the scene using SfM techniques. The specific characteristics described also tend to complicate the extraction of the epipolar geometry between image pairs [9, 10, 11, 12], rendering the navigation task challenging.
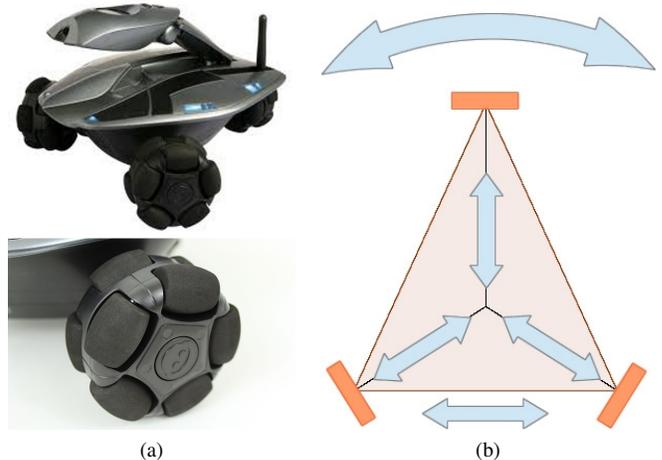
Fig. 1: a) omnidirectional Rovio robot and omni-wheel detail; b) possible motion directions for the omnidirectional robot platform.

This paper describes a sequential SfM system that addresses these problems, using a specific point filtering strategy, and a novel noise resilient feature tracking method based on the relationship created between a 3D point and its *bundle* of image features detected at the input views. By increasing the number and accuracy of the 3D features in the map, the reconstruction process becomes more robust.

We efficiently manage the *bundle*s and propose suitable filters that optimise the addition of new features to a *bundle*, as well as the merging between *bundles* via the matching of their features. Furthermore, we develop specific methods to robustly cope with noise levels which are typical of such holonomic platforms.

We evaluate these techniques, and obtain highly robust and accurate reconstruction results on a low-budget omnidirectional robot over a range of different environments, under noisy and sparse feature matching conditions and in presence of frequent narrow baseline configurations which are typical in this platform genre. These results outperform comparable techniques in the field for this configuration.

2. PRIOR WORK

There is a range of prior work considering monocular Structure from Motion (SfM) on mobile robot platforms [13, 7, 8]. The motion of a robot equipped with a single camera and moving on a planar environment was estimated with a SfM approach in [6] and [7]. Mouragnon *et al.* [14] performed 3D reconstruction on a mobile robot via an embedded system based matching process and a local bundle adjustment technique, albeit within a known environment. By contrast, our work is not specifically constrained to planar motions and is capable of reconstructing unknown environments on an omnidirectional platform in the presence of the ill-conditioned image baselines outlined.

The last decade has seen increasing research on omnidirectional

robots due to reduced production costs and design improvements [4, 5, 15]. However, the majority is devoted to motion modelling [15], and little research has been done on visual navigation [16]. To the best of our knowledge, no SfM system has previously addressed the specific problems of this kind of robotic platform.

Using central panoramic cameras [9] achieves the reconstruction with small baselines. The authors in [12] propose a robust algorithm for extracting the epipolar geometry assuming planar motion with small baselines. The present work extends these approaches to cope with ill-configured conditions which are ever present in the motion characteristics of an omnidirectional platform, by using non-linear methods and an appropriate feature matching selection policy.

Central to the SfM problem is the concept of feature matching and tracking [17, 13, 18]. Generally, prior work concentrates on variation and feature type, rather than the cross-image tracking methodology. In order to make the most of sparse matches provided by a low resolution camera, we have developed a novel feature tracking system which handles the inter-bundle relationships via robust and light filters.

The feature tracking system developed in this work also deals with the other key issues that we have encountered - *noise*. The analysis of Hartley *et al.* [19] shows the huge influence of noisy correspondences in the 3D point triangulation, where the authors estimate a noise level of $\sigma = 0.2$ in their real world images. Hebert [13] deals with uncertainty in SfM with noise variance up to 1 pixel, being this variation the overall trend in the field [20]. By contrast, our images present a noise level $\sigma = 10.58$ (estimated), due to the JPEG-compression artefacts inherent in wireless communication implemented on such low-cost omnidirectional platforms. Our novel feature tracking system, along with other noise filters implemented throughout the pipeline, successfully discards outliers within the overall feature matching process.

Furthermore, to date no work in the field has addressed the specific issues of ill-conditioned short baseline configurations within the context of noisy, low quality imagery found on holomonic platforms. Here we extend the state of the art with a noise-tolerant pipeline that overcomes such issues, bringing SfM to such platforms for navigation, 3D mapping and self-localisation tasks.

3. HOLONOMIC STRUCTURE FROM MOTION

In sequential Structure From Motion (SfM) we consider, at any given time, the most recent image I_n received by wireless transmission from a low-quality onboard camera as it transits the scene. This image I_n passes through a processing pipeline that recovers the global robot pose and the scene map (structure). First the image is filtered, and features are detected and matched against features detected in previous frames. Subsequently a feature tracking method is used, before the final recovery of the actual camera pose and the map update.

3.1. Feature Detection and Matching

Firstly, bilateral filtering is applied as an efficient inexpensive method to perform feature preserving noise reduction on each image received [21]. Subsequent feature extraction is performed using SURF [22] as an efficient trade-off between computational efficiency and robustness. Here 64 dimensional SURF features are extracted from image I_n. We then use k-d tree based lookup [23] on the feature descriptors to perform pairwise image matching between I_n and previous images I_{n-i}, with i increasing until the match population found in the pair $\{I_{n-i}, I_n\}$, $i = k$, is below a given

threshold τ_m (empirically, $\tau_m = 20$). We denote this recursive matching by the expression $\{I_{n-i}, I_n\}_{i=1}^{i=k}$, $1 \leq k \leq n$, and the set of feature matches created for each pair $\{I_{n-i}, I_n\}$, by S_{in}.

Our contribution at this stage is the careful selection of feature matches by quality. We assess the quality of a match between two features a and b by the *L2* difference of their descriptors, denoted by δ_{ab}. Three match quality filters are deployed. Firstly, only unique matches are considered. The uniqueness of a match is defined by the ratio δ_{ab}/δ_{ac}, where b is the closest matching feature to a, and c is the second closest matching feature [24], based on *L2* difference of the SURF descriptors. Ratios lower than a threshold τ_u do not generate a match (we set $\tau_u = 0.4$). Secondly, only the best matches of S_{in} are selected. This selection is accomplished by taking certain percentile rank, τ_κ of the score on δ_{ab} population over S_{in} (we use $\tau_\kappa = 0.8$). Finally, we enforce one-to-one feature matching between image pairs. This combination of filters counteracts the effect of noise on the feature matching process but additionally results in a significantly sparse set of feature matches S'_{in} from which we then have to perform SfM.

3.2. Relative Pose Estimation

Based on the identified set of filtered matches S'_{in}, RANdom SAmple Consensus (RANSAC) [25] is performed to find an inlier subset of matches S''_{in}, using the epipolar equation $x'^T E x = 0$ as parametrising model. In the case of S''_{1n}, where the relative pose is required, we subsequently recover the essential matrix E with the algebraic error minimisation approach described in [26]. The extraction of E leads to the estimation of the relative camera pose of I_n. Subsequently S_{in} is examined and added to the structure population.

3.3. Feature Tracking

A key problem implicit in all SfM approaches is the feature registration problem, where multiple pair-wise feature correspondences must be merged into a single multiple-view feature track, or *bundle* of features for a given 3D point X.

Three main computational operations should be enabled when efficiently feature tracking matches over a sequence:- 1) direct access to X referenced from any feature in its *bundle* and vice versa, 2) addition of new features to a *bundle* and 3) merging of two *bundles*. In our tracking method, novelly we devise *bundles* as dynamic lists, a structure which allows us to efficiently perform these tasks. Furthermore, when a new feature is added to the *bundle* of X, our specific implementation of *bundle* will automatically link it to X and to the rest of features of the *bundle*.

Given the sparsity of the 3D point cloud produced by our matching filters (See Section 3.1), it is necessary to properly manage the addition of features to a *bundle* and the merging between *bundles*, in order to create sufficient duration feature tracks. This is handled by two filter checks. The first filter f_1 checks, when a feature m_a from image I_a is matched with a feature m_b from image I_b, whether the *bundle* associated to m_b has already a feature from image I_a. The analogue check is done with the bundle associated to m_a. When this is the case it compares the values of the coordinates of the features involved to establish whether they are truly the same feature. This ensures that a *bundle* is linked to one feature per image. The second filter f_2 compares whether two 3D points p_i and p_j are close enough to be considered the same 3D point. For each axis $\imath \in \{x, y, z\}$ we define $\delta^\imath = \|p_i^\imath - p_j^\imath\|$, and $\mu^\imath = mean\left\{p_i^\imath, p_j^\imath\right\}$. The filter f_2 checks that $\delta^\imath < k \cdot \mu^\imath$. In this case, they are assumed to be the same point. Empirically we use $k = 0.02$.

For every feature match of S''_{in} three possible cases arise:- 1) none of the features belong to any *bundle*, 2) one feature of the match belongs to a *bundle* and 3) both features belong already to different *bundles*. In the first case, a new 3D point $\{0, 0, 0\}$ and its *bundle* is initialised. The actual value of the corresponding 3D point will be estimated in the triangulation step (Section 3.4). At this point the *bundle* is composed of the two matching features. In the second case the filter f_1 is conducted. In case of success the *bundle*-less feature is added to the *bundle* of the other feature. Otherwise, the new feature is discarded. In the third case, additionally, the filter f_2 is applied. If the pair of *bundles* passes this last filter, they are merged.

The specific creation and management of the structure of *bundles*, along with the filters associated to it, allows us to obtain precise camera poses and a reliable point cloud out of sparse matches populations (in our experiments, at this stage an average image has 755 views, with 3.56 projections per 3D point, see Fig. 2).

3.4. Joint Pose and Structure estimation

The introduction of the sets $\left\{S''_{in}\right\}_{i=1}^{i=k}$ increases the structure population and widens the range of the *bundles*. With this new information the scale of the camera pose of I_n is adjusted to be coherent with the rest of the sequence. This refinement is performed via the resection method proposed in [27].

Once the global camera poses have been calculated the triangulation process over the updated point cloud takes place, where the new 3D points are estimated and those whose *bundles* have increased are recomputed. Subsequently, the structure undergoes filtering based on reprojection error and cheirality [26] (i.e. those 3D points behind the camera are deleted).

The last stage of the reconstruction involves the application of Bundle Adjustment (BA), where camera poses and 3D points are simultaneously optimised by minimising the reprojection error function cost. This work runs the implementation of [28] which efficiently applies Levenberg-Marquardt minimization method by exploiting the sparseness of the SfM problem. We employ BA in two scopes, locally and globally, as [14, 29] propose. The local BA is conducted within the process pipeline, as a last refining step on the new camera and 3D points. The global BA is executed parallel to the sequential pipeline over the whole point cloud and the last n camera poses (empirically, n=10).

3.5. Final Scene Recovery

The combination of limited camera resolution, image noise and small baselines inherent within the use of an omnidirectional mobile platform forces our core SfM method to be highly restrictive over the quality of matches. This produces a sparse scene reconstruction resulting in a sparse 3D point cloud of scene surfaces compared to traditional SfM approaches [14]

In order to provide a dense surface reconstruction (e.g. as shown in Fig. 4), a variant of the SfM pipeline is run as a data post-process. This variant makes use of the estimated camera poses and the extracted features. Since the motion is fixed, there is no inherent risk in now including noisy matches and thus we can relax the thresholds of the match quality filters (from Section 3.1). Particularly, τ_u is more benign (set to $\tau_u = 0.65$) and there is no selection over the score on δ_{ab}. This arrangement produces a point cloud whose population is increased up to 200% in terms of recovered 3D scene surface points (see Fig. 2). Note that 4,303 features are extracted from an average image, and the final point cloud has 1,675 views per image, which gives 38.82% of features matched over the total features extracted
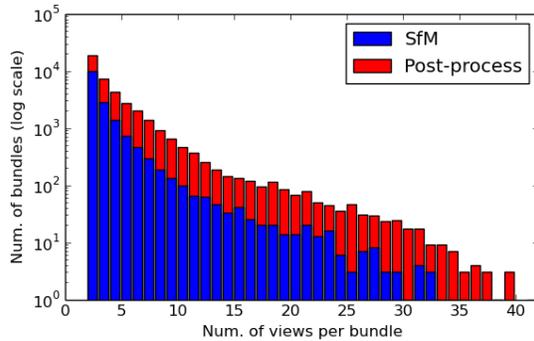
Fig. 2: Histogram of views per bundle for the laboratory sequence. With the post-processing the number of views per bundle increases drastically. The number of bundles containing 3 or more views is raised from from 6.5K to 21.7K.

(a) Laboratory. (b) Industrial environment.
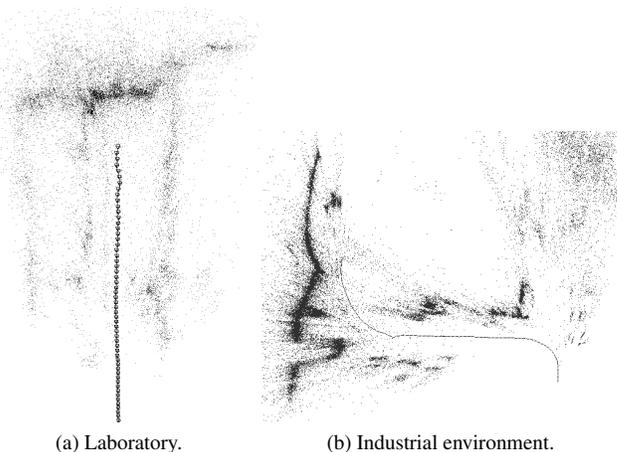
Fig. 3: Odometry and 3D points obtained for two of the sequences used in the experiments.

per image. Fig. 2 compares the length of the tracks, or bundles, obtained with the SfM reconstruction before and after applying this post-process variant for 3D point improvement. The relaxation on the matching filters produces a larger number of image projections, that will be available for a posterior bundle adjustment. At this stage the only filtering realised is commanded by the fixed camera poses through filters on reprojection error.

The final point cloud is filtered by statistical techniques [30] over which a smooth surface is estimated by Moving Least Squares surface reconstruction [31] and using a Poisson method [32] (see Fig. 4).

4. RESULTS

In our experiments we used the low-budget mobile robot Rovio (WowWee Rovio). This robot platform is controlled by three wheels on a radial axis (Fig. 1a) which endows it with omnidirectional movement. The Rovio platform is controlled by wireless communication based on an established API [33]. As a low-budget platform, it is equipped with a 640×480 resolution camera which can be craned within a height range of 10-30 cm, from the surface being transited. We present two experiments in different environments and

	3D Points	Projections	Avg. Rep. Error
SfM with feature tracking	40,128	157,004	1.22
PhotoScan	15,067	59,671	11,74
VisualSfM	4,401	31,498	2.31

Table 2: Comparison of reconstruction accuracy obtained on the industrial sequence with our system, PhotoScan, and VisualSfM. The latter is only able to reconstruct 66 cameras out of 75.

Fig. 5 evaluates the accuracy in the estimation of the camera poses of our system. Here the trajectories estimated by each system and the ground-truth of the path followed by the platform on the industrial experiment are shown. VisualSfM is not in this comparison since it only manages to reconstruct 66 cameras in this experiment. Fig. 5 only shows the last stretch of the sequence as the difference of the camera poses with the ground-truth in the first cameras is negligible. Although both PhotoScan and our system perform similarly, it is notable that the path described by our system consistently matches the trajectory of the ground-truth.

Fig. 4: **Top row**: sample images in the laboratory sequence. **Bottom row:** 3D surface representation obtained with our SfM system.

compare our system with two state of the art implementations: the commercial package PhotoScan (version 1.1.0) from AgiSoft LLC, used in other research works [34, 35] and VisualSfM [36, 37], an interactive application for 3D reconstruction using SfM techniques.

In our experiments a bilateral filter is applied with the diameter of 3 pixels and both colour and spatial filter sizes as $\sigma = 50$. Based upon this pre-filtering, up to 5,000 features are extracted per image, varying on inter-image overlap. Our proposed filtering method identifies a maximum of 700 pair-wise feature matches in optimal matching conditions.

In our first test scenario 55 images were taken over a distance of 6 metres. Here the robot platform performed an approximately straight translation. The 2 dimensional map derived from the estimated camera poses is shown in Fig. 3a. Fig. 4 shows the reconstructed 3D scene. Fig. 4 shows two 3D representations of the laboratory environment where, despite significant noise, the key scene features remain apparent.

In the second experiment the platform performs specific omni-directional movements along a sequence of 75 images. The path and orientation of the robot estimated by our system can be seen in figure 3b. In this experiment SIFT [24] descriptors were used.

Tables 1 and 2 show a comparison in the results given by our system, PhotoScan and VisualSfM. Our system clearly outperforms the other two, providing more 3D structure points at lower reprojection error. The reprojection error is measured as the averaged root mean square of the residuals.

	3D Points	Projections	Reproj. Error
SfM with feature tracking	24,393	100,753	1.53
PhotoScan	8,783	38,534	44.06
VisualSfM	4,288	35,789	4.51

Table 1: Comparison on the laboratory sequence with PhotoScan and VisualSfM.

Fig. 5: Comparison of ground-truth translation, and the odometry estimated using our SfM system, and PhotoScan, for the industrial environment data-set. The image shows the estimated camera location for the last part of the tracking.

5. CONCLUSIONS

We have demonstrated that the proposed noise-tolerant feature tracking method facilitates the effective implementation of Structure From Motion on low-cost omnidirectional robots. These low-budget holonomic platforms produce high levels of image noise ($\sigma \backsim 10$) and narrow inter-image baselines, which, after the application of strict noise filters result in sparse but reliable image feature matches. The feature tracking system maximises the length of feature tracks by an efficient management of the *bundles* created between a 3D point and its views on the image sequence.

Our SfM reconstruction system, which includes this tracking method, succeeds in producing reliable scene reconstruction with low reprojection errors. We compared its performance with different state of the art SfM systems, showing the advantages of our approach in terms of quantity and quality of the resulting 3D scene reconstruction.

6. REFERENCES

[1] C. Ren and S. Ma, "Dynamic modeling and analysis of an omnidirectional mobile robot," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference on*, 2013, pp. 4860–4865.

[2] A. El-Shenawy, A. Wellenreuther, A.S. Baumgart, and E. Badreddin, "Comparing different holonomic mobile robots," in *Systems, Man and Cybernetics, 2007. ISIC. IEEE International Conference on*, Oct 2007, pp. 1584–1589.

[3] M. Udengaard and K. Iagnemma, "Design of an omnidirectional mobile robot for rough terrain," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*, May 2008, pp. 1666–1671.

[4] Y. Ueno, T. Ohno, K. Terashima, and H. Kitagawa, "The development of driving system with differential drive steering system for omnidirectional mobile robot," in *Proc. in Mechatronics and Automation*, 2009, pp. 1089–1094.

[5] Y. Liu, J. Zhao, J. Apple, T. Frank, M. Saylor, and T. Siegel, "An autonomous omnidirectional robot," *Journal of Robotics*, 2010.

[6] D. Ortin and J. M. M. Montiel, "Indoor robot motion based on monocular images," *Robotica*, vol. 19, pp. 331–342, 2001.

[7] L. Maohai, H. Bingrong, and L. Ronghua, "Novel method for monocular vision based mobile robot localization," in *Proc. in International Conference on Computational Intelligence and Security*. IEEE, 2006, vol. 2, pp. 949–954.

[8] E. Royer, M. Lhuillier, M. Dhome, and J. Lavest, "Monocular vision for mobile robot localization and autonomous navigation," *International Journal of Computer Vision*, vol. 74, no. 3, pp. 237–260, 2007.

[9] O. Shakernia, R. Vidal, and S. Sastry, "Structure from small baseline motion with central panoramic cameras," in *Proc. in Computer Vision and Pattern Recognition Workshop*, 2003, vol. 7, pp. 83–89.

[10] R. Szeliski and S. B. Kang, "Recovering 3d shape and motion from image streams using nonlinear least squares," *Journal of Visual Communication and Image Representation*, vol. 5, no. 1, pp. 10 – 28, 1994.

[11] T. Jebara, A. Azarbayejani, and A. Pentland, "3D structure from 2D motion," *Proc. in Signal Processing Magazine, IEEE*, vol. 16, no. 3, pp. 66–84, 1999.

[12] R. Vidal and J. Oliensis, "Structure from planar motions with small baselines," in *Proc. in European Conference in Computer Vision*, pp. 383–398. Springer, 2002.

[13] M. Hebert, "Robust tracking and structure from motion with sample based uncertainty representation," *Proc. in 2002 IEEE International Conference on Robotics and Automation*, pp. 3030–3037.

[14] E. Mouragnon, M. Lhuillier, M. Dhome, F. Dekeyser, and P. Sayd, "Real time localization and 3d reconstruction," in *Proc. in Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, 2006, vol. 1, pp. 363–370.

[15] R. L. Williams, B. E. Carter, P. Gallina, and G. Rosati, "Dynamic model with slip for wheeled omnidirectional robots," *IEEE Transactions on Robotics and Automation*, vol. 18, no. 3, pp. 285–293, 2002.

[16] A. Begum, M. Lee, and Y. J. Kim, "A simple visual servoing and navigation algorithm for an omnidirectional robot," in *Proc. in 3rd International Conference on Human-Centric Computing*. 2010, pp. 1–5, IEEE.

[17] Z. Zhang, "A robust technique for matching two uncalibrated images through the recovery of the unknown epipolar geometry," *Artificial Intelligence*, vol. 78, no. 1-2, pp. 87–119, 1995.

[18] G. Zhang, Z. Dong, J. Jia, T. Wong, and H. Bao, "Efficient non-consecutive feature tracking for structure-from-motion," in *Proc. in European Conference on Computer Vision: Part V*. 2010, pp. 422–435, Springer-Verlag.

[19] R. I. Hartley and P. Sturm, "Triangulation," *Computer Vision and Image Understanding*, vol. 68, no. 2, pp. 146–157, 1997.

[20] J. I. Thomas and J. Oliensis, "Dealing with noise in multiframe structure from motion," *Computer Vision and Image Understanding*, vol. 76, no. 2, pp. 109–124, 1999.

[21] C. Tomasi and R. Manduchi, "Bilateral filtering for gray and color images," in *In Proc. Int. Conf. Computer Vision*. IEEE, 1998, pp. 839–846.

[22] H. Bay, T. Tuytelaars, and L. Van Gool, "SURF: Speeded up robust features," in *Proc. in European Conference in Computer Vision*, 2006, pp. 404–417.

[23] S. Arya, D. M. Mount, N. S. Netanyahu, R. Silverman, and A. Y. Wu, "An optimal algorithm for approximate nearest neighbor searching in fixed dimensions," in *Proc. in ACM-SIAM symposium on discrete algorithms*, 1994, pp. 573–582.

[24] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *International Journal of Computer Vision*, vol. 60, pp. 91–110, 2004.

[25] M. A. Fischler and R. C. Bolles, "Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography," *Commun. ACM*, vol. 24, no. 6, pp. 381–395, 1981.

[26] R. I. Hartley and A. Zisserman, *Multiple view geometry in computer vision*, Cambridge University Press, 2004.

[27] V. Lepetit, F. Moreno-Noguer, and P. Fua, "EPnP: An Accurate O(n) Solution to the PnP Problem," *International Journal of Computer Vision*, vol. 81, no. 2, pp. 155–166, 2009.

[28] M. I. A. Lourakis and A. A. Argyros, "SBA: A Software package for generic sparse bundle adjustment," *ACM Trans. Math. Software*, vol. 36, no. 1, pp. 1–30, 2009.

[29] C. Engels, H. Stewénius, and D. Nistér, "Bundle adjustment rules," *Photogrammetric computer vision*, vol. 2, 2006.

[30] R. B. Rusu and S. Cousins, "3D is here: Point Cloud Library (PCL)," in *Proc. in International Conference on Robotics and Automation*, 2011.

[31] M. Alexa, J. Behr, D. Cohen-Or, S. Fleishman, D. Levin, and C. T. Silva, "Computing and rendering point set surfaces," *IEEE Transactions on Visualization and Computer Graphics*, vol. 9, no. 1, pp. 3–15, 2003.

[32] M. Kazhdan, M. Bolitho, and H. Hoppe, "Poisson surface reconstruction," in *Proc. in the Fourth Eurographics Symposium on Geometry Processing*, 2006, pp. 61–70.

[33] "Rovio api c++ class library," https://github.com/tobybreckon/roviolib, 2011-2013.

[34] G. Verhoeven, M. Doneus, Ch. Briese, and F. Vermeulen, "Mapping by matching: a computer vision-based approach to fast and accurate georeferencing of archaeological aerial photographs," *Journal of Archaeological Science*, vol. 39, no. 7, pp. 2060 – 2070, 2012.

[35] G. Verhoeven, "Taking computer vision aloft – archaeological three-dimensional reconstructions from aerial photographs with photoscan," *Archaeological Prospection*, vol. 18, no. 1, pp. 67–73, 2011.

[36] C. Wu, "Visualsfm: A visual structure from motion system," http://ccwu.me/vsfm/, 2011.

[37] C. Wu, S. Agarwal, B. Curless, and S.M. Seitz, "Multicore bundle adjustment," in *2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2011, pp. 3057–3064.