3D model-based video mapping

Thomas Guerneve
Heriot Watt University
SeeByte Ltd
Edinburgh, Scotland
Email: thomas.guerneve@seebyte.com

Kartic Subr
Edinburgh University
Edinburgh, Scotland
Email: K.Subr@ed.ac.uk

Yvan Petillot
Heriot Watt University
Edinburgh, Scotland
Email: Y.R.Petillot@hw.ac.uk

***Abstract*—Acquiring accurate and coloured 3D representations of underwater scenes is of prime importance in many applications such as ship hull inspection, offshore structure assessment, archaeology or the detection of hazards such as underwater mines. While video and laser-based 3D reconstruction approaches are limited to very short ranges, the use of SONAR sensors enable the recognition and localisation of man-made structures at centimetre-level. We combine mid-range SONAR mapping with video mapping onto a high-resolution CAD model of the object of interest. We apply this 3D model-based video mapping approach on field data gathered by AUV during lake trials. Through this experiment, we demonstrate the interest in using multi-modal sensing and reference CAD models to obtain high-resolution coloured 3D representations of man-made objects.**

I. Introduction

The acquisition of accurate 3D representations of submerged man-made structures is of great interest when performing offshore structure maintenance, ship-hull inspection or mine detection and classification. While the acquisition of this 3D representation can be made using various modalities, all these solutions exhibit major drawbacks.

Video mapping approaches typically require pristine lighting conditions and dense 3D reconstruction traditionally depend on the presence of texture on the reconstructed surface [1]. Laser-based systems enable accurate range measurements of texture-less objects but typically feature low footprints [2], yielding poor surface coverage or conversely requiring low inspection speed. Additionally, both of these methods require good visibility conditions [3], [4] making their employability weather-dependent and restricted to very short range observations. Operating underwater vehicles at close distances to man-made structures is in general risky if not impractical. Thanks to favourable propagation properties in water, acoustic methods have been consistently employed to acquire 3D representations at short to large observation ranges with centimetre-level accuracies [5]. While SONAR measurements are typically corrupted by noise and multi-path effects, 3D reconstructions from SONAR data still feature enough detail for human-assisted or automated object recognition. Once identified, CAD model representations of the object of interest can be employed to provide a complete high-resolution 3D representation of the object.

Using data collected by the Subsea 7 AIV (Autonomous Inspection Vehicle) prototype, we demonstrate here a high-resolution underwater 3D optical mapping method based on the SONAR-aided localisation of a 3D model in the scene. Once the object is fully reconstructed using standard 2D imaging SONARs, its associated CAD model is registered, providing a reference 3D representation allowing both online relocalisation and high-resolution optical mapping.

II. Related work

As reviewed in [6], the reconstruction of 3D models of underwater environments has been investigated through the use of various sensing modalities. The addition of colour information using video data has been consistently gaining interest to provide photo-realistic models.

Through the use of robust triangulation methods [7] and multiple 2D optical images, the geometry of a coloured 3D scene can be recovered. When enough features are present in the scene, both the camera calibration and the 3D geometry of the scene can be estimated [8]. As a result, underwater reconstruction from monocular imagery has been investigated with the employment of structure-from-motion techniques [1] but typically provide sparse reconstructions due to the low amount of features in underwater images. Stereo cameras facilitate the reconstruction process by providing on-the-shelf calibration between the two views, they have been applied successfully for so-called dense reconstructions [9] but remain sensitive to outliers such as particles. Recent work [10] addressed the robustness of the reconstruction process to noise and outliers typically provided by underwater sensing modalities but remain dependent on good visibility conditions and require large computation durations. In poor visibility conditions, the methods relying only on video data typically require short-range inspections putting the inspection equipment or the inspected elements at risk. In practice, safety margins are applied when operating around man-made structures, typically imposing stand-off distances of at least a metre.

In this situation, there is an interest in acquiring a 3D representation using a midrange sensing modality enabling a second inspection at shorter range to acquire video data with improved visibility. When both sensors are mounted on the same platform, joint calibration procedures have been investigated for laser-camera systems [11] as well as on notoriously more noisy multibeam SONARs [12]. Experimental results of video mapping on representations acquired from multiple depth sensors including infrared scanner have been presented in [13], exhibiting good accuracy but remained limited to very short range and depth (less than a metre). Large-scaled underwater mapping was demonstrated in [14] where the authors fused multibeam SONAR and video data to simultaneously fix navigation drift and generate a coloured 3D map of the seabed.

When the acquisition of respectively the 3D geometry and the visual information are performed separately, a registration procedure between the two representations is needed. The mapping of video data on a 3D model acquired from a laser-scanner in the air has been investigated in [15]. Interestingly, the authors registered the video data onto the 3D representation by associating the 2D lines observed in the images to the 3D lines present in the scene.

In opposition to the previously cited methods, we propose to operate midrange 3D mapping and take advantage of an available CAD model representation of the object to achieve high-resolution and dense video mapping. To the best of our knowledge, this work is the first demonstration of model-based underwater video mapping.

III. Method

We first generate a rough 3D representation using multiple 2D SONAR images of the object of interest. We then apply a two-step procedure to register a CAD model of the object in the 3D point cloud. We finally map the video data onto the registered model to obtain a photo-realistic 3D representation.

A. 3D reconstruction from 2D SONAR images

Based on on-board navigation and 2D pencil-beam SONAR data, a first 3D representation is obtained. In order to solve the illness of the 3D reconstruction problem from 2D samples, we applied a space carving reconstruction method [5] where the space occupancy along the missing dimension (vertical aperture of the SONAR) is solved by observing empty spaces. A 3D anisotropic map is maintained in an Octree structure [16] and refined every time a new SONAR image is acquired. Multiple 2D views of the object are successively acquired by moving along the direction of uncertainty (perpendicular to the SONAR plane). An occlusion resolution step enables to solve the acoustic shadowing phenomenon happening during the imaging process. The reconstruction method finally provides a 3D point cloud of the surface of the scene observed by the multibeam SONAR.

B. CAD model registration

Once the object reconstructed, a CAD model of the object is registered in 6D with a two-step registration procedure. The object of interest is assumed to be laid on the seafloor, therefore featuring an orientation based on the local geometry of the seabed. While the model registration is a 6D problem, the rotation angle of the model on two of the axis (roll and pitch) is expected to be very small. As a consequence the first registration step is performed in 4D while the second step provides 6D registration.

1) 4D elevation map registration: In order to adopt a representation adapted to the 4D matching problem, the point cloud issued from the 3D reconstruction is first converted to an elevation map where each pixel represents a (*North,East*) coordinate and contain an elevation value (*Depth*) as an intensity. Due to the small size of the patch, an exhaustive 4D matching is performed, considering a discrete set of translations along the North and East axis and structure orientations in $[0°, 180°]$ or $[0°, 360°]$ angular intervals, depending on the symmetry of the object. For each of these configurations, the depth of the model is first set to the depth of the map by aligning the signed median values of the elevation distributions of the two maps to be compared. Once the depth of the model computed, an elevation map of the model is generated and compared to the scene using an euclidean distance, normalized over the number of overlapping pixels. The optimal registration is then chosen as the minimum euclidean distance between the transformed model and the scene. This registration 4D offset is then applied to the model and constitutes the initial position of the second registration step.

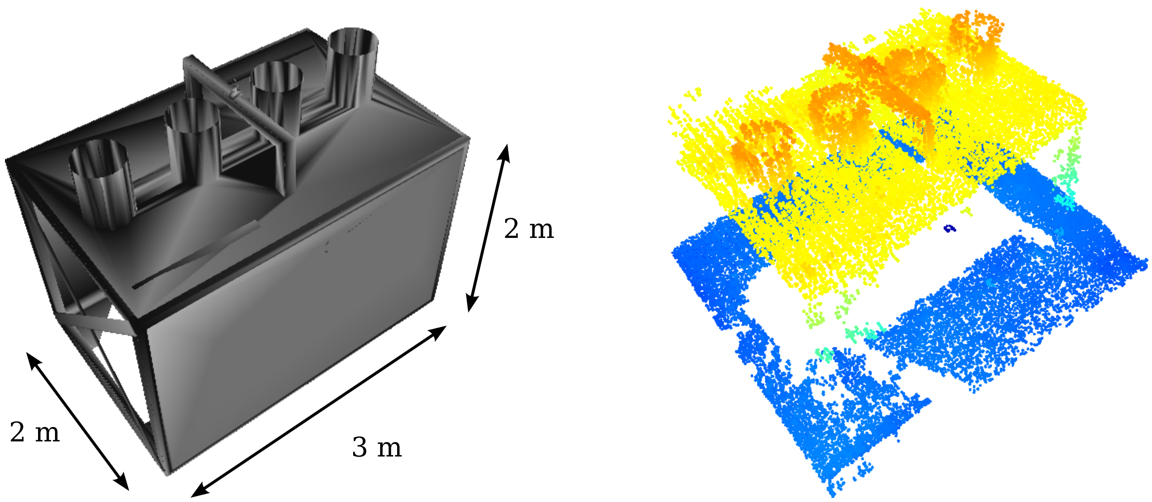
2) 6D ICP-based registration: In situations where objects are placed irregularly or when the seabed below the object is not flat, the model can feature small *Pitch* and *Roll* angles. In order to account for possible offsets in these two dimensions, a robust ICP-based method is applied [17], [18] by discarding the furthest points (outliers) for registration.

C. Video mapping

Once the model registered in the 3D scene, the position of the vehicle relatively to the object is known. In particular, when the video camera is present on the same platform operated for the SONAR inspection, a simple measurement of the position of the camera relatively to the SONAR sensor gives the positions of the camera relatively to the reconstructed point cloud. In our case, we assume these sensor offsets known and assume that both the SONAR data and the video data have been acquired on the same platform during the same inspection (see discussion section). Once the position of the camera known for each image, a direct projection of the colour information is performed through a raytracing process [19]. Knowing the parameters of the camera (horizontal and vertical field of view, focal point), the model is discretized in a fine voxel grid and parsed in an Octree structure allowing look-up requests and raytracing with the PCL library [20]. Since multiple views of each 3D point can be acquired during the structure inspection, the colour reading with the highest intensity is kept in each voxel.

IV. Experimental results

In order to demonstrate our 3D model-based video mapping approach, we present field experiments on a dataset acquired by an AUV. The dataset was gathered by the Subsea 7 AIV during trials in Loch Eil in Scotland. On the site of the trials, a structure mimicking an oil-field rig structure (see fig 1a) was placed on the seabed at approximately 30 metres depth. A pencil-beam (1° aperture) multibeam



(a) CAD model of the inspected structure (b) 3D reconstruction from SONAR data

Fig. 1. 3D reconstruction of oil-field-type structure from multibeam SONAR observations.

BlueView MB2250 SONAR mounted in a downward-looking configuration acquired vertical samples during the inspection with a range resolution of 1cm and a bearing resolution of 1°. The vehicle inspected the structure from above at a distance of 2m, following a rectilinear motion along the principal direction of the structure with images acquired at 4cm along-track sampling period. A downward-looking high-definition M12 CATHX video camera provided 10 images of the top of the structure. The images were captured at a frequency of 2Hz corresponding to 20cm motion between each frame. The navigation data was provided by an on-board navigation module at 10Hz frequency, integrating readings from a DVL, a gyroscope and a compass. Bilinear interpolation was applied on the navigation samples to provide an estimate of the exact sampling position of respectively SONAR and video measurements.

Figure 1b presents the 3D reconstruction of the structure obtained using the space carving reconstruction method. An elevation map at 8cm resolution was generated and the two-step matching procedure described in section III-B was applied on the CAD model in 50sc. In order to quantify the quality of the registration, the points representing the seabed were removed manually after registration and the median value of the unsigned distances between the remaining points and the CAD model was computed. The median values obtained after respectively the first step and the second step were 8.3cm and 2.4cm. Figure 2a depicts the result of the model registration.

The last part of the experiment consisted in projecting the images gathered during the top inspection. For reference, we present in figure 2b the coloured 3D point cloud resulting from the projection of video data onto the 3D reconstruction of figure 1b. Figure 3a depicts the result of the video reprojection on the registered CAD model.

V. Discussion

A. Analysis

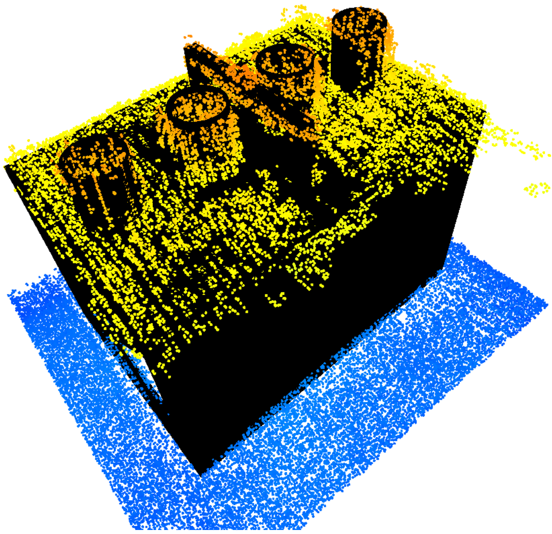
The 3D reconstruction presented in fig 1b exhibits enough accuracy for visual identification and manual registration of the structure. Our two-step registration method enabled an automatic registration with only 2.4cm median registration error. Considering the relatively large along-track sampling period (4cm) and the artefacts typically observed in SONAR images (multipath returns and noise), 2cm is the expected reconstruction error suggesting that the final registration is close to optimal.

The projection of the video data on the reconstructed point cloud (fig 2b) exhibits the limited interest of mapping high-resolution information (images) onto a low-resolution 3D representation. The resulting coloured point cloud features low-resolution colour information which does not provide a significant improvement in the visualization of the object. On the contrary, the video mapping onto the model exhibits high-resolution information such as the visible presence of mud on the top of structure.

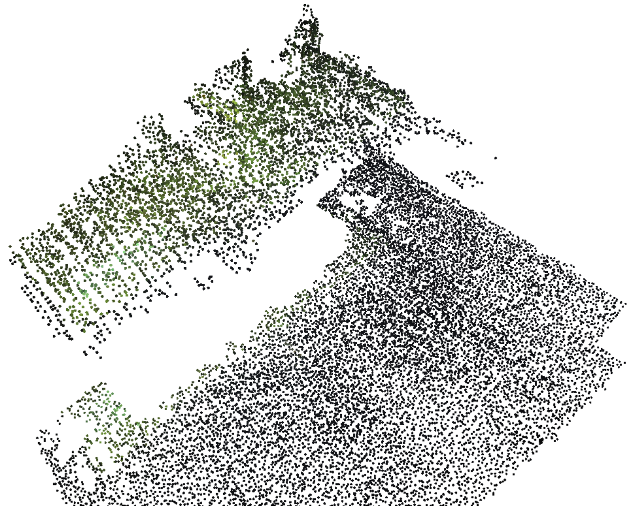
B. Limitations and future work

Although a lot of details are visible on the coloured model, a few artefacts are visible. Due to the limitations of the raytracing algorithm used, unrealistic coloured points appear at the bottom of the structure. These artefacts could be addressed by performing the raytracing onto a mesh (continuous surface) rather than a discrete set of points.

More importantly, discontinuities appear in the coloured model as visible on the right-hand long edge of the model. These discontinuities are expected to be due to local navigation inaccuracies (in particular roll and pitch estimation) as well as the limited accuracy in the estimation of camera parameters. The employment of camera calibration techniques such as [21] would benefit to a more accurate reconstruction. Once the

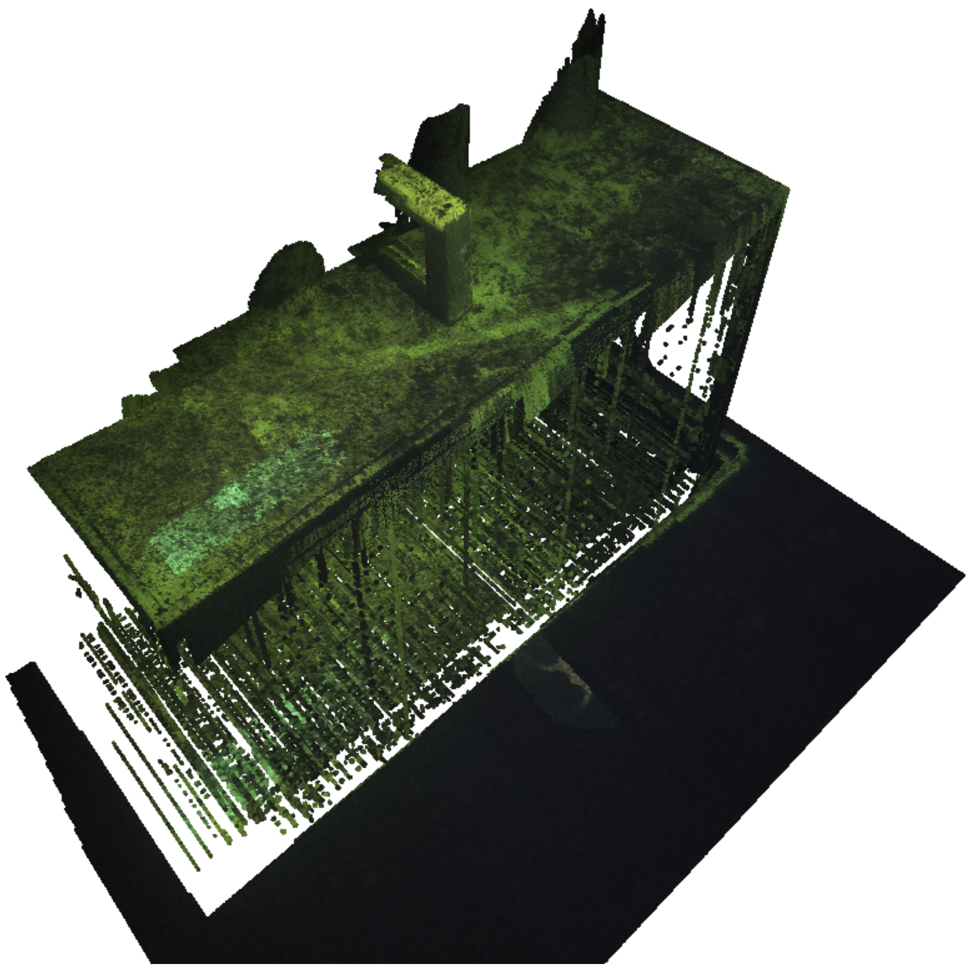


(a) Registered CAD model (black) after the two step registration procedure



(b) Video projection on the 3D reconstruction

Fig. 2. 3D registration and video projection on 3D reconstruction from SONAR data.



(a) Video projection on the registered CAD model

Fig. 3. 3D registration and video projection on 3D reconstruction from SONAR data.

intrinsic parameters known, the presence of features in the images could correct the small scale navigation drift between each image by feature matching and bundle adjustment.

While this experiment was performed on video and SONAR data acquired at the same time, similar results could be obtained using data collected during different inspections. In particular, one could imagine a first long range inspection covering the whole structure of interest in a safely manner. Based on the SONAR data acquired during this first inspection, a 3D reconstruction of the whole structure would be obtained, enabling a robust localisation of the model through a registration. A second inspection at shorter range would then improve the visibility and allow the acquisition of detailed views of the structures or part of the structures. In this situation, the eventual navigation offset between the two inspections could be mitigated by matching a new 3D representation acquired during the second inspection using the same registration technique as detailed in section III-B.

A limitation of model-mapping approach arises in situation when the geometry of the object differs from the available model. In particular, the accumulation of mud, marine growth and corrosion typically alter the geometry of submerged objects. In this situation, a CAD model of the structure made before deployment on the field is not an accurate 3D representation of the inspected object but the use of a high-accuracy ranging device such as a laser scanner instead of a SONAR could provide a similar level of accuracy to the model.

VI. Conclusion

This paper presented a method providing a high-resolution, 3D textured representation based on 3D mapping from a multibeam SONAR data and the registration of a CAD model of the object of interest. A field experiment with an AUV demonstrated the performance of the method as well as the interest in using multiple sensing modalities to achieve high-resolution inspection of man-made structures. While SONAR-based 3D reconstructions are limited in resolution, they feature enough information for object identification and centimetre-level localisation enabling increased autonomy. In contrast with the low-resolution 3D reconstruction obtained from multibeam SONAR data, the availability of a high-resolution 3D model enabled the generation of a dense coloured 3D model of a man-made structure. Future work will focus on improving both the camera calibration and the local navigation accuracy through the use of visual features. The use of a laser-scanner will also be considered to provide accurate coloured reconstructions of elements of the scene for which no CAD model is available.

Acknowledgment

The authors would like to thank Subsea 7 for providing access to the AIV data. This project was funded under the Marie Curie ITN program Robocademy FP7-PEOPLE-2013-ITN-608096.

References

[1] C. Beall, B. J. Lawrence, V. Ila, and F. Dellaert, "3d reconstruction of underwater structures," in *Intelligent Robots and Systems (IROS), 2010 IEEE/RSJ International Conference on*. IEEE, 2010, pp. 4418–4423.

[2] C. Roman, G. Inglis, and J. Rutter, "Application of structured light imaging for high resolution mapping of underwater archaeological sites," in *OCEANS 2010 IEEE-Sydney*. IEEE, 2010, pp. 1–9.

[3] A. Sedlazeck, K. Koser, and R. Koch, "3d reconstruction based on underwater video from rov kiel 6000 considering underwater imaging conditions," in *OCEANS 2009-EUROPE*. IEEE, 2009, pp. 1–10.

[4] F. Bruno, G. Bianco, M. Muzzupappa, S. Barone, and A. Razionale, "Experimentation of structured light and stereo vision for underwater 3d reconstruction," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, no. 4, pp. 508–518, 2011.

[5] T. Guerneve and Y. Petillot, "Underwater 3d reconstruction using blueview imaging sonar," in *OCEANS 2015-Genova*. IEEE, 2015, pp. 1–7.

[6] M. Massot-Campos and G. Oliver-Codina, "Optical sensors and methods for underwater 3d reconstruction," *Sensors*, vol. 15, no. 12, pp. 31 525–31 557, 2015.

[7] R. I. Hartley and P. Sturm, "Triangulation," *Computer vision and image understanding*, vol. 68, no. 2, pp. 146–157, 1997.

[8] Z. Zhang, Q.-T. Luong, and O. Faugeras, "Motion of an uncalibrated stereo rig: Self-calibration and metric reconstruction," *IEEE Transactions on Robotics and Automation*, vol. 12, no. 1, pp. 103–113, 1996.

[9] M. Johnson-Roberson, O. Pizarro, S. B. Williams, and I. Mahon, "Generation and visualization of large-scale three-dimensional reconstructions from underwater robotic surveys," *Journal of Field Robotics*, vol. 27, no. 1, pp. 21–51, 2010.

[10] R. Campos, R. Garcia, P. Alliez, and M. Yvinec, "A surface reconstruction method for in-detail underwater 3d optical mapping," *The International Journal of Robotics Research*, vol. 34, no. 1, pp. 64–89, 2015.

[11] Q. Zhang and R. Pless, "Extrinsic calibration of a camera and laser range finder (improves camera calibration)," in *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, vol. 3. IEEE, 2004, pp. 2301–2306.

[12] N. Hurtós, X. Cufi, and J. Salvi, "Calibration of optical camera coupled to acoustic multibeam for underwater 3d scene reconstruction," in *OCEANS 2010 IEEE-Sydney*. IEEE, 2010, pp. 1–7.

[13] A. Dancu, M. Fourgeaud, Z. Franjcic, and R. Avetisyan, "Underwater reconstruction using depth sensors," in *SIGGRAPH Asia 2014 Technical Briefs*. ACM, 2014, p. 2.

[14] C. Kunz and H. Singh, "Map building fusing acoustic and visual information using autonomous underwater vehicles," *Journal of field robotics*, vol. 30, no. 5, pp. 763–783, 2013.

[15] I. Stamos and P. Allen, "3-d model construction using range and image data," in *Computer Vision and Pattern Recognition, 2000. Proceedings. IEEE Conference on*, vol. 1. IEEE, 2000, pp. 531–536.

[16] D. Meagher, "Geometric modeling using octree encoding," *Computer graphics and image processing*, vol. 19, no. 2, pp. 129–147, 1982.

[17] P. J. Besl and N. D. McKay, "Method for registration of 3-d shapes," in *Robotics-DL tentative*. International Society for Optics and Photonics, 1992, pp. 586–606.

[18] T. Masuda, K. Sakaue, and N. Yokoya, "Registration and integration of multiple range images for 3-d model construction," in *Pattern Recognition, 1996., Proceedings of the 13th International Conference on*, vol. 1. IEEE, 1996, pp. 879–883.

[19] A. S. Glassner, *An introduction to ray tracing*. Elsevier, 1989.

[20] R. B. Rusu and S. Cousins, "3d is here: Point cloud library (pcl)," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 1–4.

[21] Z. Zhang, "A flexible new technique for camera calibration," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 22, no. 11, pp. 1330–1334, 2000.