

## COMPARISON OF TWO MATHEMATICAL MODELS FOR FISHEYE CAMERAS APPLIED IN THE ORB-SLAM

T. A. C. Garcia <sup>1\*</sup>, A. M. G. Tommaselli <sup>1</sup>, L. F. Castanheiro <sup>1</sup>, M. B. Campos <sup>2</sup>

<sup>1</sup> Department of Cartography, Faculty of Sciences and Technology, São Paulo State University (UNESP),  
São Paulo 19060-900, Brazil - (thaisa.correia, a.tommaselli, leticia.ferrari)@unesp.br

<sup>2</sup> Department of Remote Sensing and Photogrammetry, Finnish Geospatial Research Institute (FGI), National Land Survey of  
Finland, Espoo, Finland - mariana.campos@nls.fi

**KEY WORDS:** SLAM, ORB-SLAM, real-time, omnidirectional system, fisheye lenses, Ricoh Theta S.

### ABSTRACT:

Fisheye lens cameras are becoming increasingly popular for vSLAM applications due to their wide field of view (FoV), providing more features to be tracked in a single image shot. However, the complex lens geometry involved in the image formation process still limits their full potential, especially when points in the hyperhemispherical field are unmodeled. In this paper, we compare two adaptations of ORB-SLAM for fisheye lens cameras, considering the use of the rigorous projection model (equisolid-angle) versus the use of the generic projection model (EUCM). The ORB-SLAM versions were adapted for real-time processing on the Nvidia Jetson TX2 board. The experiment was conducted using hyperhemispherical images obtained with a Ricoh Theta S camera. Our results showed that the trajectory estimated with the equisolid-angle ORB-SLAM had smaller discrepancies, compared to the reference trajectory, than the EUCM ORB-SLAM. This suggests that a rigorous photogrammetric model with a suitable treatment of hyperhemispherical points is beneficial for trajectory estimation.

### 1. INTRODUCTION

The 3D reconstruction of an unknown environment with simultaneous determination of the sensor orientation is commonly known as SfM (Structure from Motion) in Computer Vision, Phototriangulation in Photogrammetry, and SLAM (Simultaneous Localization and Mapping) in robotics (Cadena et al., 2016; Durrant-Whyte and Bailey, 2006; Granshaw, 2018). The classic problem of SLAM is to sequentially estimate the position and orientation of an agent platform in real-time based on remote sensors, such as cameras and laser scanning systems, mounted in mobile mapping platforms. Thus, SLAM methods enable a consistent estimate of the trajectory based on the map of the environment that contains 3D points. SLAM methods based only on sequential images information are well-known as Visual SLAM (vSLAM) (Li et al., 2019; Taketomi et al., 2017; Torresani and Remondino, 2019).

In recent years, we have witnessed significant advances in vSLAM technology to address the challenges of building accurate and robust maps in various environments. These advances were especially driven by the increased use of low-cost and lightweight optical sensors and high computational power and performance, resulting in state-of-the-art approaches such as: Mono-SLAM (Davison et al., 2007), PTAM (Klein and Murray, 2007), DTAM (Newcombe et al., 2011), SVO (Forster et al., 2014), LSD-SLAM (Engel et al., 2014), ORB-SLAM (Mur-Artal et al., 2015) e DSO (Matsuki et al., 2018).

ORB-SLAM is considered one of the leading state-of-the-art SLAM solutions due to its input data flexibility, versatility and accurate estimative. It uses the ORB operator to extract keypoints and match their features, which allows it to work with images acquired using different configurations, such as monocular, stereo, and RGB-D (red, green, blue, and depth) cameras. The matching points enable the estimation of the

trajectory and sparse 3D cloud in real-time, in addition to loop closures and relocating (Mur-Artal et al., 2015). However, a vSLAM solution combining fisheye images and ORB-SLAM has not yet been fully explored.

Fisheye lenses provide a wider field of view (FoV), enabling the capture of more features in a single image shot. As a result, they are an attractive choice for use in vSLAM solutions. However, many challenges in the use of large FoV systems (e.g., fisheye) in mobile mapping applications still remain. The main issues are related to the complex camera model and fisheye lens geometry (Wang et al., 2018).

Several works have proposed generic models to cope with the fisheye lens geometry, which allow the application in different camera systems (Usenko et al., 2018). Campos et al. (2020) and Liu et al. (2019) used generic camera models to extend ORB-SLAM for use with fisheye cameras. Campos et al. (2020) utilized the camera model developed by Kannala and Brandt (2006), while Liu et al. (2019) employed the Enhanced Unified Camera Model (EUCM), introduced originally by Geyer and Konstantinos (2000). The choice of these models was driven by their generality and simplicity, which enabled simple adaptation for use in ORB-SLAM.

Applications that demand high accuracy may require rigorous projection models that consider the physical imaging principles of fisheye lenses, such as equidistant, orthogonal, stereographic, and equisolid-angle (Schneider et al., 2009).

In this paper, we compare the accuracy of real-time ORB-SLAM solution (Aldegheri et al., 2019) when using generic versus rigorous models for fisheye lenses. The real-time ORB-SLAM, proposed by Aldegheri et al. (2019), was modified into two versions. The first one applies the generic EUCM model

\* Corresponding author

(EUCM ORB-SLAM), and the second one the rigorous equisolid-angle model (equisolid ORB-SLAM), which also allows the treatment of hyperhemispherical points (Castanheiro et al., 2021). Comparative analyses were performed with a fisheye image dataset acquired with backpack-mounted mobile system. Developed by Campos et al. (2018), the mobile mapping system is equipped with a Ricoh Theta S omnidirectional camera, which captures the entire scene with a hyperhemispherical geometry, covering a 360° FoV.

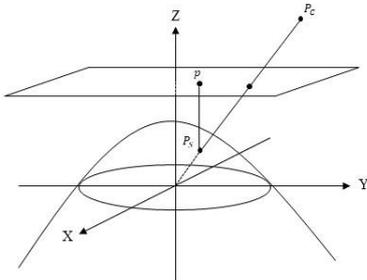
## 2. MATERIALS AND METHODS

### 2.1 Mathematical Models

Fisheye lens cameras have a FoV of approximately 180° (for hemispherical lenses) or even more (for hyperhemispherical lenses), covering a wider view of a scene in a single image compared to conventional perspective cameras. Previous works proposed a different mathematical model to deal with the geometric distortions caused by fisheye lenses and address this specific lens geometry (Usenko et al., 2018). Here, we describe in detail the two mathematical models used in this work to adapt the ORB-SLAM for fisheye lens geometry: the generic EUCM and the rigorous equisolid-angle model.

#### 2.1.1 Enhanced Unified Camera Model

The EUCM is a generic projection model for fisheye cameras based on a unified camera model that is considered simple and does not require additional distortion coefficients. EUCM requires only two more parameters to cope with the distortions. This model allows the inverse projection function to be expressed in an explicit closed form. As shown in Figure 1, the coordinates of a point  $P_C (X_C, Y_C \in Z_C)$ , in the camera reference system are projected to  $P_S$  onto an ellipsoid and then,  $P_S$  is mapped to  $u$  in the image plane using the pinhole model.



**Figure 1.** EUCM geometric model (Khomutenko et al., 2016). Adapted from Liu et al. (2019).

The vector  $u$  with image coordinates  $(u$  and  $v)$ , in pixels, with origin at the bottom-left pixel of the image, are calculated by Equation 1, where  $f$  is the principal distance in pixels,  $c_x$  and  $c_y$  are the coordinates of the origin of a centered system.  $\alpha \in [0.5, 1]$  and  $\beta > 0$ , are the projection parameters on the ellipsoid, which allow approximation of the lens properties, despite the strong distortions, with  $\rho$  being calculated by Equation 2.

$$u = f \frac{X_c}{\alpha\rho + (1-\alpha)Z_c} + c_x \quad (1)$$

$$v = f \frac{Y_c}{\alpha\rho + (1-\alpha)Z_c} + c_y$$

$$\rho = \sqrt{\beta(X_c^2 + Y_c^2) + Z_c^2} \quad (2)$$

The coordinates for the inverse projection from the image plane to point  $P_S$  on the ellipsoid in the EUCM model are obtained using Equation 3, in which  $r^2 = s_x^2 + s_y^2$ . Additionally, the maximum value of  $s_x$  is set to 1. The ellipsoid is the projection surface, and the inverse function is a ray's collinearity function since  $P_S$  and  $P_C$  are on the same ray.

$$s_x = \frac{u - c_x}{f}$$

$$s_y = \frac{v - c_y}{f} \quad (3)$$

$$s_z = \frac{1 - \beta\alpha^2 r^2}{\alpha\sqrt{1 - (2\alpha - 1)\beta r^2} + (1 - \alpha)}$$

Generic models, however, do not consider the physical principles of imaging. Conventional models considers the physical properties of the fisheye lens, which usually are based on the projection of a sphere on the image plane, with several projection models, such as equidistant, orthogonal, stereographic, and equisolid-angle (Abraham and Förstner, 2005; Hughes et al., 2010; Ray, 1994; Schneider et al., 2009). Castanheiro et al. (2021) presented a comparison of fisheye projection models in the camera calibration of the Ricoh Theta S dual-fisheye system to verify which one is suitable for hyperhemispherical lenses. The results were better when using the equisolid-angle projection model.

#### 2.1.2 Equisolid-angle projection model

Figure 2 illustrates the projection of point  $P_S$  onto the image plane as point  $p$ , following the equisolid-angle projection. In the equisolid-angle model, the radial distance  $r_d$  is obtained by the relation  $r_d = 2f \sin(\alpha/2)$ , equivalent to the chord length of the arc segment on the sphere. Equation 4 presents the equisolid-angle projection model, with  $X_c, Y_c$  and  $Z_c$  being the coordinates of a point  $P$  in the camera reference system, which are related to their coordinates in the object reference system  $(X, Y$  and  $Z)$  through a rigid body transformation (Equation 5). This transformation is a function of  $X_0, Y_0 \in Z_0$  (3D coordinates of the camera position) and  $m_{ij}$ , components of the rotation matrix. The coordinates  $x_f$  and  $y_f$  are the projection of the point in the photogrammetric reference system, with  $f$  being camera principal distance.

$$x_f = 2f \frac{X_c}{\sqrt{2(X_c^2 + Y_c^2)}} \sqrt{1 - \frac{Z_c}{X_c^2 + Y_c^2 + Y_c^2}}$$

$$y_f = 2f \frac{Y_c}{\sqrt{2(X_c^2 + Y_c^2)}} \sqrt{1 - \frac{Z_c}{X_c^2 + Y_c^2 + Y_c^2}} \quad (4)$$

$$\begin{aligned} X_c &= m_{11}(X - X_0) + m_{12}(Y - Y_0) + m_{13}(Z - Z_0) \\ Y_c &= m_{21}(X - X_0) + m_{22}(Y - Y_0) + m_{23}(Z - Z_0) \\ Z_c &= m_{31}(X - X_0) + m_{32}(Y - Y_0) + m_{33}(Z - Z_0) \end{aligned} \quad (5)$$



### 3.1 Data acquisition and dataset

The dataset consists of fisheye images collected with a Ricoh Theta S dual-fisheye camera. The camera was attached to a backpack-mounted mobile system, and images were captured over a 140 m long path. Only images captured with one of the camera sensors with a frame rate of 5 fps were used in the experiments. The study area is a path covered by high and low vegetation and urban features (buildings, sidewalks, and posts), as shown in Figure 4.

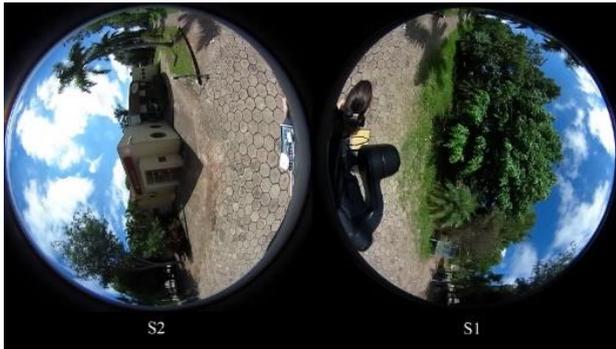


Figure 4. Images from sensor 1 (S1) and sensor 2 (S2).

Additionally, two reference trajectories, in UTM projection coordinates, were generated to assess performance and positional accuracy: (1) sensor trajectory calculated indirectly by simultaneous bundle adjustment with the Agisoft Metashape, based on a set of selected images with 1 fps and 48 GCPs (Ground Control Points) collected using GNSS (Global Navigation Satellite System) RTK (Real Time Kinematic); (2) sensor trajectory estimated by post-processing dual-frequency GNSS positioning, using a Hiper GPS receiver attached to a backpack while walking along the path. The trajectory obtained from the dual-frequency GNSS receiver data presented an average precision of 0.02 m at the beginning of the trajectory (where there was no signal loss), while the trajectory calculated by Agisoft Metashape had a precision of 0.07 m in camera positions and an average discrepancy of 0.15 m compared to the positions at the beginning of the trajectory obtained from the dual-frequency GNSS receiver. The results obtained with ORB-SLAM are referenced to a local coordinate system with an arbitrary origin, orientation, and scale. Therefore, parameters of a similarity transformation were calculated using GCPs to convert the two estimated ORB-SLAM trajectory points coordinates to UTM projection. The seven estimated similarity parameters (3 translations, 3 rotations and scale) were applied to transform the coordinates of the camera stations, estimated in a local reference system by EUCM ORB-SLAM and equisolid ORB-SLAM, to the UTM projection coordinates.

### 3.2 Performance Assessment

Comparative analyses were performed by selecting 296 common frames from: (1) the reference trajectory calculated with Agisoft Metashape (Metashape); (2) the trajectory obtained using the equisolid-angle model with ORB-SLAM (equisolid ORB-SLAM), and (3) the trajectory calculated using the ORB-SLAM with EUCM model (EUCM ORB-SLAM). The camera trajectory (exposure stations) estimated by Metashape (used as the reference), equisolid ORB-SLAM, and EUCM ORB-SLAM are depicted in Figure 5. The trajectories are shown in black, red, and green, respectively. The camera positions are plotted on the UTM reference system, providing a visual representation of the trajectory for each solution along the 140 m path. The

trajectory was intentionally straight to simulate walking in a path where there is no expectation of return and loop closure based on revisiting a place, such as in forests, agricultural corridors, and other similar scenarios.

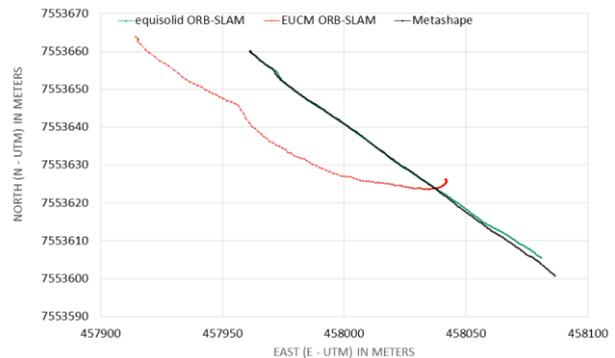


Figure 5. Trajectories equisolid ORB-SLAM, EUCM ORB-SLAM and Metashape.

The use of hyperhemispherical images from the Ricoh camera shows that the EUCM model exhibited significant drifts in position and scale when compared to both the reference trajectory and the equisolid ORB-SLAM trajectory. The equisolid ORB-SLAM presented more consistent results compared to the EUCM ORB-SLAM. However, it is possible to observe a drift at the end of the trajectory, which affected not only the position but also the scale. We selected the first and last ten estimated camera positions from both the equisolid ORB-SLAM and EUCM ORB-SLAM trajectories. Then, we calculated the RMSE (Root Mean Squared Error) with respect to the corresponding camera positions obtained from the reference trajectory (Metashape bundle adjustment). The results are presented in Table 1 and Table 2, which show the RMSE of the discrepancies in camera positions between the first and last ten frames of the reference trajectory (Metashape), and the trajectories of equisolid ORB-SLAM and EUCM ORB-SLAM, respectively. The equisolid ORB-SLAM had an RMSE of 0.114 m, 0.111 m, and 0.005 m for E, N, and h coordinates in the first ten frames, while for the last ten frames, the RMSE increased to 3.941 m, 3.592 m, and 0.113 m for E, N, and h coordinates, respectively, as indicated in Table 1. These results show that the discrepancies in the camera positions gradually increased as new observations were acquired and sequentially processed, leading to error propagation and drift at the end of the trajectory. Table 2 shows that the EUCM ORB-SLAM had a RMSE of 44.114 m, 2.402 m, and 17.580 m for E, N, and h coordinates, respectively, for the first ten frames in the estimated trajectory, and a RMSE of 40.859 m, 23.125 m, and 6.854 m for E, N, and h coordinates, respectively, for the last ten frames. The higher errors of the ORB-SLAM EUCM when compared to the equisolid ORB-SLAM performance can be explained by its generic nature, which was not able to cope with the hyperhemispheric geometry of the Ricoh Theta S, mainly due to the considerably long path and a lower frame rate (5 fps).

RMSE	E (m)	N (m)	h (m)
Beginning	0.114	0.111	0.005
End	3.941	3.592	0.113

Table 1. RMSE of the beginning and end of the trajectory for equisolid ORB-SLAM.

RMSE	E (m)	N (m)	h (m)
Beginning	44.114	2.402	17.580
End	40.859	23.125	6.854

**Table 2.** RMSE of the beginning and end of the trajectory for EUCM ORB-SLAM.

#### 4. CONCLUSION

The complex lens geometry involved in the image formation process still limits the full potential of fisheye lenses cameras, especially when points in the hyperhemispherical field are unmodeled either by a generic model or by some rigorous model (Castanheiro et al., 2021). Based on our results, it is possible to conclude that when using hyperhemispherical images, the EUCM model showed more discrepancies to the reference trajectory than the equisolid-angle model. One possible reason is the unfit of EUCM model to the Ricoh Theta S camera geometry. Furthermore, a longer path can also contribute to the trajectory drift since the trajectory resulting from equisolid-angle model was also affected by drift. It is important to emphasize that in the conducted experiments, the loop closure option of ORB-SLAM was not accomplished, which could significantly reduce the effect caused by the trajectory drift. Additionally, a higher frame rate may improve the performance of the EUCM model, but at the cost of increasing computational complexity. Overall, we can conclude that the use of a rigorous photogrammetric model, with a suitable treatment of hyperhemispherical points, proved to be beneficial, enabling the use of image observations throughout the full field of view for trajectory estimation.

#### ACKNOWLEDGEMENTS

This study was funded by the Brazilian Federal Agency for Support and Evaluation of Graduate Education (Scholarship, CAPES - Finance Code 001 and Program CAPES-PrInt, process number 88887.310313/2018-00, National Council for Scientific and Technological Development (CNPq - grant n. 303670/2018-5) and São Paulo Research Foundation (FAPESP) – grant 2021/06029-7.

#### REFERENCES

- Abraham, S., Förstner, W., 2005. Fish-eye-stereo calibration and epipolar rectification. *ISPRS J. Photogramm. Remote Sens.* 59, 278–288. <https://doi.org/10.1016/j.isprsjprs.2005.03.001>
- Aldegheri, S., Bombieri, N., Bloisi, D.D., Farinelli, A., 2019. Data Flow ORB-SLAM for Real-time Performance on Embedded GPU Boards, in: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). Presented at the 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), IEEE, Macau, China, pp. 5370–5375. <https://doi.org/10.1109/IROS40897.2019.8967814>
- Brown, D.C., 1971. Close-range camera calibration. *Photogramm. Eng.* 37, 855–866.
- Cadena, C., Carlone, L., Carrillo, H., Latif, Y., Scaramuzza, D., Neira, J., Reid, I., Leonard, J.J., 2016. Past, present, and future of simultaneous localization and mapping: Toward the robust-perception age. *IEEE Trans. Robot.* 32, 1309–1332.
- Campos, C., Elvira, R., Rodríguez, J., J. G. Montiel, J.M.M., Tardós, J.D., 2020. ORB-SLAM3: An Accurate Open-Source Library for Visual, Visual-Inertial and Multi-Map SLAM. <https://doi.org/10.48550/ARXIV.2007.11898>
- Campos, M., Tommaselli, A., Honkavaara, E., Prol, F., Kaartinen, H., El Issaoui, A., Hakala, T., 2018. A Backpack-Mounted Omnidirectional Camera with Off-the-Shelf Navigation Sensors for Mobile Terrestrial Mapping: Development and Forest Application. *Sensors* 18, 827. <https://doi.org/10.3390/s18030827>
- Campos, M.B., 2019. A Portable Mobile Terrestrial System With Omnidirectional Camera For Close Range Applications (PhD Thesis). Faculdade de Ciências e Tecnologia de Presidente Prudente, Presidente Prudente.
- Castanheiro, L.F., Tommaselli, A.M.G., Berveglieri, A., Campos, M.B., Junior, J.M., 2021. Modeling Hyperhemispherical Points and Calibrating a Dual-Fish-Eye System for Close-Range Applications. *Photogramm. Eng. Remote Sens.* 87, 375–384. <https://doi.org/10.14358/PERS.87.5.375>
- Conrady, A.E., 1919. Decentred Lens-Systems. *Mon. Not. R. Astron. Soc.* 79, 384–390. <https://doi.org/10.1093/mnras/79.5.384>
- Davison, A.J., Reid, I.D., Molton, N.D., Stasse, O., 2007. MonoSLAM: Real-Time Single Camera SLAM. *IEEE Trans. Pattern Anal. Mach. Intell.* 29, 1052–1067. <https://doi.org/10.1109/TPAMI.2007.1049>
- Durrant-Whyte, H., Bailey, T., 2006. Simultaneous localization and mapping: part I. *IEEE Robot. Autom. Mag.* 13, 99–110.
- Engel, J., Schöps, T., Cremers, D., 2014. LSD-SLAM: Large-Scale Direct Monocular SLAM, in: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (Eds.), *Computer Vision – ECCV 2014*, Lecture Notes in Computer Science. Springer International Publishing, Cham, pp. 834–849. [https://doi.org/10.1007/978-3-319-10605-2\\_54](https://doi.org/10.1007/978-3-319-10605-2_54)
- Forster, C., Pizzoli, M., Scaramuzza, D., 2014. SVO: Fast semi-direct monocular visual odometry, in: 2014 IEEE International Conference on Robotics and Automation (ICRA). Presented at the 2014 IEEE International Conference on Robotics and Automation (ICRA), IEEE, Hong Kong, China, pp. 15–22. <https://doi.org/10.1109/ICRA.2014.6906584>
- Geyer, C., Konstantinos, D., 2000. A Unifying Theory for Central Panoramic Systems and Practical Applications, in: *Proceedings of the 6th European Conference on Computer Vision-Part II, ECCV '00*. Presented at the 6th European Conference on Computer Vision-Part II, Springer-Verlag, Berlin, Heidelberg, pp. 445–461.
- Granshaw, S.I., 2018. Structure from motion: origins and originality. *Photogramm. Rec.* 33, 6–10. <https://doi.org/10.1111/phor.12237>
- Hughes, C., Denny, P., Jones, E., Glavin, M., 2010. Accuracy of fish-eye lens models. *Appl. Opt.* 49, 3338. <https://doi.org/10.1364/AO.49.003338>
- Kannala, J., Brandt, S.S., 2006. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Trans. Pattern Anal. Mach. Intell.* 28, 1335–1340. <https://doi.org/10.1109/TPAMI.2006.153>
- Klein, G., Murray, D., 2007. Parallel Tracking and Mapping for Small AR Workspaces, in: 2007 6th IEEE and ACM International Symposium on Mixed and Augmented Reality. Presented at the 2007 6th IEEE International Symposium on

Mixed and Augmented Reality (ISMAR), IEEE, Nara, Japan, pp. 1–10. <https://doi.org/10.1109/ISMAR.2007.4538852>

Kummerle, R., Grisetti, G., Strasdat, H., Konolige, K., Burgard, W., 2011. G2o: A general framework for graph optimization, in: 2011 IEEE International Conference on Robotics and Automation. Presented at the 2011 IEEE International Conference on Robotics and Automation (ICRA), IEEE, Shanghai, China, pp. 3607–3613. <https://doi.org/10.1109/ICRA.2011.5979949>

Li, A., Ruan, X., Huang, J., Zhu, X., Wang, F., 2019. Review of vision-based Simultaneous Localization and Mapping, in: 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC). Presented at the 2019 IEEE 3rd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC), IEEE, Chengdu, China, pp. 117–123. <https://doi.org/10.1109/ITNEC.2019.8729285>

Liu, S., Guo, P., Feng, L., Yang, A., 2019. Accurate and Robust Monocular SLAM with Omnidirectional Cameras. *Sensors* 19, 4494. <https://doi.org/10.3390/s19204494>

Matsuki, H., Von Stumberg, L., Usenko, V., Stuckler, J., Cremers, D., 2018. Omnidirectional DSO: Direct Sparse Odometry with Fisheye Cameras. <https://doi.org/10.48550/ARXIV.1808.02775>

Mur-Artal, R., Montiel, J.M.M., Tardos, J.D., 2015. ORB-SLAM: A versatile and accurate monocular SLAM system. *IEEE Trans. Robot.* 31, 1147–1163.

Newcombe, R.A., Lovegrove, S.J., Davison, A.J., 2011. DTAM: Dense tracking and mapping in real-time, in: 2011 International Conference on Computer Vision. Presented at the 2011 IEEE International Conference on Computer Vision (ICCV), IEEE, Barcelona, Spain, pp. 2320–2327. <https://doi.org/10.1109/ICCV.2011.6126513>

Ray, S.F., 1994. Applied photographic optics: lenses and optical systems for photography, film, video, and electronic imaging, 2 ed. ed. Focal Press, Oxford ; Boston.

Schneider, D., Schwalbe, E., Maas, H.-G., 2009. Validation of geometric models for fisheye lenses. *ISPRS J. Photogramm. Remote Sens.* 64, 259–266. <https://doi.org/10.1016/j.isprsjprs.2009.01.001>

Taketomi, T., Uchiyama, H., Ikeda, S., 2017. Visual SLAM algorithms: a survey from 2010 to 2016. *IPSJ Trans. Comput. Vis. Appl.* 9, 16. <https://doi.org/10.1186/s41074-017-0027-2>

Torresani, A., Remondino, F., 2019. Videogrammetry VS Photogrammetry for heritage 3D reconstruction. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* XLII-2/W15, 1157–1162. <https://doi.org/10.5194/isprs-archives-XLII-2-W15-1157-2019>

Usenko, V., Demmel, N., Cremers, D., 2018. The Double Sphere Camera Model. <https://doi.org/10.48550/ARXIV.1807.08957>

Wang, S., Yue, J., Dong, Y., Shen, R., Zhang, X., 2018. Real-time Omnidirectional Visual SLAM with Semi-Dense Mapping, in: 2018 IEEE Intelligent Vehicles Symposium (IV). Presented at the 2018 IEEE Intelligent Vehicles Symposium (IV), IEEE, Changshu, pp. 695–700. <https://doi.org/10.1109/IVS.2018.8500460>