

CONCLUSÃO

Foram apresentados um detector de bordas subpíxel, o vinculador de bordas e a aprendizagem de máquina especificamente ajustados para as ondas do tanque de ondas da COPPE/UFRJ, projetados para a posterior reconstrução 3D das ondas através de curvas. A abordagem utilizada pode ser estendida e generalizada para outras cenas, já que o aprendizado é feito sobre os dados de treinamento, restando apenas o ajuste fino na escolha das características finais e parâmetros dos métodos, específicos de cada aplicação.

São duas as contribuições mais significativas deste trabalho, o sistema funcional para o tanque de ondas, que pode ser expandido e modificado livremente junto com a ferramenta gráfica para a construção de referências para qualquer cena, e a análise do aprendizado aplicado a cada característica utilizada, servindo como guia para futuras expansões e construção de outros sistemas.

Partes deste trabalho foram apresentadas em [Fabbri \(2016\)](#), [Fabbri \(2017\)](#), [Souza e Fabbri \(2017\)](#).

Trabalhos futuros

Algumas direções para trabalhos futuros incluem:

- alimentação das curvas extraídas ao sistema de reconstrução 3D. Marcações manuais de correspondência para a avaliação do sistema de reconstrução 3D seriam necessárias. A interface gráfica para este trabalho está disponível no [The VXD... \(2017\)](#). O aprendizado aplicado em 3D também pode fornecer grandes avanços na reconstrução;
- avaliação e ajuste mais refinados do sistema a fim de melhorar ainda mais a extração de contornos eliminando *outliers*, e talvez introduzindo novas características para a quebra, junção e seleção de curvas;
- aplicação das mesmas técnicas em ambientes abertos como as ondas do mar;
- melhoria na escalabilidade do sistema. Embora o sistema por processos e arquivos seja eficaz e tenha uma escalabilidade boa, pode-se conseguir ainda mais reduzindo a granularidade ao nível de *threads*, tornando a comunicação mais direta entre processos por *pipes*, utilizando memória compartilhada ou outras formas mais sofisticadas;
- automação da escolha das áreas de interesse das quatro câmeras;
- ajuste automático de tamanho proporcional ao tamanho da imagem.

REFERÊNCIAS

13TH International Conference on Computer Vision, Barcelona, Spain.

ABUHASHIM, T.; SUKKARIEH, S. Incorporating geometric information into gaussian process terrain models from monocular images. In: *Intelligent Robots and Systems (IROS), 2012 IEEE/RSJ International Conference on*. [S.l.: s.n.], 2012. p. 4162–4168. ISSN 2153-0858.

AGARWAL, S. et al. Building rome in a day. In: *Twelfth IEEE International Conference on Computer Vision (ICCV 2009)*. Kyoto, Japan: IEEE, 2009. Disponível em: <https://www.microsoft.com/en-us/research/publication/building-rome-in-a-day/>. Acesso em: 31 jul. 2017.

AMORIM, M. de et al. An image-based fluid surface pattern model. *International Review of Chemical Engineering*, v. 5, n. 5, 2013.

ARBELAEZ, P. et al. Contour detection and hierarchical image segmentation. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 33, n. 5, p. 898–916, 2011.

BAATZ, G. et al. Large scale visual geo-localization of images in mountainous terrain. In: *Proceedings of the 12th European Conference on Computer Vision - Volume Part II*. Berlin, Heidelberg: Springer-Verlag, 2012. (ECCV'12), p. 517–530. ISBN 978-3-642-33708-6.

BASU, M. Gaussian-based edge-detection methods-a survey. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, IEEE, v. 32, n. 3, p. 252–260, 2002.

BENETAZZO, A. Measurements of short water waves using stereo matched image sequences. *Coastal engineering*, Elsevier, v. 53, n. 12, p. 1013–1032, 2006.

BENETAZZO, A. et al. Stereo wave imaging from moving vessels: Practical use and applications. *Coastal Engineering*, Elsevier, v. 109, p. 114–127, 2016.

BENETAZZO, A. et al. Offshore stereo measurements of gravity waves. *Coastal Engineering*, Elsevier, v. 64, p. 127–138, 2012.

BERGAMASCO, F. et al. Multi-view horizon-driven sea plane estimation for stereo wave imaging on moving vessels. *Computers & Geosciences*, Elsevier, v. 95, p. 105–117, 2016.

BERTHILSSON, R.; ÅSTRÖM, K.; HEYDEN, A. Reconstruction of general curves, using factorization and bundle adjustment. *International Journal of Computer Vision*, v. 41, n. 3, p. 171–182, 2001. Disponível em: <http://dx.doi.org/10.1023/A:1011104020586>. Acesso em: 31 jul. 2017.

BOUCHARA, J. C.; ZANETIC, V. L. C.; HELLMEISTER, A. C. P. *Cálculo integral avançado*. [S.l.]: Edusp, 1996.

CALAKLI, F. et al. High resolution surface reconstruction from multi-view aerial imagery. In: *IEEE. 3DIMPVT'12*. [S.l.], 2012. p. 25–32.

- CARLSON, F. B.; VUONG, N. D.; JOHANSSON, R. Polynomial reconstruction of 3D sampled curves using auxiliary surface data. In: IEEE. *Robotics and Automation (ICRA), 2014 IEEE International Conference on*. [S.l.], 2014. p. 4411–4416.
- CHEN, T.-Y.; KLETTE, R. Animated non-photorealistic rendering in multiple styles. In: SPRINGER. *Image and Video Technology–PSIVT 2013 Workshops*. [S.l.], 2014. p. 12–23.
- CIPOLLA, R.; GIBLIN, P. *Visual Motion of Curves and Surfaces*. [S.l.]: Cambridge University Press, 2000.
- COLE, F. et al. How well do line drawings depict shape? *ACM Trans. Graph.*, ACM, New York, NY, USA, v. 28, n. 3, p. 28:1–28:9, jul. 2009. ISSN 0730-0301. Disponível em: <http://doi.acm.org/10.1145/1531326.1531334>. Acesso em: 31 jul. 2017.
- DISKIN, Y.; ASARI, V. Dense point-cloud representation of a scene using monocular vision. *Journal of Electronic Imaging*, v. 24, n. 2, p. 023003, 2015. Disponível em: <http://dx.doi.org/10.1117/1.JEI.24.2.023003>. Acesso em: 31 jul. 2017.
- DOLLÁR, P.; ZITNICK, C. L. Fast edge detection using structured forests. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 37, n. 8, p. 1558–1570, 2015.
- ESPINO, J. R. et al. Vision system for 3D reconstruction with telecentric lens. In: *Pattern Recognition*. [S.l.]: Springer, 2012, (Lecture Notes in Computer Science, v. 7329). p. 127–136. ISBN 978-3-642-31148-2.
- ESTEBAN, C. H.; SCHMITT, F. Silhouette and stereo fusion for 3d object modeling. *Comput. Vis. Image Underst.*, Elsevier Science Inc., New York, NY, USA, v. 96, n. 3, p. 367–392, dez. 2004. ISSN 1077-3142. Disponível em: <http://dx.doi.org/10.1016/j.cviu.2004.03.016>. Acesso em: 31 jul. 2017.
- FABBRI, R. *Multiview Differential Geometry in Application to Computer Vision*. Tese (Ph.D. Dissertation) — Division Of Engineering, Brown University, Providence, RI, 02912, July 2010.
- FABBRI, R. Image-based fluid modeling. In: *First International Brazilian Symposium on Water Waves (invited talk)*. Rio de Janeiro, Brazil: [s.n.], 2016.
- FABBRI, R. Current techniques for the 3d photogrammetry of water waves. In: *Second International Brazilian Symposium on Water Waves (invited talk)*. Rio de Janeiro, Brazil: [s.n.], 2017.
- FABBRI, R.; GIBLIN, P. J.; KIMIA, B. B. Camera pose estimation using first-order curve differential geometry. In: *Proceedings of the IEEE European Conference in Computer Vision*. [S.l.]: Springer, 2012. (Lecture Notes in Computer Science).
- FABBRI, R.; KIMIA, B. B. High-order differential geometry of curves for multiview reconstruction and matching. In: SPRINGER. *Energy Minimization Methods in Computer Vision and Pattern Recognition*. [S.l.], 2005. p. 645–660.
- FABBRI, R.; KIMIA, B. B. 3D curve sketch: Flexible curve-based stereo reconstruction and calibration. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA: IEEE Computer Society Press, 2010.

FABBRI, R.; KIMIA, B. B. Multiview differential geometry of curves. *International Journal of Computer Vision*, Springer, v. 117, p. 1–23, 2016. Disponível em: <http://dx.doi.org/10.1007/s11263-016-0912-7>. Acesso em: 31 jul. 2017.

FATHI, H.; DAI, F.; LOURAKIS, M. Automated as-built 3D reconstruction of civil infrastructure using computer vision: Achievements, opportunities, and challenges. *Advanced Engineering Informatics*, 2015. ISSN 1474-0346. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1474034615000245>. Acesso em: 31 jul. 2017.

FEDELE, F. et al. Space-time measurements of oceanic sea states. *Ocean Modelling*, Elsevier, v. 70, p. 103–115, 2013.

FEDELE, F. et al. Euler characteristics of oceanic sea states. *Mathematics and Computers in Simulation*, Elsevier, v. 82, n. 6, p. 1102–1111, 2012.

FELZENSZWALB, P.; MCALLESTER, D. A min-cover approach for finding salient curves. In: *Computer Vision and Pattern Recognition Workshop, 2006. CVPRW '06. Conference on*. [S.l.: s.n.], 2006. p. 185.

FENG, A. et al. Rapid avatar capture and simulation using commodity depth sensors. In: *ACM SIGGRAPH 2014 Talks*. New York, NY, USA: ACM, 2014. (SIGGRAPH '14), p. 16:1–16:1. ISBN 978-1-4503-2960-6. Disponível em: <http://doi.acm.org/10.1145/2614106.2614182>. Acesso em: 31 jul. 2017.

FERRARI, V. et al. Groups of adjacent contour segments for object detection. *IEEE Trans. Pattern Analysis and Machine Intelligence*, v. 30, n. 1, p. 36–51, 2008.

FFMPEG. 2017. Disponível em: <https://ffmpeg.org/>. Acesso em: 31 jul. 2017.

FISCHLER, M.; TENENBAUM, J.; WOLF, H. Detection of roads and linear structures in low-resolution aerial imagery using a multisource knowledge integration technique. *Computer Graphics and Image Processing*, v. 15, n. 3, p. 201 – 223, 1981.

FLEMING, R. W.; HOLTMANN-RICE, D.; BÜLTHOFF, H. H. Estimation of 3D shape from image orientations. *Proceedings of the National Academy of Sciences*, National Acad Sciences, v. 108, n. 51, p. 20438–20443, 2011.

FURUKAWA, Y.; PONCE, J. Accurate, dense, and robust multi-view stereopsis. In: *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), 18-23 June 2007, Minneapolis, Minnesota, USA*. [s.n.], 2007. Disponível em: <http://dx.doi.org/10.1109/CVPR.2007.383246>. Acesso em: 31 jul. 2017.

GALLEGO, G. et al. A variational stereo method for the three-dimensional reconstruction of ocean waves. *IEEE Transactions on Geoscience and Remote Sensing*, IEEE, v. 49, n. 11, p. 4445–4457, 2011.

GALLEGO, G. et al. Variational stereo imaging of oceanic waves with statistical constraints. *IEEE Transactions on Image Processing*, IEEE, v. 22, n. 11, p. 4211–4223, 2013.

- GOESELE, M. et al. Multi-view stereo for community photo collections. In: *IEEE 11th International Conference on Computer Vision, ICCV 2007, Rio de Janeiro, Brazil, October 14-20, 2007*. [s.n.], 2007. p. 1–8. Disponível em: <http://dx.doi.org/10.1109/ICCV.2007.4408933>. Acesso em: 31 jul. 2017.
- GUO, Y.; KIMIA, B. B. On evaluating methods for recovering image curve fragments. In: *Proceedings of IEEE Workshop on Perceptual Organization in Computer Vision, POCV*. [S.l.: s.n.], 2012. p. 9–16.
- GUO, Y. et al. A multi-stage approach to curve extraction. In: *CVPR'14*. [S.l.: s.n.], 2014.
- GUO, Y. et al. A multi-stage approach to curve extraction. In: *CVPR'14*. [S.l.: s.n.], 2014.
- GUO, Y. et al. A multi-stage approach to curve extraction. In: *Proceedings of European Conference on Computer Vision*. [S.l.: Springer, 2014. (Lecture Notes in Computer Science, v. 8695), p. 663–678.
- GUO, Y.; NARAYANAN, M.; KIMIA, B. A multi-stage approach to curve extraction. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, IEEE, 2017. Submitted.
- HABBECKE, M.; KOBELT, L. A surface-growing approach to multi-view stereo reconstruction. In: *2007 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2007), 18-23 June 2007, Minneapolis, Minnesota, USA*. [s.n.], 2007. Disponível em: <http://dx.doi.org/10.1109/CVPR.2007.383195>. Acesso em: 31 jul. 2017.
- HARTLEY, R.; ZISSERMAN, A. *Multiple View Geometry in Computer Vision*. [S.l.: Cambridge University Press, 2000.
- HEINLY, J. et al. Reconstructing the world in six days. In: *Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2015.
- HU, W.; WU, Y. N.; ZHU, S.-C. Image representation by active curves. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 1808–1815.
- IMAGEMAGICK. 2017. Disponível em: www.imagemagick.org. Acesso em: 31 jul. 2017.
- IVERSON, L.; ZUCKER, S. Logical/linear operators for image curves. *PAMI*, v. 17, n. 10, p. 982–996, October 1995.
- KENNEDY, R.; GALLIER, J.; SHI, J. Contour cut: Identifying salient contours in images by solving a hermitian eigenvalue problem. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.: IEEE, 2011. p. 2065–2072.
- KOENDERINK, J.; DOORN, A. van; WAGEMANS, J. SfS? not likely. . . . *i-Perception*, Pion Publications, v. 4, n. 5, p. 299, 2013.

- KOKKINOS, I. Highly accurate boundary detection and grouping. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. San Francisco, CA, USA: IEEE Computer Society Press, 2009. p. 2520–2527.
- KOWDLE, A. et al. imodel: Interactive co-segmentation for object of interest 3D modeling. In: *Proceedings of the 11th European Conference on Trends and Topics in Computer Vision - Volume Part II*. [S.l.]: Springer-Verlag, 2012. (ECCV'10), p. 211–224. ISBN 978-3-642-35739-8.
- KUANG, Y.; OSKARSSON, M.; ÅSTRÖM, K. Revisiting trifocal tensor estimation using lines. In: IEEE. *Pattern Recognition (ICPR), 2014 22nd International Conference on*. [S.l.], 2014. p. 2419–2423.
- KUNSBURG, B.; ZUCKER, S. W. Why shading matters along contours. In: *Neuromathematics of Vision*. [S.l.]: Springer, 2014, (Lecture Notes in Morphogenesis). p. 107–129. ISBN 978-3-642-34443-5.
- LEBEDA, K.; HADFIELD, S.; BOWDEN, R. 2D or not 2D: Bridging the gap between tracking and structure from motion. In: *Proc. of ACCV*. [S.l.: s.n.], 2014.
- LEGGITT, J. *Drawing Shortcuts: Developing Quick Drawing Skills Using Today's Technology*. Wiley, 2015. ISBN 9780470639139. Disponível em: <https://books.google.com/books?id=ZrDzCQAAQBAJ>. Acesso em: 31 jul. 2017.
- LIN, L. et al. Layered graph matching by composite cluster sampling with collaborative and competitive interactions. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Miami, Florida, USA: IEEE Computer Society Press, 2009. p. 1351–1358.
- LITVINOV, V.; YU, S.; LHUILLIER, M. 2-manifold reconstruction from sparse visual features. In: IEEE. *3D Imaging (IC3D), 2012 International Conference on*. [S.l.], 2012. p. 1–8.
- LIU, S. et al. Free-form object reconstruction from silhouettes, occluding edges and texture edges: A unified and robust operator based on duality. *IEEE Trans. Pattern Anal. Mach. Intell.*, v. 30, n. 1, p. 131–146, 2008. Disponível em: <http://dx.doi.org/10.1109/TPAMI.2007.1143>. Acesso em: 31 jul. 2017.
- MARTIN, D. R.; FOWLKES, C. C.; MALIK, J. Learning to detect natural image boundaries using local brightness, color, and texture cues. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, IEEE Computer Society, Los Alamitos, CA, USA, v. 26, n. 5, p. 530–549, 2004. ISSN 0162-8828.
- MATTINGLY, W. A. et al. 3D modeling of branching structures for anatomical instruction. *Journal of Visual Languages & Computing*, v. 29, n. 0, p. 54 – 62, 2015. ISSN 1045-926X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1045926X15000191>. Acesso em: 31 jul. 2017.
- MOREELS, P.; PERONA, P. Evaluation of features detectors and descriptors based on 3d objects. *Int. J. Comput. Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 73, n. 3, p. 263–284, jul. 2007. ISSN 0920-5691. Disponível em: <http://dx.doi.org/10.1007/s11263-006-9967-1>. Acesso em: 31 jul. 2017.

- NARAYANAN, M.; KIMIA, B. To complete or not to complete: Gap completion in real images. In: *Proceedings of IEEE Workshop on Perceptual Organization in Computer Vision, POCV*. [S.l.: s.n.], 2012. p. 47–54.
- NURUTDINOVA, I.; FITZGIBBON, A. Towards pointless structure from motion: 3d reconstruction and camera parameters from general 3d curves. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2015. p. 2363–2371.
- PAYET, N.; TODOROVIC, S. From a set of shapes to object discovery. In: *Proceedings of European Conference on Computer Vision*. [S.l.: Springer, 2010. (Lecture Notes in Computer Science), p. 57–70.
- PAYET, N.; TODOROVIC, S. From contours to 3d object detection and pose estimation. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.: s.n.], 2011. p. 983–990.
- POLLEFEYS, M. et al. Visual modeling with a hand-held camera. *Int. J. Comput. Vision*, Kluwer Academic Publishers, Hingham, MA, USA, v. 59, n. 3, p. 207–232, set. 2004. ISSN 0920-5691. Disponível em: <http://dx.doi.org/10.1023/B:VISI.0000025798.50602.3a>. Acesso em: 31 jul. 2017.
- PöTSCH, K.; PINZ, A. 3D geometric shape modeling by ‘3D contour cloud’ reconstruction from stereo videos. In: CITESEER. *16th Computer Vision Winter Workshop*. [S.l.], 2011. p. 99.
- CVPR’10: IEEE Conference on Computer Vision and Pattern Recognition, San Francisco, California*. San Francisco, CA, USA: IEEE Computer Society Press, 2009.
- Second International Brazilian Symposium on Water Waves*.
- RAO, D.; CHUNG, S.-J.; HUTCHINSON, S. CurveSLAM: An approach for vision-based navigation without point features. In: *IEEE/RSJ Intelligent Robots and Systems (IROS)*. [S.l.: s.n.], 2012. p. 4198–4204. ISSN 2153-0858.
- RAPIZO, H. et al. Simple techniques for retrieval of wind wave periods and directions from optical images sequences in wave tanks. *Journal of Coastal Research*, The Coastal Education and Research Foundation, v. 32, n. 5, p. 1227–1234, 2015.
- REN, X.; FOWLKES, C.; MALIK, J. Learning probabilistic models for contour completion in natural images. *International Journal of Computer Vision*, v. 77, n. 1-3, p. 47–63, 2008.
- RESTREPO, M. I.; ULUSOY, A. O.; MUNDY, J. L. Evaluation of feature-based 3-D registration of probabilistic volumetric scenes. *ISPRS Journal of Photogrammetry and Remote Sensing*, Elsevier, v. 98, p. 1–18, 2014.
- SEITZ, S. M. et al. A comparison and evaluation of multi-view stereo reconstruction algorithms. In: *Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 1*. Washington, DC, USA: IEEE Computer Society, 2006. (CVPR ’06), p. 519–528. ISBN 0-7695-2597-0. Disponível em: <http://dx.doi.org/10.1109/CVPR.2006.19>. Acesso em: 31 jul. 2017.

- SHI, J.; MALIK, J. Normalized cuts and image segmentation. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. [S.l.]: IEEE Computer Society, 1997. p. 731–737.
- SHINOZUKA, Y.; SAITO, H. Sharing 3D object with multiple clients via networks using vision-based 3D object tracking. In: *Proceedings of the 2014 Virtual Reality International Conference*. New York, NY, USA: ACM, 2014. (VRIC '14), p. 34:1–34:4. ISBN 978-1-4503-2626-1. Disponível em: <http://doi.acm.org/10.1145/2617841.2620723>. Acesso em: 31 jul. 2017.
- SIMOES, F. et al. Challenges in 3D reconstruction from images for difficult large-scale objects. *2014 XVI Symposium on Virtual and Augmented Reality*, IEEE, Los Alamitos, CA, USA, v. 0, p. 74–83, 2012.
- SOUZA, C.; FABBRI, R. Low-level machine learning for tracking curves in water wave videos. In: *International Brazilian Symposium on Water Waves (poster)*. Rio de Janeiro, Brazil: [s.n.], 2017.
- TAMRAKAR, A. *Image Contour Extraction Using Geometric Consistency*. Providence, RI, 02912, 2008.
- TAMRAKAR, A.; KIMIA, B. B. No grouping left behind: From edges to curve fragments. In: *IEEE 11th International Conference on Computer Vision, ICCV 2007, Rio de Janeiro, Brazil, October 14-20, 2007*. [s.n.], 2007. p. 1–8. Disponível em: <http://dx.doi.org/10.1109/ICCV.2007.4408919>. Acesso em: 31 jul. 2017.
- TANGE, O. et al. Gnu parallel-the command-line power tool. *The USENIX Magazine*, v. 36, n. 1, p. 42–47, 2011.
- TAUBIN, G. et al. Shape from depth discontinuities. In: *Emerging Trends in Visual Computing, LIX Fall Colloquium, ETVC 2008, Palaiseau, France, November 18-20, 2008. Revised Invited Papers*. [S.l.: s.n.], 2008. p. 216–237.
- TENEY, D.; PIATER, J. Sampling-based multiview reconstruction without correspondences for 3D edges. In: *Proceedings*. Washington, DC, USA: IEEE, 2012. (3DIMPVT '12), p. 160–167. ISBN 978-0-7695-4873-9. Disponível em: <http://dx.doi.org/10.1109/3DIMPVT.2012.28>. Acesso em: 31 jul. 2017.
- THE VXD Development C++ Computer Vision Libraries. 2017. Disponível em: <http://github.com/rfabbri/vxd>. Acesso em: 31 jul. 2017.
- THE VXL C++ Computer Vision Libraries. 2017. Disponível em: <http://vxl.sourceforge.net>. Acesso em: 31 jul. 2017.
- USUMEZBAS, A.; FABBRI, R.; KIMIA, B. B. From multiview image curves to 3D drawings. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. [S.l.: s.n.], 2016. Submitted.
- USUMEZBAS, A.; FABBRI, R.; KIMIA, B. B. From multiview image curves to 3D drawings. In: *Proceedings of the European Conference in Computer Vision*. [S.l.: s.n.], 2016.

- USUMEZBAS, A.; FABBRI, R.; KIMIA, B. B. From multiview image curves to 3d drawings. In: *Computer Vision - ECCV 2016 - 14th European Conference, Amsterdam, The Netherlands, October 11-14, 2016, Proceedings, Part IV*. [S.l.: s.n.], 2016. p. 70–87.
- USUMEZBAS, A.; FABBRI, R.; KIMIA, B. B. The surfacing of multiview 3D drawings via lofting and occlusion reasoning. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, Hawaii, USA: IEEE Computer Society Press, 2017.
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. Kauai, Hawaii, USA: IEEE Computer Society Press, 2001. p. 511–518.
- WANG, R.; CHOI, J.; MEDIONI, G. 3D modeling from wide baseline range scans using contour coherence. In: IEEE. *Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on*. [S.l.], 2014. p. 4018–4025.
- YEE, R. *Architectural Drawing: A Visual Compendium of Types and Methods*. Wiley, 2012. ISBN 9781118310403. Disponível em: <<https://books.google.com/books?id=UwWai1uM-HsC>>. Acesso em: 31 jul. 2017.
- YU, S. X.; SHI, J. Multiclass spectral clustering. In: *Proceedings of the IEEE International Conference on Computer Vision*. Nice, France: IEEE Computer Society, 2003. p. 313–319. ISBN 0-7695-1950-4.
- ZHANG, L. *Line Primitives and Their Applications in Geometric Computer Vision*. Tese (Doutorado) — Department of Computer Science, Univ., 2013.
- ZHU, Q.; SONG, G.; SHI, J. Untangling cycles for contour grouping. In: *Proceedings of the IEEE International Conference on Computer Vision*. [S.l.]: IEEE Computer Society, 2007. p. 1–8.
- ZHU, Q. et al. Contour context selection for object detection: A set-to-set contour matching approach. In: *Proceedings of European Conference on Computer Vision*. [S.l.]: Springer, 2008. (Lecture Notes in Computer Science, v. 5303), p. 774–787. ISBN 978-3-540-88685-3.
- ZIA, M.; STARK, M.; SCHINDLER, K. Towards scene understanding with detailed 3D object representations. *International Journal of Computer Vision*, Springer US, v. 112, n. 2, p. 188–203, 2015. ISSN 0920-5691. Disponível em: <<http://dx.doi.org/10.1007/s11263-014-0780-y>>. Acesso em: 31 jul. 2017.
- ZUCKER, S. Stereo, shading, and surfaces: Curvature constraints couple neural computations. *Proceedings of the IEEE*, v. 102, n. 5, p. 812–829, May 2014.

APÊNDICE A – Geometria e Calibração do Sistema de Câmeras

Neste apêndice é apresentado o processo de calibração interna e externa do sistema de câmeras utilizado no experimento com o tanque *LabOceano* realizado neste trabalho. A intenção é facilitar o uso destas informações juntamente com os resultados do rastreador como entrada para sistemas de reconstrução 3D. A maior parte deste material é uma descrição original de técnicas clássicas de álgebra linear e geometria de múltiplas perspectivas (HARTLEY; ZISSERMAN, 2000), detalhando-se os conceitos e usando-se uma notação explícita.

A.1 Notação

Um *sistema de coordenadas* para um espaço vetorial V é dado por uma base de vetores e uma origem. Considere-se uma base $A = \{a_1, a_2, \dots, a_n\}$ de V . Ela forma um sistema de coordenadas quando se especifica uma origem O_A . Note-se que tanto a_i , $i = 1, \dots, n$ como O_A são elementos de V .

O vetor numérico em $\mathbb{R}^n$ das coordenadas de um vetor $v \in V$ relativas à base A é aqui explicitamente denotado $\mathcal{X}_A(v)$. Trata-se do conjunto dos únicos escalares $x_1, \dots, x_n$ tal que $v = x_1 a_1 + \dots + x_n a_n$. Lembre-se que um espaço vetorial consiste dos chamados vetores, junto com um corpo de escalares K , aqui tomado como os reais $\mathbb{R}$. Assim, as coordenadas são dadas por um isomorfismo entre os vetores de um espaço V de dimensão n , e o espaço de n -tuplas ordenadas de escalares (coordenadas), $\mathbb{R}^n$. Se nenhuma mudança de coordenadas é realizada por determinada aplicação, pode-se abusar a notação e usar o próprio vetor para denotar suas coordenadas, ou seja, escreve-se v como um atalho para $\mathcal{X}_A(v)$.

Se indicarmos por 1 o sistema de coordenadas $\{A, O_A\}$, então as coordenadas de um vetor $v \in V$ no sistema de coordenadas 1 é dado por

$$\mathcal{X}_1(v) := \mathcal{X}_A(v - O_A). \quad (15)$$

Definição A.1. Dado um mapa linear $L : V \rightarrow V$ e duas bases $A = \{a_1, a_2, \dots, a_n\}$ e $B = \{b_1, b_2, \dots, b_n\}$, a matriz de L relativa às bases A e B é a única matriz $\mathcal{M}_B^A(L)$ tal que:

$$\mathcal{X}_B(L(v)) = \mathcal{M}_B^A(L) \cdot \mathcal{X}_A(v), \quad (16)$$

para todo v em V .

Teorema A.1. A matriz $\mathcal{M}_B^A(L)$ é dada por:

$$\mathcal{M}_B^A(L) = \begin{bmatrix} M_B^A(L)\mathcal{X}_A(a_1) & M_B^A(L)\mathcal{X}_A(a_2) & \cdots & M_B^A(L)\mathcal{X}_A(a_n) \end{bmatrix} \quad (17)$$

Isto é, primeiro escreve-se cada vetor base transformado a_i de A em termos da base B , e colocam-se os mesmos na coluna da matriz.

Demonstração. Como $\mathcal{X}_A(a_i)$ é dado por $(0, \dots, 1, \dots, 0)^\top$, onde 1 ocorre na i -ésima entrada, aplicando-se a matriz o resultado é sua i -ésima coluna. $\square$

Definição A.2. (*Matriz de mudança de base*). Considere-se duas bases A e B para um espaço vetorial V . A matriz de mudança de base da base A à base B é a matriz do mapa identidade id relativa às bases A e B , ou seja, $\mathcal{M}_B^A(id)$. Trata-se da única matriz que leva os vetores coordenada $\mathcal{X}_A(v)$ aos vetores coordenada $\mathcal{X}_B(v)$ via multiplicação usual $\mathcal{X}_B(v) = \mathcal{M}_B^A(id)\mathcal{X}_A(v)$.

A.1.1 Projeção Perspectiva

Uma câmera é numericamente modelada como um dispositivo que mapeia um ponto em coordenadas globais – denominadas coordenadas do “mundo” – a um ponto projetado em coordenadas de uma imagem. No caso de um modelo *pinhole* ou estenopeico, isso é dado pela composição de três mudanças de coordenadas:

1. 3D–3D: transformação rígida (coordenadas do mundo às da câmera)
2. 3D–2D: projeção (coordenadas da câmera às coordenadas normalizadas de imagem do ponto projetado)
3. 2D–2D: transformação (coordenadas de imagem normalizadas às coordenadas finais da imagem)

Classicamente, em visão computacional a transformação de coordenadas homogêneas relativas ao mundo, a coordenadas homogêneas de imagem, é descrita por uma matriz de projeção 3×4 P , que pode ser decomposta como

$$P = K \begin{bmatrix} I & | & 0 \end{bmatrix} \begin{bmatrix} R & T \\ 0^\top & 1 \end{bmatrix} = K[R | T], \quad (18)$$

onde

- a matriz $\begin{bmatrix} R & T \\ 0^\top & 1 \end{bmatrix}$ (contendo os *parâmetros extrínsecos*) transforma das coordenadas do mundo às chamadas coordenadas 3D relativas à câmera;

- a matriz $[I | 0]$ realiza a projeção perspectiva, tendo como entrada coordenadas de câmera a coordenadas normalizadas 2D de imagem;
- a matriz K (*matriz de calibração*), contendo os *parâmetros intrínsecos* transforma de coordenadas normalizadas de imagem às coordenadas finais na imagem.

A seguir, tais transformações são detalhadas. As coordenadas de mundo são dadas por qualquer sistema de coordenadas 3D, denotado por $W_c = \{O, W\}$, com $W = \{e_1, e_2, e_3\}$. Um ponto neste sistema tem coordenadas indicadas por $(x^w, y^w, z^w)^\top$. A Figura 65 mostra este sistema, posicionado pela equipe do *LabOceano* no canto do tanque correspondente ao canto inferior esquerdo da figura; na Figura 25, este sistema está indicado por eixos $(x, y, z)^\top$ em vez de e_1, e_2 e e_3 , por simplicidade.

O sistema de coordenadas da imagem é definido como as coordenadas de píxel (u, v) com origem no canto superior esquerdo da imagem, onde u é o eixo horizontal, e v é o eixo vertical apontando para baixo, de forma a refletir indexação da matriz de imagens iniciando em $(0, 0)$. A cada câmera é associado um sistema ortonormal de coordenadas $\{c, B\}$, $B = \{b_1, b_2, b_3\}$, com as coordenadas indicadas por $(x, y, z)^\top$. O sistema da câmera é convenientemente definido de acordo com as seguintes propriedades:

- sua origem está no centro da câmera, c ;
- o plano $x - y$ é tomado como paralelo ao plano da imagem;
- o eixo z é normal ao plano da imagem;
- o sistema é orientado de acordo com a regra da mão direita (assumindo-se o mesmo para o sistema do mundo);
- escolhe-se z formando um sistema orientado pela regra da mão direita com as direções de u e v ;
- o eixo x é alinhado com o eixo u da imagem, por convenção;
- o eixo y é definido tal que x, y, z é ortogonal e orientado pela regra da mão direita;
- o sistema da câmera tem as mesmas unidades que as coordenadas do mundo (no caso do *LabOceano*, milímetros).

Define-se o *ponto principal* geometricamente como a projeção do centro da câmera c no plano da imagem, podendo ser expresso em qualquer sistema de coordenadas desejado. O plano da imagem, no sistema de coordenadas da câmera, é dado por pontos da forma (x, y, f) , onde f é a distância de c ao plano da imagem, chamado de *distância focal*. O plano do detector, em realidade, está espelhado em $(x, y, -f)$. Denota-se *coordenadas normalizadas de imagem* as coordenadas (x, y) no plano virtual $(x, y, 1)$ do sistema de

coordenadas da câmera. O plano de imagem normalizado é o plano $x - y$ transladado a uma distância unitária a partir de c . É uma imagem virtual tendo uma base ortonormal, origem no ponto principal (e não no canto da imagem), e tal como se a distância focal fosse igual a uma unidade do mundo.

De modo geral, para se definir os parâmetros de uma câmera em uma modelagem numérica, tem-se que:

1. identificar um sistema de coordenadas de imagem e um sistema de coordenadas do mundo;
2. definir um sistema de coordenadas de câmera de forma conveniente, como supracitado;
3. encontrar uma matriz de rotação e um vetor translação de forma a mudar as coordenadas do mundo para coordenadas relativas à câmera;
4. projetar os pontos em coordenadas de câmera para pontos em coordenadas de plano de imagem normalizado;
5. aplicar uma escala, translação e cisalhamento nas coordenadas normalizadas para obter as coordenadas finais.

Para obter as coordenadas finais, nesse último passo, tem-se que determinar:

- a distância focal f ;
- os fatores de conversão s_u e s_v entre unidades do mundo e de pixels, medidos nas direções u e v , respectivamente;
- a posição do ponto principal relativa ao canto superior esquerdo da imagem;
- o ângulo de cisalhamento (*skew*) entre u e v , já que em alguns casos (raros) u e v podem não ser ortogonais.

Pode-se mostrar que, a princípio, em uma modelagem por álgebra linear em coordenadas homogêneas, são necessários 6 parâmetros intrínsecos à câmera. Abaixo será mostrado que f , s_u e s_v podem ser combinados sem alterar o efeito final. Portanto, são necessários 5 parâmetros intrínsecos e 6 extrínsecos (rotação e translação), dando um total de 11 parâmetros para um modelo de uma câmera *pinhole*. Define-se o *aspect ratio* como s_u/s_v .

Na notação por coordenadas homogêneas em (18), os parâmetros intrínsecos estão codificados na chamada *matriz de calibração* (interna) K ,

$$K = \begin{bmatrix} \alpha_u & s & u_o \\ 0 & \alpha_v & v_o \\ 0 & 0 & 1 \end{bmatrix}, \quad (19)$$

onde

α_u	distância focal $\times u$ píxeis / unidade do mundo (mm)
α_v	distância focal $\times y$ píxeis / unidade do mundo (mm)
u_0	coordenada u do ponto principal
v_0	coordenada v do ponto principal
s_θ	fator de cisalhamento ou <i>skew</i>

$$\alpha_u = \frac{f}{\text{largura de um píxel em unidades do mundo}} \quad (20)$$

$$\alpha_v = \frac{f}{\text{altura de um píxel em unidades do mundo}}. \quad (21)$$

Problema A.1. *Dado um ponto $p^w = (x^w, y^w, z^w)^\top$ em coordenadas de mundo, qual é (u, v) em píxeis?*

Esta operação é mais facilmente realizada usando-se coordenadas projetivas e multiplicação de matrizes (HARTLEY; ZISSERMAN, 2000). No entanto, abaixo são detalhadas passo-a-passo cada transformação envolvida. Primeiro, transforma-se de coordenadas de mundo a coordenadas de câmera $p = (x, y, z)^\top$:

$$p = Rp^w + T, \quad (22)$$

onde R e T são os parâmetros extrínsecos. Em seguida, projeta-se o ponto ao plano de imagem normalizado usando-se semelhança de triângulos,

$$\begin{cases} \tilde{x} = \frac{x}{z} \\ \tilde{y} = \frac{y}{z}. \end{cases} \quad (23)$$

$$\quad (24)$$

Após a projeção, obtêm-se as coordenadas de imagens, primeiramente incluindo-se a distância focal:

$$\begin{cases} \bar{x} = f \frac{x}{z} \\ \bar{y} = f \frac{y}{z}. \end{cases} \quad (25)$$

$$\quad (26)$$

Em seguida, converte-se a píxeis

$$\begin{cases} \bar{x}_{pix} = s_u f \frac{x}{z} \\ \bar{y}_{pix} = s_v f \frac{y}{z}, \end{cases} \quad (27)$$

$$\quad (28)$$

Tabela 4 - Resumo dos parâmetros de uma câmera em visão computacional

Graus de Liberdade	Descrição
3	rotação
3	translação
2	escala ao longo de x, y
1	distância focal (pode ser incorporado como fator de escala)
2	ponto principal
1	<i>skew</i> ou cisalhamento
6	Total extrínsecos
6	Total intrínsecos
12	Total
11	Total únicos (distância focal como fator de escala)

e translada-se ao canto superior esquerdo, somando-se um vetor translação $(t_{\tilde{u}}, t_{\tilde{v}})$,

$$\begin{cases} \tilde{u} = s_u f \frac{x}{z} + t_{\tilde{u}} \\ \tilde{v} = s_v f \frac{y}{z} + t_{\tilde{v}}. \end{cases} \quad (29)$$

Porfim, converte-se para coordenadas não-ortogonais com *skew*,

$$\begin{cases} u = \tilde{u} - \tilde{v} \cot \theta \\ v = \tilde{v} \csc \theta, \end{cases} \quad (31)$$

dando

$$\begin{cases} u = \overbrace{s_u f}^{\alpha_u} \frac{x}{z} + \overbrace{(-\cot \theta s_v f)}^{s_\theta} \frac{y}{z} + \overbrace{t_{\tilde{u}} - \cot \theta t_{\tilde{v}}}^{u_0} \\ v = \underbrace{s_v f \csc \theta}_{\alpha_v} \frac{y}{z} + \underbrace{\csc \theta t_{\tilde{v}}}_{v_0}. \end{cases} \quad (33)$$

$$\quad (34)$$

A.2 Geometria Epipolar

A geometria epipolar é a representação da pose relativa e da diferença entre parâmetros intrínsecos entre câmeras, em termos de seu efeito em transformações entre as imagens (fotografias) 2D dessas câmeras. Essa geometria surge ao responder a seguinte questão: dados dois pontos projetados p^1 e p^2 em duas vistas, como saber se eles podem corresponder a um mesmo ponto 3D? Primeiramente, definem-se alguns conceitos básicos

Definição A.3. (*Entidades básicas da geometria epipolar*)

- A *reta epipolar* el_1^2 de um ponto p^1 , na imagem 2, ou $el_1^2(p^1)$, é a projeção na imagem 2 de uma reta no espaço 3D constituída por todos os pontos 3D que projetam em p^1 . Define-se $el_2^1(p^2)$ de maneira análoga.
- O *epipolo* e_1^2 da câmera 1 na imagem 2 é a projeção na imagem 2 do centro de câmera c_1 . Define-se e_2^1 de maneira análoga.
- O *plano epipolar* de um ponto 3D p^w é o plano formado por c_1, c_2, p^w .

Sendo assim, a reta epipolar el_1^2 de p^1 reduz o espaço de busca pelo correspondente de p^1 na imagem 2 a uma busca unidimensional. A razão por termos definido estes termos aqui é simplesmente para a compreensão das retas mostradas na Figura 67. Nessa figura, os extremos da curva mostrada em azul na primeira imagem, no canto superior esquerdo da placa de PVC, estão representados por pontos sólidos, um amarelo e o outro verde. Nas demais imagens, tem-se as retas epipolares com cores correspondentes. Para que a geometria das câmeras esteja correta, é necessário que as retas epipolares passem pelos pontos correspondentes. A geometria epipolar é explorada a fundo em sistemas de reconstrução 3D visados por este trabalho (FABBRI, 2010; USUMEZBAS; FABBRI; KIMIA, 2017; USUMEZBAS; FABBRI; KIMIA, 2016c; FABBRI; KIMIA, 2010).

A.3 Calibração Interna Inicial

Para as câmeras de vídeo do *LabOceano*, foram fornecidas as seguintes especificações internas:

- modelo Kodak KAI 02150, com Kowa Motorized Zoom Lens LMZ 1177-IR;
- distância focal: $11 - 77mm$;
- tamanho do píxel: $5.5\mu m \times 5.5\mu m$;
- resolução: $1920 \times 1080px$.

Foi utilizado um ponto principal de acordo com a resolução da imagem, resultando numa mesma matriz de calibração para as quatro câmeras, da forma:

$$K = \begin{bmatrix} \alpha_u & s & u_o \\ 0 & \alpha_v & v_o \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} f/(5.5 \cdot 10^{-3}) & 0 & 1920/2 \\ 0 & f/(5.5 \cdot 10^{-3}) & 1080/2 \\ 0 & 0 & 1 \end{bmatrix}, \quad (35)$$

onde f é desconhecido, porém inicializado em $11mm$ durante as otimizações (*bundle adjustment*) descritas a seguir.

Figura 64 - Câmera FULL HD utilizada no tanque de ondas



Fonte: Equipe Lab Oceano, 2017.

A.4 Calibração Externa Inicial

A disposição das quatro câmeras de vídeo segue a Figura 65. As câmeras 1 e 3 formam uma chamada cabeça estéreo, com linha de base (separação entre os centros das câmeras) curta (*baseline*) de $240mm$ na vertical. As câmeras 2 e 4 formam a mesma cabeça estéreo rotacionada 45° . Dessa forma, tem-se um baseline grande entre as câmeras 1 e 2, assim como entre as câmeras 3 e 4. Para efeitos de inicialização da otimização das câmeras, não será considerada uma rotação entre as câmeras de uma mesma cabeça estéreo, ou seja, será considerado $R_1 = R_3$ e $R_2 = R_4$.

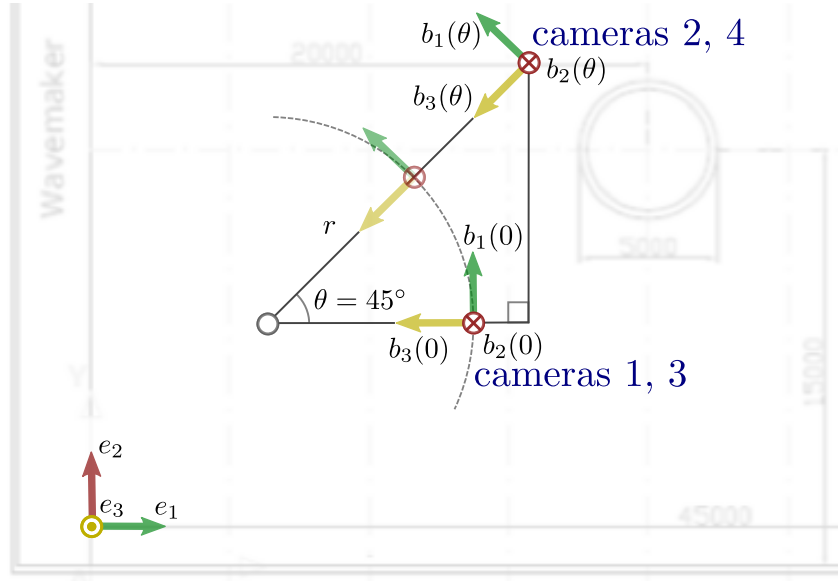
Sejam coordenads de mundo dadas por um sistema de coordenadas 3D fixo no canto do tanque, como indicado na Figura 65. Sejam as coordenadas de mundo $I = \{W, O\}$, e coordenadas de câmera num ângulo θ dadas por $II(\theta) = \{B(\theta), c(\theta)\}$, como na Figura 65. Assim, a solução é $R_1 = R_3 = R(0)$, e $R_2 = R_4 = R(\pi/4)$. Sabe-se que $R(0)$ é a matriz de mudança de base do mundo à câmera:

$$R(0) = \mathcal{M}_{B_0}^W(id) = [\mathcal{X}_{B_0}(e_1) \quad \mathcal{X}_{B_0}(e_2) \quad \mathcal{X}_{B_0}(e_3)], \quad (36)$$

que vem de A.1. Logo, da Figura 65,

$$R(0) = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & -1 \\ -1 & 0 & 0 \end{bmatrix}. \quad (37)$$

Figura 65 - Disposição das cabeças estéreo, vistas de cima.



Fonte: O autor, 2017.

Da mesma forma, da Figura 65, pode-se escrever

$$R(\pi/4) = \begin{bmatrix} -\sin(\pi/4) & \cos(\pi/4) & 0 \\ 0 & 0 & -1 \\ -\cos(\pi/4) & -\cos(\pi/4) & 0 \end{bmatrix}. \quad (38)$$

As translações são obtidas das coordenadas dos centros de câmera $c(\theta)$, que são dadas a partir do projeto de aquisição de dados, Figura 65,

$$\mathcal{X}_{II(\theta)}(p) = R(\theta)\mathcal{X}_W(p) + T(\theta). \quad (39)$$

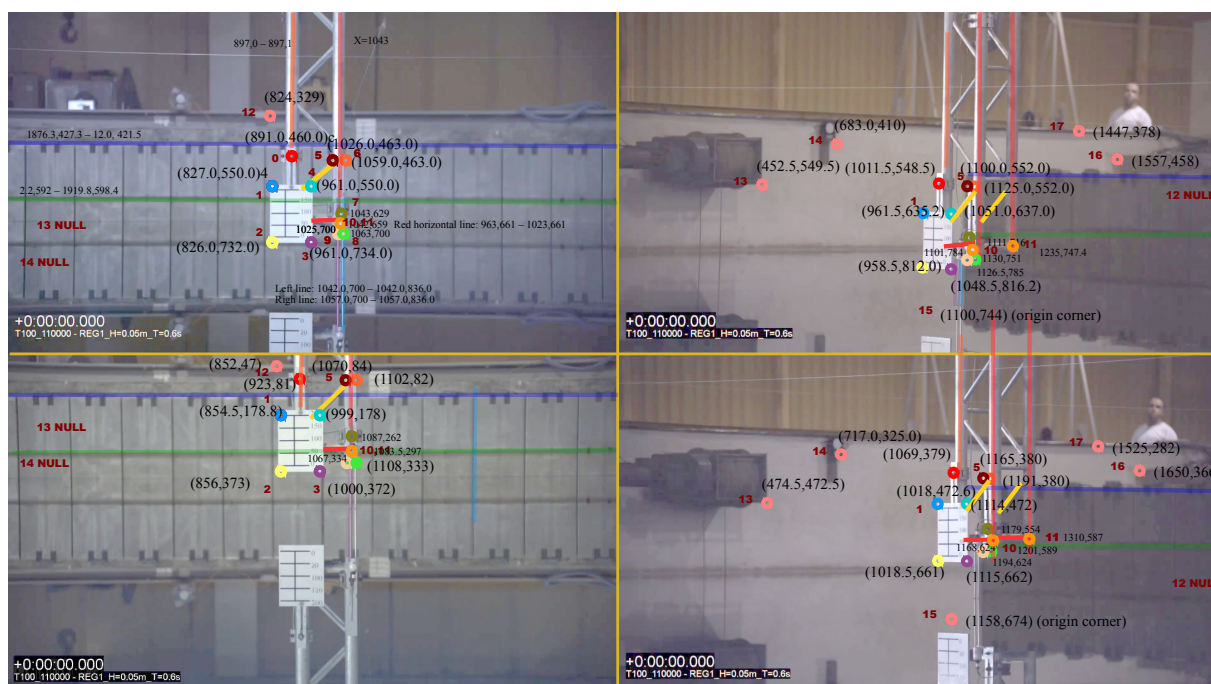
Logo,

$$T(\theta) = -R(\theta)\mathcal{X}_W(c(\theta)). \quad (40)$$

A.5 Otimização das Câmeras

Os parâmetros restantes das câmeras, que não foram especificados no experimento, são deduzidos a partir de correspondências entre as imagens. Foram utilizadas correspondências manuais entre pontos de interesse e entre retas que aparecem nas imagens, Figura 66. No entanto, a otimização a partir de pontos apenas, sem o uso de retas, foi muito mais prática por já está incluída por padrão em bibliotecas de software, como a VXL, [The VXL...](#) (2017).

Figura 66 - Correspondências entre pontos e retas, marcadas manualmente



Fonte: O autor, 2017.

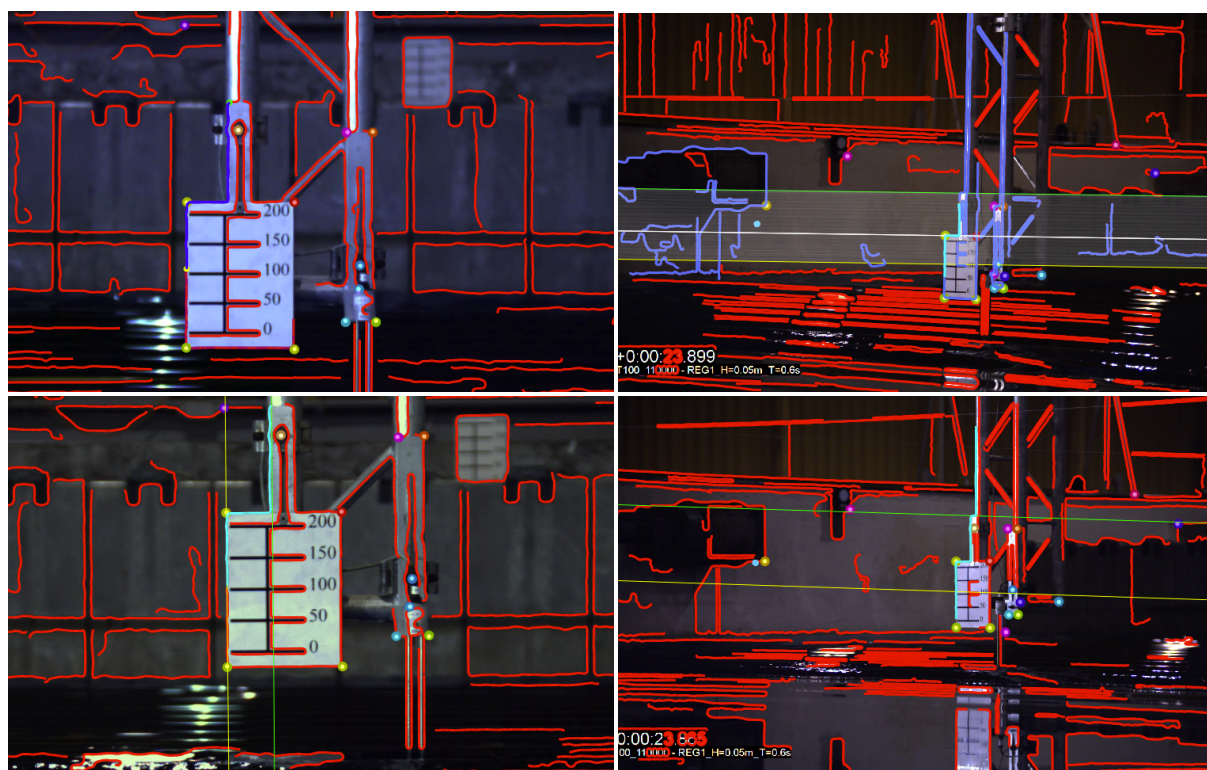
As câmeras foram otimizadas a partir das correspondências usando-se uma variante do método Levenberg-Marquardt esparsa específico para este problema, denominado *Bundle Adjustment* (HARTLEY; ZISSERMAN, 2000; FABBRI, 2010). Como resultado, pode-se visualizar em uma **GUI** a geometria epipolar e a projeção dos pontos subjacentes às correspondências, nos quatro vídeos, Figura 67.

A.5.1 Ambiguidade Projetiva Restante

Existem várias quadruplas de câmeras que produzem o mesmo efeito nas imagens, a mesma geometria epipolar, todas sendo consistentes com o mesmo conjunto de correspondências entre as imagens. Se a matriz de calibração interna K fosse fixa, a única ambiguidade de reconstrução seria uma similaridade. Mais explicitamente, a posição das câmeras obtidas na seção anterior estaria correta modulo um fator de escala global, bem como uma rotação e translação global. Por um lado, isso diz que apenas a posição relativa entre as câmeras é reconstruída, mas não sua rotação e translação absoluta em relação ao sistema de coordenadas global. Por mais que esta informação estivesse incluída na inicialização, é possível que a otimização altere essa posição em relação ao referencial do mundo.

Além disso, a ambiguidade do fator de escala pode ser interpretada da mesma

Figura 67 - Interface gráfica mostrando os pontos utilizados na otimização dos parâmetros das câmeras



Legenda: Interface gráfica mostrando os pontos utilizados na otimização (mesmas cores denotam mesmos pontos em correspondência), com pontos menores ao centro denotando pontos otimizados e projetados nas câmeras. Note-se uma perfeita projeção nos pontos marcados, usando-se as câmeras finais. Também é mostrado como uma curva na primeira imagem, marcada em azul escuro no canto da placa PVC, corresponde perfeitamente à mesma curva nas demais imagens, marcada em ciano. Note-se a disposição das linhas epipolares (linhas de correspondência) consistente com a geometria relativa entre as câmeras

Fonte: O autor, 2017.

forma, como o fato de que as câmeras poderiam estar duas vezes mais distantes entre si, e mesmo assim serem consistentes com as mesmas correspondências. Outra forma de dizer isso é que, na solução final, a unidade de medida pode ser diferente da unidade do mundo (mm). Essa ambiguidade é facilmente resolvida com conhecimento adicional do mundo; no nosso caso, usando-se a informação de que sabemos de antemão a distância entre pontos na placa de PVC no centro da Figura 65.

Como nossa otimização não varia apenas a rotação e translação, mas também a distância focal (uma vez que os técnicos do *LabOceano* não puderam determinar exatamente o valor do efeito zoom utilizado no experimento), há mais ambiguidade envolvida. O fator de escala pode ser confundido com o valor da distância focal. Vale notar que a distância focal sendo otimizada não é o f separadamente, mas sim a entrada α na matriz K :

$$\mathbf{K} = \begin{bmatrix} \alpha & 0 & 1920/2 \\ 0 & \alpha & 1080/2 \\ 0 & 0 & 1 \end{bmatrix}. \quad (41)$$

Para resolver tal ambiguidade, também pode-se utilizar mais informações sobre o mundo. No entanto, isto foge ao nosso propósito imediato e foi deixado para trabalho futuro. As câmeras obtidas, mesmo com tal ambiguidade, já são o suficiente para reconstruções 3D qualitativas.