References

[1] Y. I. Abdel-Aziz and H. M. Karara. Direct linear transformation into object space coordinates. In *Proceedings of the ASP Symposium on Close-Range Photogrammetry (CRP)*, pages 1–18, 1971. 1

[2] M. Andriluka, L. Pishchulin, P. Gehler, and B. Schiele. 2D human pose estimation: New benchmark and state of the art analysis. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 3686–3693, 2014. 4, 6

[3] M. Antunes, J. P. Barreto, D. Aouada, and B. Ottersten. Unsupervised vanishing point detection and camera calibration from a single Manhattan image with radial distortion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 6691–6699, 2017. 1, 2, 3

[4] S. T. Barnard. Interpreting perspective images. *Artificial Intelligence (AIJ)*, 21(4):435–462, 1983. 1

[5] D. C. Brown. Close-range camera calibration. *Photogrammetric Engineering*, 37(8):855–866, 1971. 3

[6] Y. Cai, H. Li, J.-F. Hu, and W.-S. Zheng. Action knowledge transfer for action prediction with partial videos. In *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019. 1

[7] Z. Cao, G. Hidalgo, T. Simon, S.-E. Wei, and Y. Sheikh. OpenPose: Realtime multi-person 2D pose estimation using part affinity fields. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 43(1):172–186, 2021. 6

[8] C.-K. Chang, J. Zhao, and L. Itti. DeepVP: Deep learning for vanishing point detection on 1 million street view images. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 4496–4503, 2018. 2

[9] S.-H. Chang, C.-Y. Chiu, C.-S. Chang, K.-W. Chen, C.-Y. Yao, R.-R. Lee, and H.-K. Chu. Generating 360 outdoor panorama dataset with reliable sun position estimation. In *Proceedings of the ACM SIGGRAPH Asia*, pages 1–2, 2018. 2, 5

[10] B. Cheng, O. Parkhi, and A. Kirillov. Pointly-supervised instance segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 2607–2616, 2022. 1

[11] H.-S. Choi, K. An, and M. Kang. Regression with residual neural network for vanishing point detection. *Image and Vision Computing (IVC)*, 91:103797, 2019. 2

[12] J. M. Coughlan and A. L. Yuille. Manhattan world: Compass direction from a single image by Bayesian inference. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 941–947, 1999. 1, 2

[13] J. S. Dai. Euler–Rodrigues formula variations, quaternion conjugation and intrinsic connections. *Mechanism and Machine Theory (MMT)*, 92:144–152, 2015. 5

[14] A. W. Fitzgibbon. Simultaneous linear estimation of multiple view geometry and lens distortion. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages I125–I132, 2001. 2, 3

[15] S. Huang, Z. Shen, Z. Huang, Z. Ding, J. Dai, J. Han, N. Wang, and S. Liu. Anchor3DLane: Learning to regress 3D anchors for monocular 3D lane detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 17451–17460, 2023. 1

[16] J. Kannala and S. S. Brandt. A generic camera model and calibration method for conventional, wide-angle, and fish-eye lenses. *IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI)*, 28(8):1335–1340, 2006. 2

[17] D. P. Kingma and J. L. Ba. ADAM: A method for stochastic optimization. In *Proceedings of the International Conference on Learning Representations (ICLR)*, 2015. 6

[18] F. Kluger, H. Ackermann, M. Y. Yang, and B. Rosenhahn. Deep learning for vanishing point detection using an inverse gnomonic projection. In *Proceedings of the German Conference on Pattern Recognition (GCPR)*, pages 17–28, 2017. 2

[19] F. Kluger, E. Brachmann, H. Ackermann, C. Rother, M. Y. Yang, and B. Rosenhahn. CONSAC: Robust multi-model fitting by conditional sample consensus. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4634–4643, 2020. 2

[20] I. Kotseruba, A. Rasouli, and J. K. Tsotsos. Benchmark for evaluating pedestrian action prediction. In *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision (WACV)*, pages 1257–1267, 2021. 1

[21] H. Law and J. Deng. CornerNet: Detecting objects as paired keypoints. *International Journal of Computer Vision (IJCV)*, 128:642–656, 2020. 3

[22] J. Lee, M. Sung, H. Lee, and J. Kim. Neural geometric parser for single image camera calibration. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 541–557, 2020. 2

[23] J. Lee, H. Go, H. Lee, S. Cho, M. Sung, and J. Kim. CTRL-C: Camera calibration TRansformer with line-classification. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 16228–16237, 2021.

[24] H. Li, J. Zhao, J.-C. Bazin, W. Chen, Z. Liu, and Y.-H. Liu. Quasi-globally optimal and efficient vanishing point estimation in Manhattan world. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 1646–1654, 2019. 2

[25] K. Li, K. Chen, H. Wang, L. Hong, C. Ye, J. Han, Y. Chen, W. Zhang, C. Xu, D.-Y. Yeung, X. Liang, Z. Li, and H. Xu. CODA: A real-world road corner case dataset for object detection in autonomous driving. In *Proceedings of the European Conference on Computer Vision (ECCV)*, pages 406–423, 2022. 1

[26] L. Li, T. Zhou, W. Wang, J. Li, and Y. Yang. Deep hierarchical semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 1236–1247, 2022. 1

[27] Y.-L. Li and S. Wang. R(Det)^2: Randomized decision routing for object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 4815–4824, 2022. 1