Evaluating Visual Inertial Odometry Using the Windy Forest Dataset

Sam Schofield
Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
sam.schofield@pg.canterbury.ac.nz

Andrew Bainbridge-Smith
Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
andrew.bainbridge-smith@canterbury.ac.nz

Richard Green
Computer Science and Software Engineering
University of Canterbury
Christchurch, New Zealand
richard.green@canterbury.ac.nz

2021 36th International Conference on Image and Vision Computing New Zealand (IVCNZ) | 978-1-6654-0645-1/21/$31.00 ©2021 IEEE | DOI: 10.1109/IVCNZ54163.2021.9653391

***Abstract*—This paper examines the effect of moving vegetation on visual-inertial-odometry performance. The evaluation was performed using a custom dataset consisting of fifteen sequences in a forest scene. The sequences are comprised of three trajectories, each with five different wind/scene motion levels — allowing the relationship between scene motion and state-estimation accuracy to be examined. The dataset was created using a "semi-synthetic" approach that combines real IMU data and synthetic images to provide increased realism while maintaining precise control over the scene. The dataset is used to assess the effect of scene motion on multiple open-source visual-inertial odometry algorithms in a variety of configurations. Additionally, the results are validated using real-world sequences containing imagery of vegetation moving in the wind. The results show that scene motion as a result of moving vegetation can worsen visual-inertial odometry performance by up to a factor of ten.**

I. INTRODUCTION

Visual inertial odometry (VIO) and simultaneous localisation and mapping (VI-SLAM) have reached a high level of maturity and have several performant open-source solutions available [1]–[5]. One factor that has allowed the field to mature so far is the availability of many high-quality benchmark datasets [6]–[8]. Datasets allow algorithms to be compared effectively and make algorithm development more accessible, as they remove the overheads associated with data capture for their users. Datasets also help to focus the community on specific problems within the field. For example, the KITTI dataset [6] focused on the challenges involved with visual pose estimation for autonomous driving. The Drone racing dataset [9], on the other hand, targeted the problem of pose estimation during aggressive trajectories. One challenge that has not been addressed by any existing dataset is performing visual-inertial state estimation around moving vegetation such as trees and grass.

Making visual odometry robust to dynamic scenes is an active area of research. Most existing datasets with dynamic scenes target urban environments [6], [10]–[12]; as such, much of the existing work on scene-motion-robust VO focuses on providing robustness in the presence of moving cars and people [13]–[15]. Robustness in vegetated environments, however, remains relatively unstudied yet is essential for expanding robotics into areas such as agriculture.

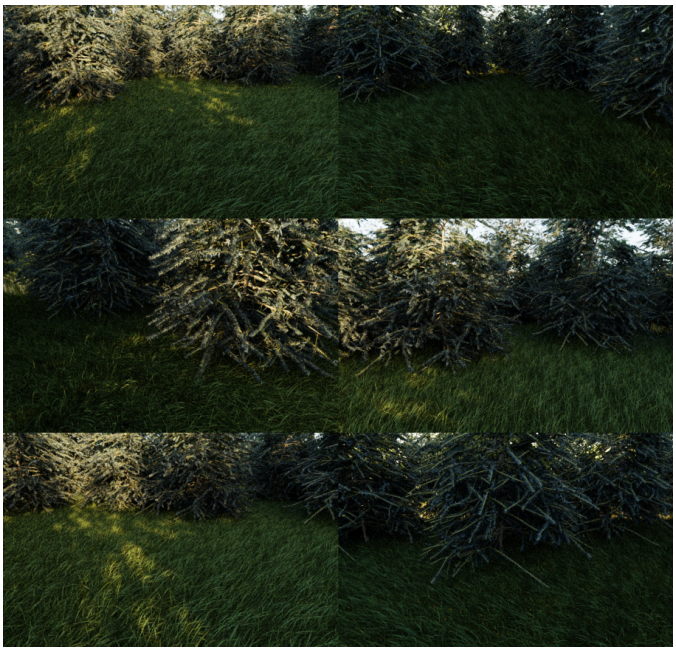


Fig. 1: Example images from the Windy Forest dataset. Images are generated using Blender with a FOV of 88 degrees, a resolution of 752×480 and a rate of 20 Hz.

Scene-motion-robust VO algorithms often work by separating the moving objects from the static background then using the background for pose estimation [13], [15]–[18]. This segmentation of foreground and background often relies on assumptions about the scene's geometry, e.g. that the background consists of lines or planes [16], [17], [19], or about the moving objects, e.g. identifying (and ignoring) cars and pedestrians using machine learning [13], [15], [18]. However, these methods are unlikely to work in a vegetated environment such as a forest where the scene's geometry is relatively unpredictable. Furthermore, even if the foreground and background can be segmented, the background is also likely to be moving, which may cause significant problems for existing methods.

The nature of scene motion in vegetated environments is also different from that found in existing datasets with moving objects. Pedestrians and cars (often found in existing datasets) can move freely about their environments and have

(relatively) simple deformations. On the other hand, plants are highly deformable but mainly oscillate about a single point. These differences make vegetated environments an interesting challenge for visual odometry. We hope that overcoming these challenges will result in more robust algorithms both around vegetation and in general.

This paper presents the Windy Forest dataset (see Fig. 1 for example images), a photorealistic visual-inertial dataset aimed at evaluating the robustness of VIO algorithms in the presence of scene motion. The dataset provides fifteen sequences consisting of three trajectories travelling through the same forest scene with different wind/motion settings (see Fig. 2). The dataset is "semi-synthetic", consisting of real IMU data and images rendered along the trajectory of a real multi-rotor. This approach allows the data to be as realistic as possible while allowing precise control over the motion in the scene. 752×480 stereo RGB images and IMU data are provided at 20 Hz and 200 Hz, respectively. Ground truth depth, optical flow and camera poses are also available. The dataset is used to evaluate many state-of-the-art VIO algorithms in different configurations to evaluate their suitability for use around moving vegetation.

II. RELATED WORK

There are many datasets for evaluating visual-inertial odometry algorithms. However, there are no existing datasets for evaluating these algorithms that provide both challenging camera motion and dynamic vegetation.

One of the primary motivations for VO research has been autonomous driving. As such, many datasets consist of sensor data captured from a car in an urban environment [6], [20], [21]. These datasets have enabled much progress in the field. However, the camera motion is relatively constrained (little/no pitch, roll or change in elevation). Additionally, while these datasets often contain scene motion, it is almost exclusively caused by cars and pedestrians. These "distractors" can often be dealt with by segmenting the objects from the background — either geometrically (using lines and planes to find the background [16], [17], [19]) or using machine learning (to detect objects likely to be moving, e.g., people and cars [13], [15], [18]). These approaches are not likely to work in natural environments where fewer assumptions about the scene can be made.

Other datasets target UAVs [8], [9], [22], [23]. These datasets typically have much more challenging camera motion than their road-vehicle counterparts — particularly those focusing on fast motion, such as the Drone racing [9] and BlackBird datasets [23]. However, these often have even less focus on moving scenery than road-vehicle alternatives. Hand-held datasets [7], [10], [24], [25] fall into a similar category. They typically provide more challenging motion than road-vehicle datasets but do not offer particularly novel scene motion.

Datasets captured in vegetated environments do exist [9], [26]–[29], but come with their own set of limitations. The Rosario [28] and Symphony lake [29] datasets provide novel scenery but have slow 3-DOF camera motion. The Drone Racing [9] and ETH3D [27] datasets provide much more challenging camera trajectories through vegetated scenes; however, the optical flow in these datasets is dominated by the camera's motion instead of scene's. The TartanAir dataset [26] provides sequences of complex camera motion through a forest (among others). However, the forest sequences appear to focus on scene motion caused by falling leaves against a static background. Although this is an interesting challenge, it is different from dealing with trees and grass blowing in the wind, which is a more common and arguably more challenging scenario.

III. DATASET

The Windy Forest dataset comprises fifteen sequences consisting of three trajectories travelling through the same scene at five different wind/motion levels. The wind levels range from zero (no motion) to one (full motion). The trajectories correspond to the V1_01, V1_02, V1_03 trajectories from the EuRoC dataset [8].

A. Semi-synthetic data

The dataset was generated using the semi-synthetic approach detailed in our previous work [30]. A semi-synthetic sequence fuses real IMU data and synthetic images to provide precise control over the scene while keeping the sequence as realistic as possible. The data is created by measuring the trajectory of an IMU while recording its measurements then rendering images along the trajectory. The real IMU data and synthetic images are then used to evaluate VIO algorithms. The specific approach used in this paper uses the tool provided in [31] to combine motion capture and IMU measurements to estimate the IMU's trajectory more accurately than motion capture alone. The additional accuracy is required to ensure the IMU data and synthetic images are consistent with each other. A visual representation of the system is given in Fig 3.

B. Sensors

The dataset provides sensor data from two (identical) synthetic cameras and a real IMU. Table I gives a summary of the relevant sensor characteristics. In addition to the images and IMU data, ground truth depth, optical flow and trajectory data is also available.

C. Wind simulation

The dataset provides five different wind/motion levels for each of its trajectories. The wind levels were created by generating wind-animated trees and grass using SpeedTree, then altering the "influence" of the animation in Blender. This approach allowed the scene to be kept consistent across wind levels, as only the magnitude of the animation varies — the position of the trees, as well as the timing and direction of the wind gusts, are all identical.

D. Scene motion analysis

To provide insight into the dataset, we perform analysis on both the scene ($\boldsymbol{F}_{\text{scene}}$) and camera-motion-induced ($\boldsymbol{F}_{\text{camera}}$) optical flow. $\boldsymbol{F}_{\text{camera}}$ is calculated by projecting the ground truth depth at time t onto the image plane at time $t+1$ using

(a) Reference image. (b) Scene motion (wind level = 0.25). (c) Scene motion (wind level = 1.00).

Fig. 2: Examples of scene motion optical flow for the same frame for the 0.25 (Fig. 2b) and 1.00 (Fig. 2c) wind levels. The images show that the scene-motion-induced optical flow increases significantly as the wind level increases.

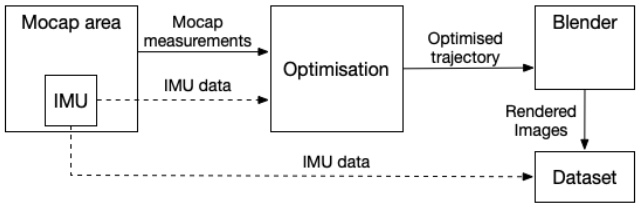


Fig. 3: The method used to generate semi-synthetic visual inertial sequences for this dataset.

TABLE I: Summary of the sensors characteristics used in the dataset.

Sensor	Parameter	Value
Camera	Resolution ($w \times h$)	752×480
	Frame rate (Hz)	20
	Sample rate (per-pixel)	256
	Camera model	Pinhole
	Distortion	None
	Sensor size (mm)	6.4
	Focal length (mm)	3.9
	FOV (diagonal, deg)	88
IMU	IMU model	ADIS16448
	IMU rate	200 Hz
Ground truth pose	Rate	20 Hz
Depth	Rate	20 Hz
	Resolution ($w \times h$)	752×480
Optical flow	Rate	20 Hz
	Resolution ($w \times h$)	752×480

the ground truth camera poses. $\boldsymbol{F}_{\text{scene}}$ is then calculated by subtracting $\boldsymbol{F}_{\text{camera}}$ from the ground truth (total) optical flow. Fig. 4 shows the deconstruction of the optical flow into its camera and scene motion components.

Fig. 5 and Fig. 6 show the distribution of camera and scene induced optical flow across the different wind levels. The optical flow induced by the camera motion is very similar across wind levels (small differences result from the wind-animation magnitude causing the branches and leaves to be in slightly different positions for equivalent frames). In contrast, the scene-motion induced optical flow shows significant differences in both the maximum magnitude and proportion of moving pixels. For reference, Fig. 7 shows how the percentage of moving pixels changes throughout the sequence.

TABLE II: Summary of the evaluated algorithms.

Algorithm	Stereo/Mono	Inertial?	Backend
OpenVINS	Stereo	Yes	MSCKF
MSCKF	Stereo	Yes	MSCKF
Kimera	Stereo	Yes	Fixed-lag smoother
VINS-Fusion (stereo VIO)	Stereo	Yes	Fixed-lag smoother
VINS-Fusion (stereo VO)	Stereo	**No**	Fixed-lag smoother
VINS-Fusion (mono VIO)	**Mono**	Yes	Fixed-lag smoother

IV. Evaluation

A. Algorithms

We evaluate multiple state of the art, open-source VIO algorithms, namely: OpenVINS [2], VINS-Fusion [3], MSCKF [5] and Kimera [4]. It should be noted that although most of these algorithms support loop closure, this feature was disabled so the pure VIO could be evaluated. VINS-Fusion provides multiple variants, including stereo, monocular VIO and stereo VIO. These variants allow the relationship between performance in dynamic scenes and sensor configuration to be examined. Additionally, a range of filtering and smoothing-based algorithms were selected to evaluate whether this had an impact on performance. Table II provides an overview of different algorithms evaluated.

B. Results

The benchmarking suite proposed in [32] was used to calculate the absolute trajectory error (ATE) for each algorithm across the different wind levels for each sequence. The results for the V1_01 sequences are given in Fig. 8. The results show that the error for each of the algorithms increases significantly with the wind level. Additionally, the spread between algorithms also increases with wind level, suggesting that scene motion impacts some methods more than others. One interesting observation is that the magnitude of the scene motion (Fig. 6) is relatively small compared to the camera motion (Fig. 5) even for the highest wind level — highlighting how sensitive the algorithms are to scene motion.

The results also give some insight into how different algorithm configurations affect performance. The addition of an IMU appears to have the most significant impact. Stereo-only VINS-Fusion outperforms its inertial counterpart for the low wind level but performs significantly worse as the wind/motion

(a) Total optical flow. (b) Camera motion optical flow. (c) Scene motion optical flow.

Fig. 4: A visualisation of the optical flow components. The camera motion optical flow (Fig. 4b) is calculated by projecting the depth map at frame i onto the image plane at time i+1 using the ground truth camera pose. The scene motion optical flow (Fig. 4c) is calculated by subtracting the camera motion (Fig. 4b) from the ground truth optical flow (Fig. 9e).

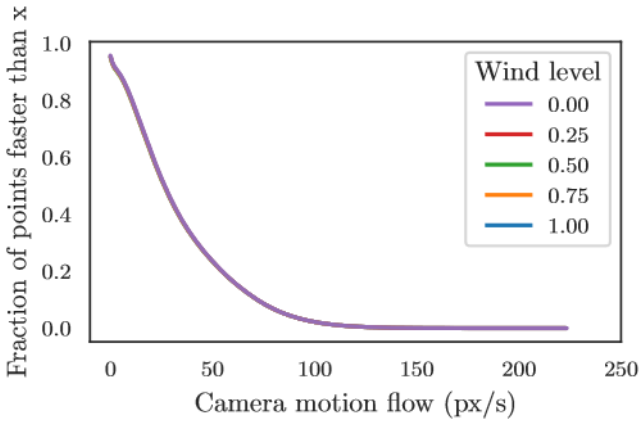


Fig. 5: A comparison of the camera-motion-induced optical flow (see Fig 4b) across the five wind levels of the V1_01 sequence. The figure shows the proportion of the pixels in the sequences that have an optical flow greater than x pixels per second.

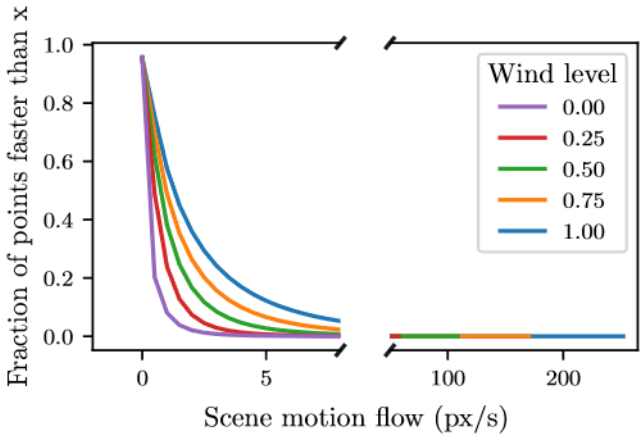


Fig. 6: A comparison of the scene-motion-induced optical flow (see Fig 4c) across the five wind levels of the V1_01 sequence. The figure shows the proportion of the pixels in the sequences that have an optical flow greater than x pixels per second.

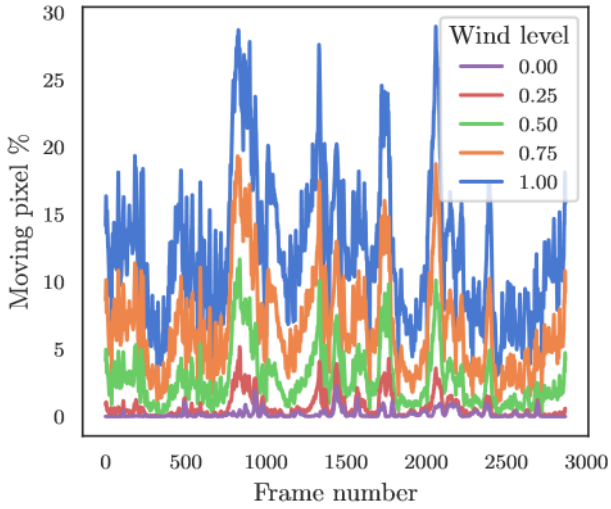


Fig. 7: A comparison of the per-frame percentage of pixels with greater than one pixel of scene motion optical flow for the V1_01 sequence.

increases. In contrast, the addition of a second camera (VINS-Fusion (mono VIO) and VINS-Fusion (stereo VIO)) appears to make little difference to performance. Similarly, the use of filtering or smoothing does not appear to be a deciding factor in performance as OpenVINS (filter) and VINS-Fusion (smoother) both appear to perform well throughout the wind ranges.

V. Comparison to Real Data

In addition to testing VIO algorithms on our semi-synthetic dataset, we also provide a comparison on five real sequences (Fig. 9). The data was captured using a synchronised visual-inertial system consisting of two FLIR Chameleon3 cameras recording 1280x1024 at 20 Hz and an ADIS16448 IMU recording at 200 Hz. The sequences were captured by moving the system around in a looping trajectory to give the algorithms a chance to initialise, then placing it on a stationary tripod facing the desired scene. The scenes consist predominantly of vegetation (trees/bushes and grass) and have varying amounts of scene motion (Fig. 10).

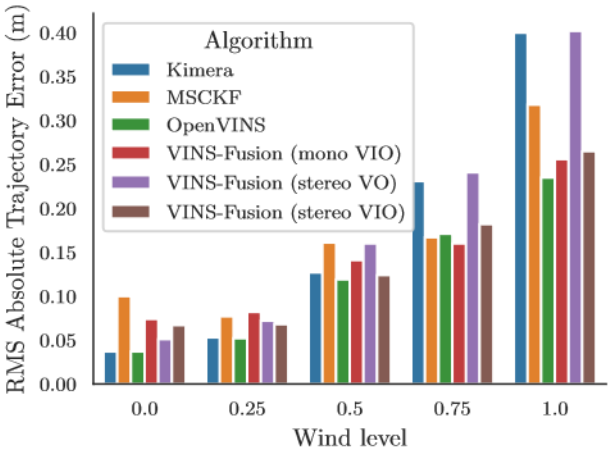
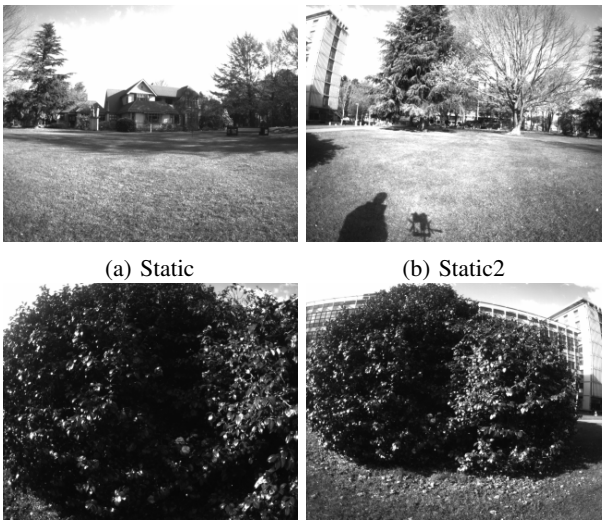


Fig. 8: Results on the V1_01 sequences.



(a) Static (b) Static2

(c) Dynamic (d) Dynamic2

(e) Dynamic3

Fig. 9: Example images from the five verification sequences.

The algorithms are evaluated on the five sequences. Each algorithm is run on the entire sequence then the estimated trajectory is clipped to the period when the camera is stationary. Clipping the trajectory to the stationary period provides a known (stationary) camera trajectory to calculate the absolute trajectory error. The ATE for each algorithm-sequence combination is shown in Fig. 11. For reference, the mean optical flow for each sequence is also given. The results show that the algorithms typically perform worse as the scene motion

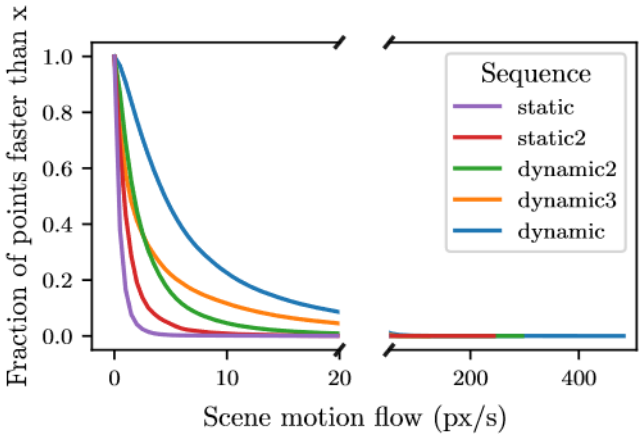


Fig. 10: Comparison of the scene-motion-induced optical flow across the five real verification sequences. The scene motion of each sequence is calculated by extracting approximately 500 features in each image and tracking them into the next image using optical flow. Since the camera is stationary, any optical flow must be induced by scene motion.

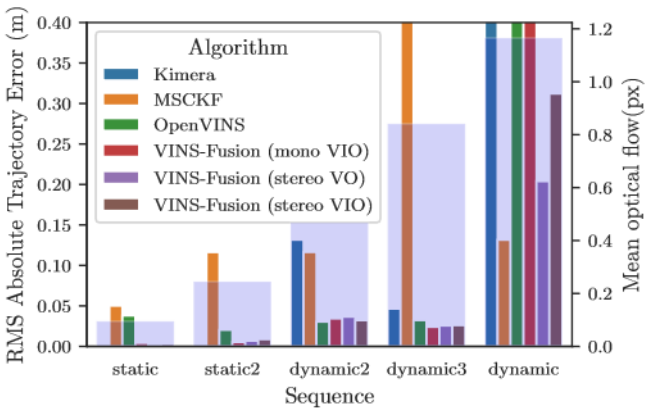


Fig. 11: Results on the real verification sequences. Note that the error on the "static" and "static2" sequences was very low for some algorithms, making them appear to be missing.

increases, although the relationship is less evident than for the semi-synthetic data (Fig. 8). However, this observation is expected as the real data does not have the same amount of control over the scene, so factors other than scene motion (e.g. brightness, abundance of features, scene geometry) may also influence VIO performance.

VI. Conclusion

This paper presented an evaluation of the effect of moving vegetation on visual-inertial odometry performance. A custom dataset was created to isolate the impact of scene motion from other factors such as camera trajectory and scene geometry. The dataset was used to evaluate multiple open-source VIO algorithms, and the results were verified against real-world data. The results showed that moving vegetation poses a

serious challenge to the existing state-of-the-art visual-inertial odometry algorithms — a challenge that needs to be addressed if VIO is going to be used in natural environments.

ACKNOWLEDGEMENTS

The research reported in this article was conducted as part of "Enabling unmanned aerial vehicles (drones) to use tools in complex dynamic environments", which is funded by the New Zealand Ministry of Business, Innovation and Employment (UOCX2113).

REFERENCES

[1] V. Usenko, N. Demmel, D. Schubert, J. Stuckler, and D. Cremers, "Visual-Inertial Mapping With Non-Linear Factor Recovery," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 422–429, 4 2020. [Online]. Available: https://ieeexplore.ieee.org/document/8938825/

[2] P. Geneva, K. Eckenhoff, W. Lee, Y. Yang, and G. Huang, "OpenVINS: A Research Platform for Visual-Inertial Estimation," in *Proc. of the IEEE International Conference on Robotics and Automation*, Paris, France, 2020. [Online]. Available: https://docs.openvins.com

[3] T. Qin, P. Li, and S. Shen, "VINS-Mono: A Robust and Versatile Monocular Visual-Inertial State Estimator," *IEEE Transactions on Robotics*, vol. 34, no. 4, pp. 1004–1020, 8 2018. [Online]. Available: https://ieeexplore.ieee.org/document/8421746/

[4] A. Rosinol, M. Abate, Y. Chang, and L. Carlone, "Kimera: an Open-Source Library for Real-Time Metric-Semantic Localization and Mapping," vol. 11680088, no. Id 100010434, 10 2019. [Online]. Available: http://arxiv.org/abs/1910.02490

[5] K. Sun, K. Mohta, B. Pfrommer, M. Watterson, S. Liu, Y. Mulgaonkar, C. J. Taylor, and V. Kumar, "Robust Stereo Visual Inertial Odometry for Fast Autonomous Flight," *IEEE Robotics and Automation Letters*, vol. 3, no. 2, pp. 965–972, 4 2018. [Online]. Available: http://arxiv.org/abs/1712.00036https://ieeexplore.ieee.org/document/8258858/

[6] A. Geiger, P. Lenz, and R. Urtasun, "Are we ready for autonomous driving? The KITTI vision benchmark suite," in *2012 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, 6 2012, pp. 3354–3361. [Online]. Available: http://ieeexplore.ieee.org/document/6248074/

[7] D. Schubert, T. Goll, N. Demmel, V. Usenko, J. Stuckler, and D. Cremers, "The TUM VI Benchmark for Evaluating Visual-Inertial Odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10 2018, pp. 1680–1687. [Online]. Available: https://ieeexplore.ieee.org/document/8593419/

[8] M. Burri, J. Nikolic, P. Gohl, T. Schneider, J. Rehder, S. Omari, M. W. Achtelik, and R. Siegwart, "The EuRoC micro aerial vehicle datasets," *The International Journal of Robotics Research*, vol. 35, no. 10, pp. 1157–1163, 9 2016. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364915620033

[9] J. Delmerico, T. Cieslewski, H. Rebecq, M. Faessler, and D. Scaramuzza, "Are We Ready for Autonomous Drone Racing? The UZH-FPV Drone Racing Dataset," in *2019 International Conference on Robotics and Automation (ICRA)*. IEEE, 5 2019, pp. 6713–6719. [Online]. Available: https://ieeexplore.ieee.org/document/8793887/

[10] S. Cortés, A. Solin, E. Rahtu, and J. Kannala, "ADVIO: An authentic dataset for visual-inertial odometry," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11214 LNCS, pp. 425–440, 2018.

[11] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of rgb-d slam systems," in *Proc. of the International Conference on Intelligent Robot Systems (IROS)*, Oct. 2012.

[12] K. M. Judd and J. D. Gammell, "The Oxford Multimotion Dataset: Multiple SE(3) Motions With Ground Truth," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 800–807, 4 2019. [Online]. Available: https://ieeexplore.ieee.org/document/8610209/

[13] M. Henein, J. Zhang, R. Mahony, and V. Ila, "Dynamic SLAM: The Need For Speed," 2 2020. [Online]. Available: http://arxiv.org/abs/2002.08584

[14] X. Bai, B. Zhang, W. Wen, L. T. Hsu, and H. Li, "Perception-aided Visual-Inertial Integrated Positioning in Dynamic Urban Areas," *2020 IEEE/ION Position, Location and Navigation Symposium, PLANS 2020*, pp. 1563–1571, 2020.

[15] P. Irmisch, D. Baumbach, and I. Ernst, "ROBUST VISUAL-INERTIAL ODOMETRY IN DYNAMIC ENVIRONMENTS USING SEMANTIC SEGMENTATION FOR FEATURE SELECTION," *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, vol. V-2-2020, pp. 435–442, 8 2020. [Online]. Available: https://www.isprs-ann-photogramm-remote-sens-spatial-inf-sci.net/V-2-2020/435/2020/

[16] Y. Sun, M. Liu, and M. Q. H. Meng, "Motion removal for reliable RGB-D SLAM in dynamic environments," *Robotics and Autonomous Systems*, vol. 108, pp. 115–128, 2018. [Online]. Available: https://doi.org/10.1016/j.robot.2018.07.002

[17] D. H. Kim and J. H. Kim, "Effective background model-based RGB-D dense visual odometry in a dynamic environment," *IEEE Transactions on Robotics*, vol. 32, no. 6, pp. 1565–1573, 2016.

[18] J. Cheng, Z. Wang, H. Zhou, L. Li, and J. Yao, "DM-SLAM: A Feature-Based SLAM System for Rigid Dynamic Scenes," *ISPRS International Journal of Geo-Information*, vol. 9, no. 4, p. 202, 3 2020. [Online]. Available: https://www.mdpi.com/2220-9964/9/4/202

[19] E. Yao, H. Zhang, H. Xu, H. Song, and G. Zhang, "Robust RGB-D visual odometry based on edges and points," *Robotics and Autonomous Systems*, vol. 107, pp. 209–220, 2018. [Online]. Available: https://doi.org/10.1016/j.robot.2018.06.009

[20] Y. Cabon, N. Murray, and M. Humenberger, "Virtual KITTI 2," pp. 1–11, 1 2020. [Online]. Available: http://arxiv.org/abs/2001.10773

[21] A. R. Gaspar, A. Nunes, A. M. Pinto, and A. Matos, "Urban@CRAS dataset: Benchmarking of visual odometry and SLAM techniques," *Robotics and Autonomous Systems*, vol. 109, pp. 59–67, 2018.

[22] A. L. Majdik, C. Till, and D. Scaramuzza, "The Zurich urban micro aerial vehicle dataset," *The International Journal of Robotics Research*, vol. 36, no. 3, pp. 269–273, 3 2017. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364917702237

[23] A. Antonini, W. Guerra, V. Murali, T. Sayre-McCord, and S. Karaman, "The Blackbird Dataset: A large-scale dataset for UAV perception in aggressive flight," pp. 130–139, 10 2018. [Online]. Available: http://link.springer.com/10.1007/978-3-030-33950-0_12http://arxiv.org/abs/1810.01987

[24] B. Pfrommer, N. Sanket, K. Daniilidis, and J. Cleveland, "PennCOSYVIO: A challenging Visual Inertial Odometry benchmark," in *2017 IEEE International Conference on Robotics and Automation (ICRA)*. IEEE, 5 2017, pp. 3847–3854. [Online]. Available: http://ieeexplore.ieee.org/document/7989443/

[25] J. Sturm, N. Engelhard, F. Endres, W. Burgard, and D. Cremers, "A benchmark for the evaluation of RGB-D SLAM systems," *IEEE International Conference on Intelligent Robots and Systems*, pp. 573–580, 2012.

[26] W. Wang, D. Zhu, X. Wang, Y. Hu, Y. Qiu, C. Wang, Y. Hu, A. Kapoor, and S. Scherer, "TartanAir: A Dataset to Push the Limits of Visual SLAM," in *2020 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10 2020, pp. 4909–4916. [Online]. Available: http://arxiv.org/abs/2003.14338https://ieeexplore.ieee.org/document/9341801/

[27] T. Schops, T. Sattler, and M. Pollefeys, "BAD SLAM: Bundle Adjusted Direct RGB-D SLAM," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 6 2019, pp. 134–144. [Online]. Available: https://ieeexplore.ieee.org/document/8954208/

[28] T. Pire, M. Mujica, J. Civera, and E. Kofman, "The Rosario dataset: Multisensor data for localization and mapping in agricultural environments," *The International Journal of Robotics Research*, vol. 38, no. 6, pp. 633–641, 5 2019. [Online]. Available: http://journals.sagepub.com/doi/10.1177/0278364919841437

[29] S. Griffith, G. Chahine, and C. Pradalier, "Symphony Lake Dataset," *International Journal of Robotics Research*, vol. 36, no. 11, pp. 1151–1158, 2017.

[30] S. Schofield, A. Bainbridge-Smith, and R. Green, "An Improved Semi-Synthetic Approach for Creating Visual-Inertial Odometry Datasets," *Submitted to - IEEE Robotics and Automation Letters*, 2021.

[31] P. Geneva and G. Huang, "Vicon2gt: Derivations and Analysis," University of Delaware, Tech. Rep. 1, 2020.

[32] Z. Zhang and D. Scaramuzza, "A Tutorial on Quantitative Trajectory Evaluation for Visual(-Inertial) Odometry," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 10 2018, pp. 7244–7251. [Online]. Available: https://ieeexplore.ieee.org/document/8593941/