Evaluation Instruction of SLAM-for-AR Competition

The Format of Submission Result

The estimated 6 DoF camera poses (from camera coordinate to the world coordinate) are required to evaluate the performance. Considering that there is a certain randomness of estimation, each sequence is required to be run for 5 times, resulting in 5 pose files and 5 running time files. We will select the median result from all five results for evaluation. It should be noted that

The format for each pose file is described as follows:

…

timestamp[i] p_x p_y p_z q_x q_y q_z q_w

…

where (p_x, p_y, p_z) is the camera position, and the unit quaternion (q_x, q_y, q_z, q_w) is the camera orientation. You should output the **real-time poses** after each frame is processed (not the poses after final global optimization), and the output of poses should be in the same frame rate as the input camera images (Otherwise, the completeness evaluation would be affected).

The format for each running time file is described as follows

…

timestamp[i] t_{pose}

…

where t_{pose} denotes the system time when the pose is estimated.

Please submit a zip file containing all the poses and running time files. The structure of zip file should follow the form described as follows:

YourSLAMName/sequence_name/Round-pose.txt

YourSLAMName/sequence_name/Round-time.txt

For example,

MY-SLAM/C0_test/0-pose.txt

MY-SLAM/C0_test/0-time.txt

Evaluation

We evaluate the overall performance of a SLAM system considering tracking accuracy, initialization quality, tracking robustness, relocalization time and the computation efficiency. The criteria are as follows:

- ε_{APE} / ε_{ARE} - absolute positional / rotational error
- ε_{RPE} / ε_{RRE} - relative positional / rotational error
- ε_{bad} – the ratio of bad poses (100% - completeness)
- ε_{init} - initialization quality
- ε_{RO} - tracking robustness
- t_{RL} - relocalization time

The detailed description of the above criteria can be found in the following paper:
Jinyu Li, Bangbang Yang, Danpeng Chen, Nan Wang, Guofeng Zhang, Hujun Bao. Survey and evaluation of monocular visual-inertial SLAM algorithms for augmented reality. Journal of Virtual Reality & Intelligent Hardware, 2019, 1(4): 386–410. DOI:10.3724/SP.J.2096-5796.2018.0011. URL: http://www.vr-ih.com/vrih/html/EN/10.3724/SP.J.2096-5796.2018.0011/article.html

We convert each criteria error ε_i into a normalized score by $s_i = \frac{\sigma_i^2}{\sigma_i^2 + \varepsilon_i^2} \times 100\%$, where σ_i is the variance controlling the normalization function shape. The complete score is a weighted sum of all the individual scores as:

$$S = w_{APE} s_{APE} + w_{ARE} s_{ARE} + w_{RPE} s_{RPE} + w_{RRE} s_{RRE} + w_{bad} s_{bad} + w_{init} s_{init} + w_R s_R + w_{RL} s_{RL}$$

The weight w and variance σ (V-SLAM / VI-SLAM) for each criteria are listed below:

w_{APE}	w_{ARE}	w_{RPE}	w_{RRE}	w_{bad}	w_{init}	w_{RO}	w_{RL}
1.0	1.0	0.5	0.5	1.0	1.0	1.0	1.0
σ_{APE}	σ_{ARE}	σ_{RPE}	σ_{RRE}	σ_{bad}	σ_{init}	σ_R	σ_{RL}
72.46/55.83	7.41/2.48	6.72/2.92	0.26/0.17	20.68/2.38	2.79/1.85	2.27/0.95	0.65/1.42

You can evaluate your SLAM system with our training dataset using the evaluation tool: https://github.com/zju3dv/eval-vislam.

In the final round competition, we will test all systems on benchmarking PCs with the same hardware configuration. The running time will be taken into account for computing the final score according to the following equation:

$$S^* = \frac{\min(30, \mathit{framerate})}{30} S$$

where *framerate* denotes the average framerate of the system.