

In the article ORB-SLAM2: an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras Mur-Artal and Tardós introduce their ORB-SLAM2 and showcase its performance in comparison to other SLAM systems. In SLAM (Simultaneous localization and mapping) a map is constructed from the sensor data of the surroundings. Usually camera images are used as sensor data and a method choice is required to choose how to fuse information from images with different viewpoints

The method, for fusing information from images, they authors chose was bundle adjustment (BA), they mention that others have used ICP (Iterative closest point) and direct methods. You would think that they choose the method that gives the most accurate map, because that is what the system is used for. But Mur-Artal and Tardós mentioned that their goal was to achieve a long-term and globally consistent localization instead of the most detailed map. They state that: "Differently to all above methods our back-end is based on bundle adjustment and builds a globally consistent sparse reconstruction. Therefore our method is lightweight and works with standard CPUs." This goal also motivates why they use different BAs in their system. One BA, for the local area, which is used frequently and another BA, for the whole map, that is used only when necessary. Doing so makes the system more lightweight but the map is not as accurate over the whole map all the time.

The authors wrote that the recent Stereo LSD-SLAM used a direct approach is not relying on features and is therefore more robust to motion blur and poorly-textured environments. The downside of direct methods is that the performance can be severely degraded by unmodeled effects like rolling shutters and non-lambertian reflectance.

When it comes to ICP, they mentioned that another SLAM system using ICP in combination with fusing of all depth data from the sensor into a volumetric dense model, was one of the reasons that system was limited to small workspaces. It is a bit vague if the limited workspace was caused by using ICP or if the ICP requires a volumetric dense model of the depth data, but I'd say still quite good justifications for not using it. I guess the people reading this article find this trivial and that is the reason they didn't mention it.

As the name suggests; ORB-SLAM2 uses ORB (Oriented FAST and Rotated BRIEF) feature descriptor but I cannot see any justification to why it was used. There are other feature descriptors like SIFT and SURF and even more. As I stated above the authors probably aimed for people who already know these things. But to me it is a problem, they don't even mention what ORB stands for. Since SURF and SIFT are not free to use and they were going to an open-source SLAM I suppose ORB was the right choice. However, there are many other free feature descriptors that could have been used according to me. I would have liked for them to give an argument why they decided to go with ORB.

In the article the authors also measure the performance of their ORB-SLAM2. To do so they need use the process of operationalization since you can't directly observe performance. They relate performance of their SLAM system by comparing the evaluation errors with other SLAM systems' published errors on popular datasets. Then what they observe is the size of the errors. Also, to get a more accurate evaluation Mur-Artal and Tardós used their system in three datasets. They evaluated their system five times for each of the three datasets and used the median of the five scores as the evaluation score. This prevents the random error a bit. Their results were in most cases better than other state-of-the-art SLAM system.

The authors mentioned that they used an Intel Core i7-4790 with 16Gb RAM, which isn't extreme specifications for a desktop. Knowing that, together with their good results shows that their system is lightweight and works with standard CPUs. Since the back end of the system where BA is performed is part of the system, it confirms their claim that BA is lightweight.

To sum up, I think Mur-Artal and Tardós could have been clearer about their methods choice especially for using the ORB feature descriptor. That they did well was to state weaknesses of other systems and by that convince the reader that their system, using BA, was better.