Discussion for vSLAM Implementation

I think MATLAB did a nice job of implementing their monocular vSLAM application. They used tried and proven methods to perform the task, but could be improved with new advances in technology. While the ORB method for feature extraction serves its purpose, it has its weaknesses in real applications. The ORB method extracts features based on changes in color, texture, and intensity of the image. If the scene is subject to volatile lighting changes or large changes in scale, the ORB method loses some accuracy. Integrating a deep learning-based feature extraction method, like a convolutional neural network (CNN), could make the process much more robust to real-world variability. A CNN can be trained to respond to various lighting and large-scale differences. This would make the system more responsive and accurate to general changes. However, if the scene is not expected to be complex, a deep learning, method like a CNN, might be overkill since they require large amounts of data to train the model. In the end, it comes down to the specific application and the necessary accuracy for the system.

A final enhancement to consider for the implementation is the introduction of adaptive keyframe selection. Given that MATLAB's current approach extracts keyframes at a fixed rate - every 20 frames for a 30-fps video - there's a risk of selecting a blurry keyframe due to fast motion or other real-world influences. By implementing an adaptive keyframe selection, we could intelligently choose the most informative frames, much like a shopper handpicking the freshest produce at a grocery store. While the development of such a feature could be complex, it holds the potential for significant performance improvements, particularly in applications where vibrations are inherent.