# POSTER: Runtime Adaptations for Energy-Efficient VSLAM

Abdullah Khalufa, Graham Riley and Mikel Luján
*Department of Computer Science*
*The University of Manchester, United Kingdom*
{*abdullah.khalufa,graham.riley,mikel.lujan*}*@manchester.ac.uk*

## I. Introduction

Visual simultaneous localisation and mapping (VSLAM) is the process of simultaneously tracking the pose of a moving visual sensor, e.g. a camera, while mapping the observed scene. Its use in a variety of vision and robotics applications has resulted in a wide range of formulations to ensure accurate estimation for each target application. The platforms in which VSLAM is increasingly deployed are essentially battery-constrained, which requires optimising power consumption with minimal impact on the estimation accuracy. This is the main objective of this research.

Trading VSLAM performance objectives (e.g. frame rate, accuracy and power use) has been studied at design time by searching the parameters space to achieve the desired trade-off [1][2]. At VSLAM runtime, however, trading these objectives requires knowledge about the nature of sensor motion and of the scene, all of which are unknown in advance. Coupled with the fact that VSLAM formulations interpret the scene in different ways, posing a portability challenge for trade-off mechanisms.

We present a novel runtime adaptation model that relies only on the change in sensor motion as a metric in a heuristic for adapting two general and portable parameters, namely DVFS in the hardware and (redundant) frame skipping in the VSLAM formulation. The aim is to reduce power consumption with minimal impact on the tracking accuracy while maintaining a degree of portability across VSLAM formulations and the devices they execute on. The adaptation model is responsive to sudden changes in the sensor motion which can lead to a major impact on the accuracy and robustness of VSLAM. We evaluate the model on two prominent keyframe-based formulations, namely ORB-SLAM [3] and DSO [4].

## II. Runtime Adaptations

The main goal of the adaptation model is to reduce power consumption with minimal impact on VSLAM tracking accuracy. To achieve this, the change in sensor motion can be used as a metric for characterising the tracking difficulty, assuming there is sufficient scene information. For example, in the case of a sensor mounted on a quad-copter, a sudden and large change in motion can lead to a significant impact on the tracking accuracy due to lack of adequate tracking information being observed by the sensor. Whereas with slow and steady motion, there are opportunities to save power with minimal impact on accuracy by taking advantage of redundancy in the tracked information and skipping frames. We define the sensor motion change $D$ at time $t$ between two subsequent frames as:

$$D_t = \Delta_p^\top \Delta_p \tag{1}$$

where $\Delta_p$ is the absolute difference between sensor pose at time $t$ and $t-1$. To cope with unknown levels of variations in the sensor motion $D$, we keep a record of its minimum and maximum values, which are then used to normalise the current change in motion $D_t$:

$$D'_t = \frac{D_t - min(D)}{max(D) - min(D)} \tag{2}$$

The normalised value $D'_t$ is then used to set the appropriate value of the adapted parameter $X$ within a predefined range ($X_{min}$ to $X_{max}$) based on the correlation between the parameter and the sensor motion:

$$X_p = X_{min} + (X_{max} - X_{min})D'_t \tag{3}$$

$$X_n = X_{max} - (X_{max} - X_{min})D'_t \tag{4}$$

Where $X_p$ is used when the correlation is positive, while $X_n$ is used for negative correlation. Equation 3 is used for DVFS adaptations (F) while equation 4 is used to determine the number of redundant frames that may be skipped in a row (S).

The model is light-weight and does not require major changes to the VSLAM implementation. It also ensures that when the change in sensor motion is at its peak ($D_t = max(D)$), the adapted parameters are quickly set to achieve high accuracy and robustness, while when the change is at its minimum, the parameters are set to achieve the highest power savings.

## III. Experimental Evaluation

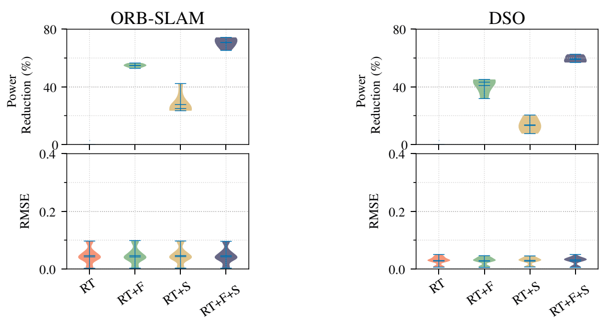
We evaluate the adaptation model on the open-source real-time monocular version of ORB-SLAM[1] and DSO[2] (for which the real-time execution performance results (RT) are used as the baseline for comparison) running on scenes from

[1]https://github.com/raulmur/ORB_SLAM2
[2]https://github.com/JakobEngel/dso

978-1-7281-3613-4/19/$31.00 ©2019 IEEE
DOI 10.1109/PACT.2019.00052

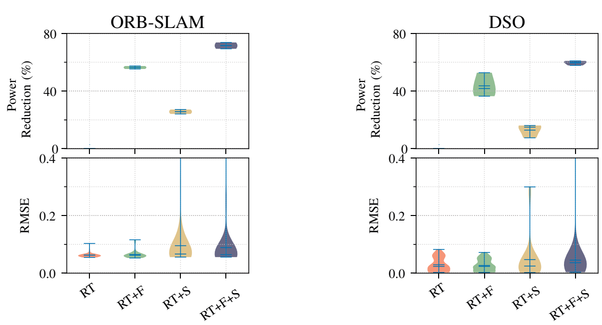
the EuRoC MAV [5] and ICL-NUIM [6] datasets. However, we exclude scenes where any of the baselines do not perform well in terms of tracking accuracy or robustness since the goal is to evaluate the adaptations not the performance of the baseline version. A total of 20 runs is performed with each scene. For each run the estimated track is scaled and aligned with the ground truth trajectory, then the Root Mean Square Error (RMSE) from the ground truth is calculated. Low variability in the 20 RMSE values implies high robustness.

Fig. 1 shows the effectiveness of the adaptation model on a number of scenes from both datasets, where the adapted parameters, (F) or (S), achieve similar RMSE values to the baseline (RT) with power reduction, relative to the baseline, up to 75% on ORB-SLAM and 64% on DSO.




(a) Scenes: MH1, MH2, MH3, V101, V201, LR0. (b) Scenes: MH2, V201, LR1, LR2, OF3.

Figure 1. Violin plots of Power Reduction (top) with marginal impact on accuracy (bottom) relative to each implementation baseline "RT" along with the proposed runtime adaptation (x-axis).




(a) Scenes: V102, V202, LR0. (b) Scenes: MH1, LR0, LR3.

Figure 2. Violin plots of Power Reduction (top) with impacted robustness (bottom) relative to each implementation baseline "RT".

Fig. 2, however, shows sequences where the use of the adaptation model incurred an impact, mainly on the robustness of the two VSLAM formulations. This impact can be attributed to *critical points* encountered during the tracking where the metric used is unable to identify and therefore respond to [7]. Fig. 3 shows an example of the occurrence of two critical points. This issue is more prevalent when redundant frames are dynamically skipped (S) where the

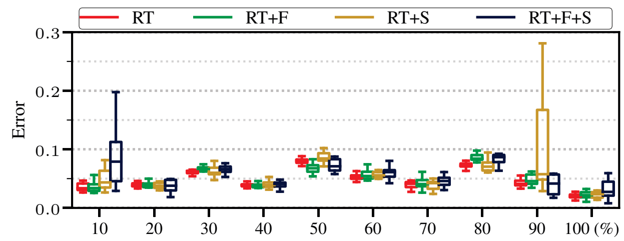



Figure 3. ORB-SLAM running on scene V202. Errors distribution for 20 runs recorded at 10 intervals over the run duration. Two critical points can be seen where the robustness is highly impacted, at 10% for (RT+F+S), and at 90% for (RT+S). ORB-SLAM recovers due to its loop closure ability.

change in sensor motion may not reflect the level of tracking and scene difficulty. This motivates the search for improved metrics for identifying specific types of motions or scene complexities that give rise to such critical points.

## Acknowledgment

Abdullah Khalufa is sponsored by Umm Al-Qura University, Mecca, Saudi Arabia. We also acknowledge the support of EPSRC grants RAIN Hub (EP/R026084/1) and PAMELA (EP/K008730/1). Mikel Luján is supported by an Arm/RAEng Research Chair appointment.

## References

[1] B. Bodin, L. Nardi, M. Z. Zia, H. Wagstaff, G. S. Shenoy, M. Emani, J. Mawer, C. Kotselidis, A. Nisbet, M. Lujan, B. Franke, P. H. J. Kelly, and M. O'Boyle, "Integrating algorithmic parameters into benchmarking and design space exploration in 3D scene understanding," in *2016 International Conference on Parallel Architecture and Compilation Techniques (PACT)*, sep 2016, pp. 57–69.

[2] M. Z. Zia, L. Nardi, A. Jack, E. Vespa, B. Bodin, P. H. Kelly, and A. J. Davison, "Comparative design space exploration of dense and semi-dense SLAM," *Proceedings - IEEE International Conference on Robotics and Automation*, vol. 2016-June, pp. 1292–1299, 2016.

[3] R. Mur-Artal and J. D. Tardós, "ORB-SLAM2: An Open-Source SLAM System for Monocular, Stereo, and RGB-D Cameras," *IEEE Transactions on Robotics*, vol. 33, no. 5, pp. 1255–1262, 2017.

[4] J. Engel, V. Koltun, and D. Cremers, "Direct Sparse Odometry," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 40, no. 3, pp. 611–625, 2018.

[5] 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, jan 2016. [Online]. Available: https://doi.org/10.1177/0278364915620033

[6] A. Handa, T. Whelan, J. McDonald, and A. J. Davison, "A benchmark for RGB-D visual odometry, 3D reconstruction and SLAM," in *2014 IEEE International Conference on Robotics and Automation (ICRA)*, 2014, pp. 1524–1531.

[7] A. Khalufa, G. Riley, and M. Luján, "A Dynamic Adaptation Strategy for Energy-Efficient Keyframe-Based Visual SLAM," in *The 25th International Conference on Parallel and Distributed Processing Techniques and Applications*, July 2019.