A High-Performance and Energy-Efficient Accelerator with the RISC-V Core for Optimization in Visual SLAM System

Speaker : 刘檬　Dr. Meng Liu

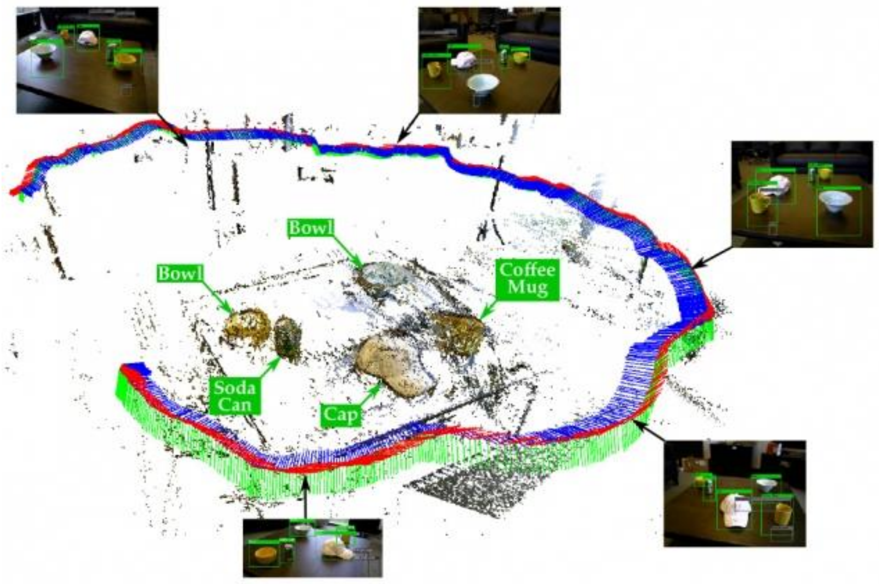
Inventors: 李任伟，吴军宁，刘檬，周沈刚，陈祖玎

Organization : Institute of Automation, Chinese Academy of Sciences

Beijing Haawking Technology Co., Ltd

单位：中科院自动化所；北京中科昊芯科技有限公司

SLAM

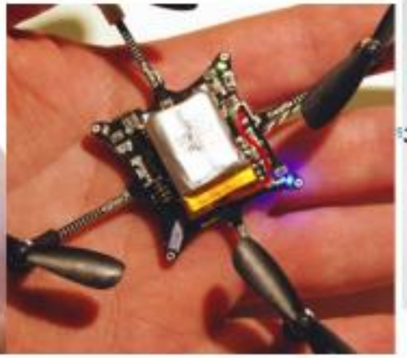


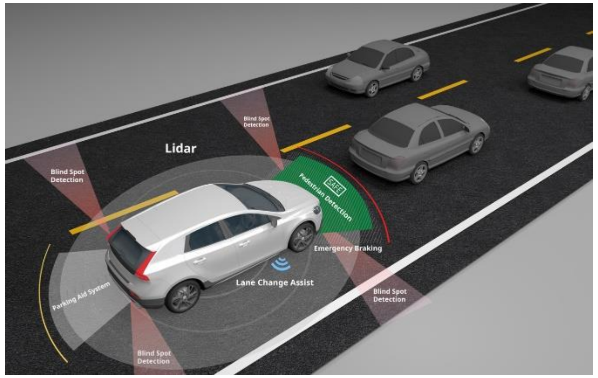
- Simultaneous Localization and Mapping
- Estimate the pose of a robot and the map of the environment at the same time , this means:
  - Map the location, creating a 3D virtual map
  - Locate itself inside the map
  - In sum, it is a ***hot topic***.

- Visual SLAM (vSLAM)
  - Computer Vision
  - Sensor is camera
  - One of the most ***challenging*** open problems

Key Applications of Visual SLAM







- Low-cost robotics (e.g. a mobile robot with a cheap camera)
- Agile robotics (e.g. drones)
- Smartphones
- Wearables
- AR/VR: inside-out tracking, gaming

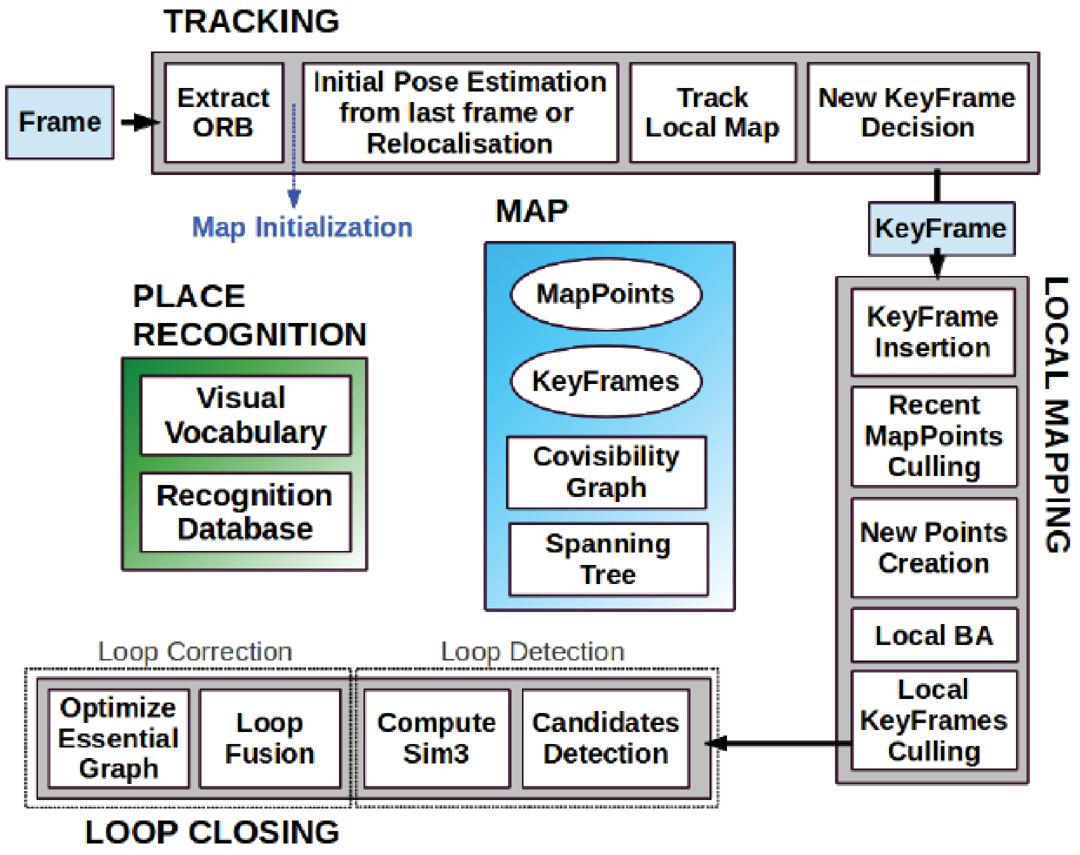
The Need for Efficiency in Advanced Real-Time Vision

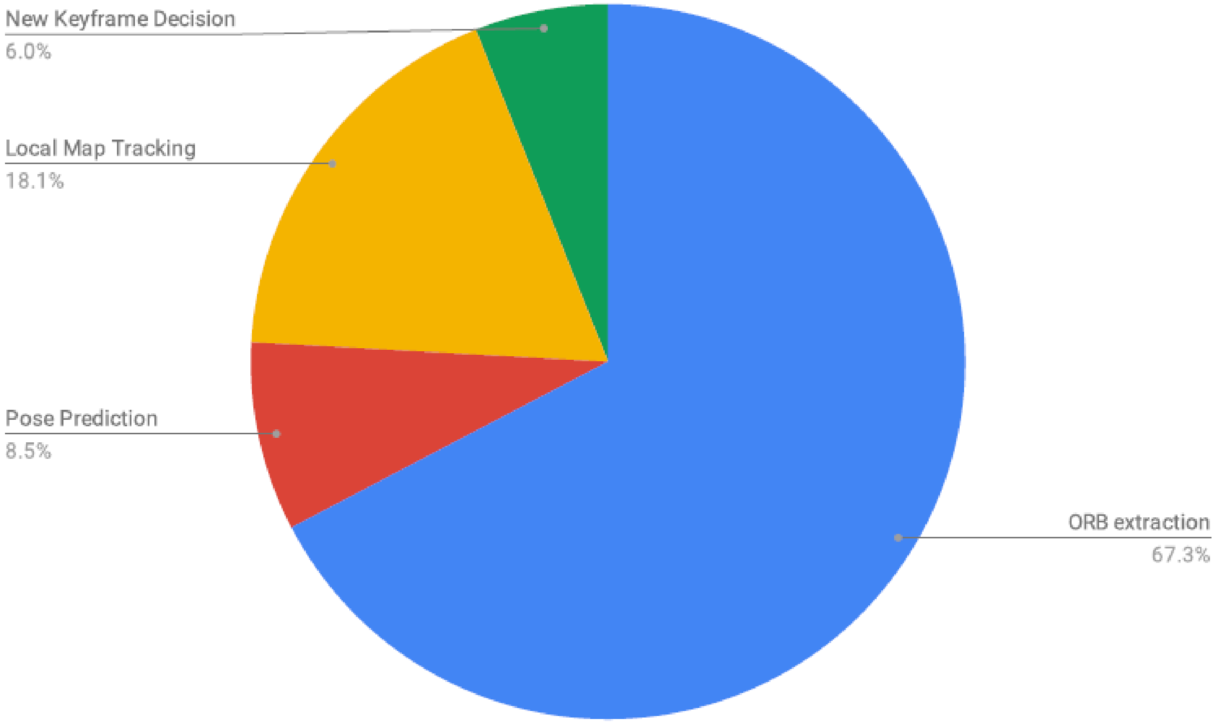
- We need 1000x power efficiency for truly capable always-on tiny devices; or to do much more with larger devices.

Trends

- Sensors, algorithms and processors together
- ASIC

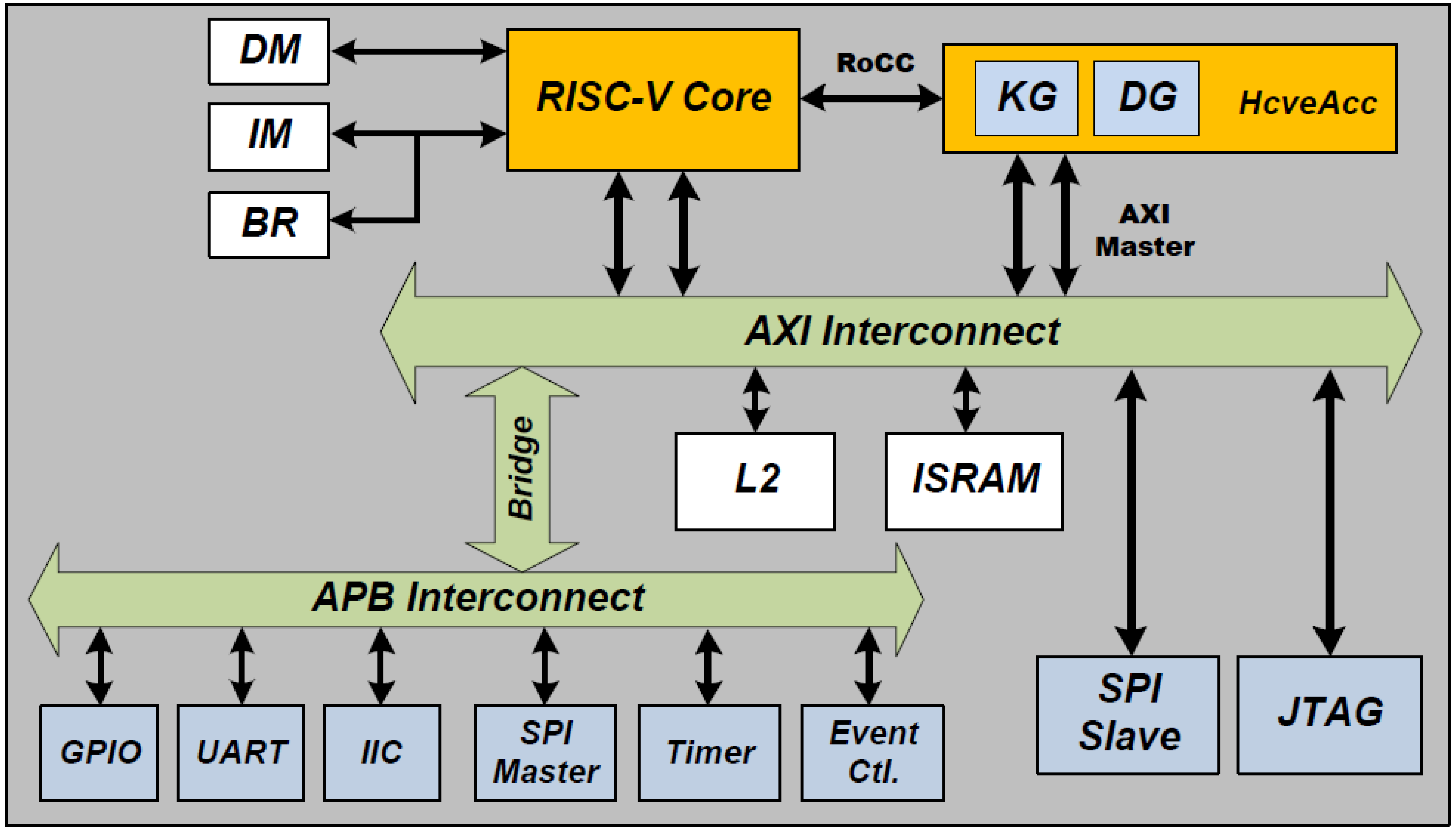
Aim of our work





- ORB-SLAM（open-source, high-ranking）
- ASIC way, providing comparable performance and accuracy (CPU； FPGA)
- Of course, we should save power cost （within 200mW）
- Currently tracking task

Framework (SoC)



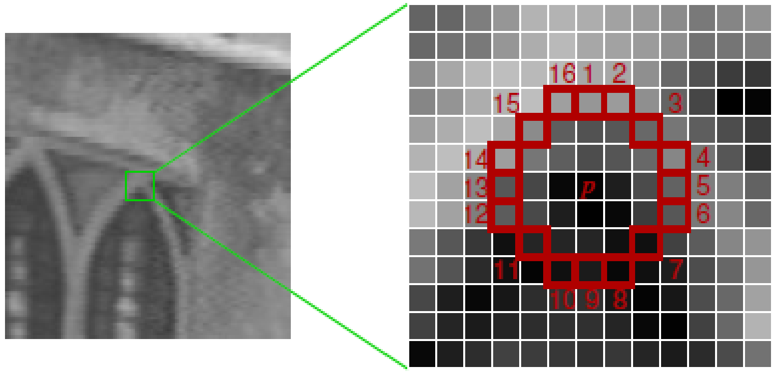
- RISC-V Core(custom design)
- HcveAcc (co-processor)
- Task pipelined
- Maximize the use of bandwidth resources

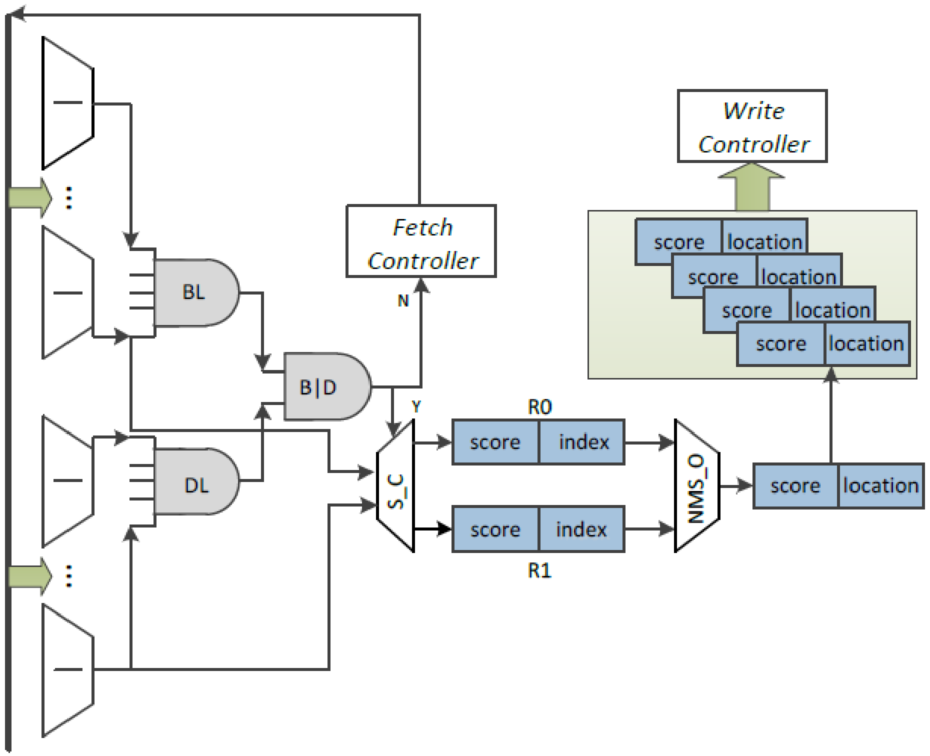
Details of Architecture

- FAST feature detection

$$S_{p\to x} = \begin{cases} d, & I_{p\to x} \le I_p - t \ (darker) \\ s, & I_p - t \le I_{p\to x} \le I_p + t \ (similar) \\ b, & I_p + t \le I_{p\to x} \ (brighter) \end{cases}$$

$$V = \max\left(\sum_{x \in S_{bright}} |I_{p\to x} - I_p| - t, \ \sum_{x \in S_{dark}} |I_p - I_{p\to x}| - t \right)$$





Details of Architecture

- Pixel Cache Design

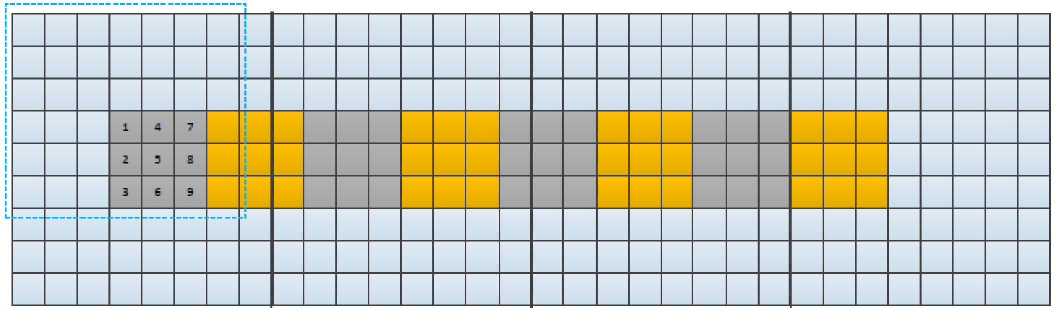


Figure 7: The Window Fusion module example.

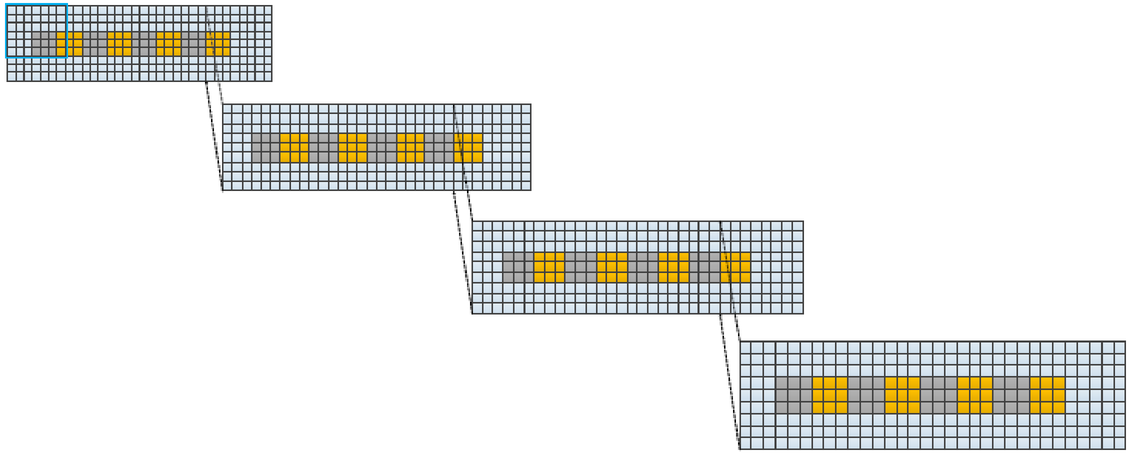


Figure 8: The example of Several Window Fusion modules with horizontal data splicing.

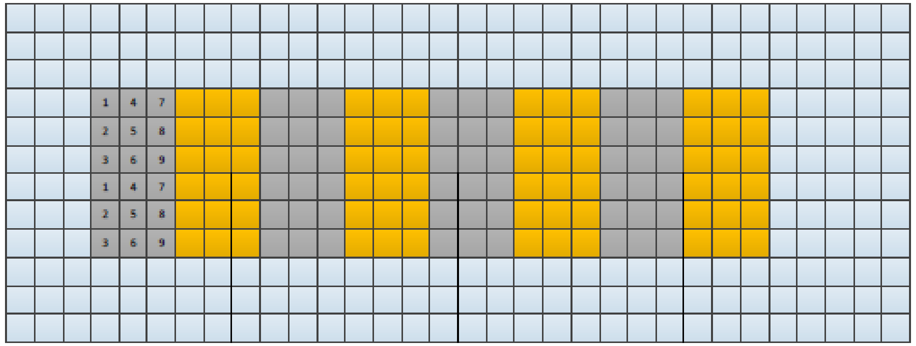
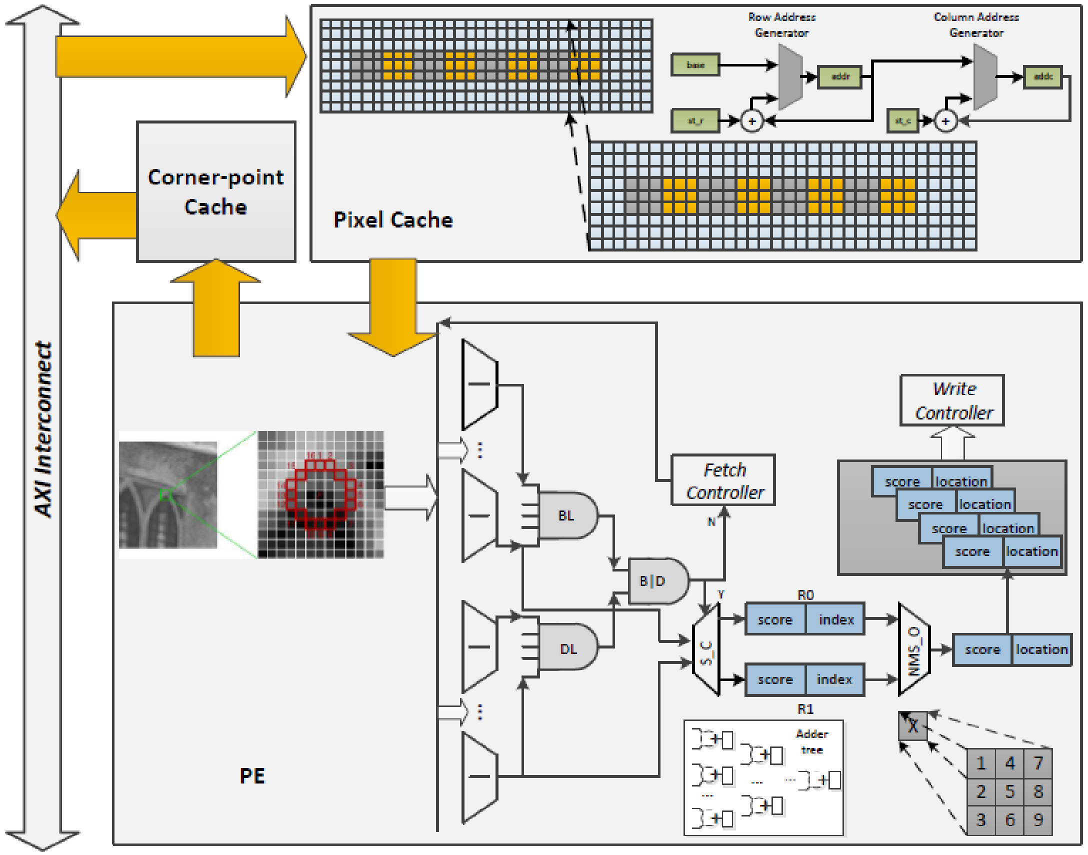


Figure 9: The example of two Window Fusion modules with vertical data splicing.

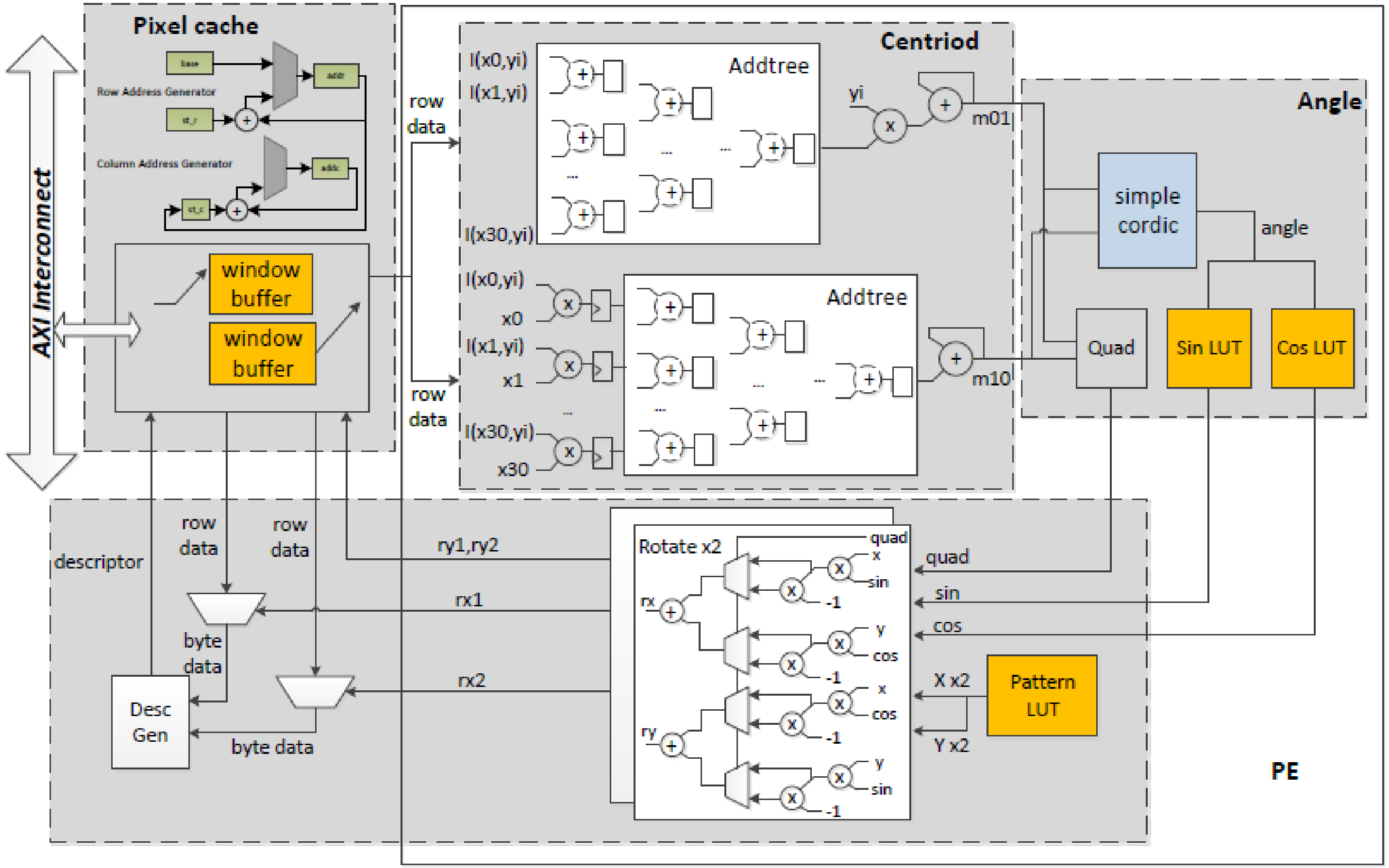
KG Architecture

- KG （Key-point Generation） module design



DG Architecture

- DG (Descriptor Generation) module design



$$\begin{cases} \widetilde{x_{i+1}} = x_i + y_i tan\theta \\ \widetilde{y_{i+1}} = y_i - x_i tan\theta \\ \theta_{i+1} = \theta_i + \theta \end{cases}$$

$$\begin{cases} \widetilde{x_{i+1}} = x_i + y_i \times 2^{-i} \\ \widetilde{y_{i+1}} = y_i - x_i \times 2^{-i} \\ \theta_{i+1} = \theta_i + tan^{-1}\left(2^{-i}\right) \end{cases}$$

- ➢ The rBRIEF algorithm
- ➢ CORDIC calculation

Experimental Results

TABLE II: The comparison of performance and power results.

Platform	*Image Resolution*	*Frame Latency (ms)*	*Throughput (MPix/s)*	*Speedup Ratio*	*Power (W)*
CPU [3]	640*480	32.5	9.5	15.4	47
ARM [3]	640*480	291.6	1.1	133.2	1.574
FPGA [4]	640*480	14.8	20.1	7.3	4.6
eSLAM [3]	640*480	9.1	33.8	4.3	1.936
HcveAcc	752*480	2.3	146.5	N/A	0.181

TABLE I: Result comparisons between FAST-BRIEF and HcveAcc.

Design Name		FAST-BRIEF [5]		HcveAcc	
CMOS Tech.		65nm		28nm	
Components		Detector	Descriptor	Detector	Descriptor
Quantization		N/A	*256bits (N)*	FAST12	*256bits (Y)*
Number of PEs		N/A		2	2
Area cost (sq um)	*Comb*	10146.4	11923.9	2837.6	13345.6
	Seq	9066.3	42154.9	968.1	7175.4
	Total	73291.5		24326.7	
Freq.(GHz)		N/A		1.0	
Power(mW)		1131		50	

- Power consumption much saved
- Area saved up to 66.8%

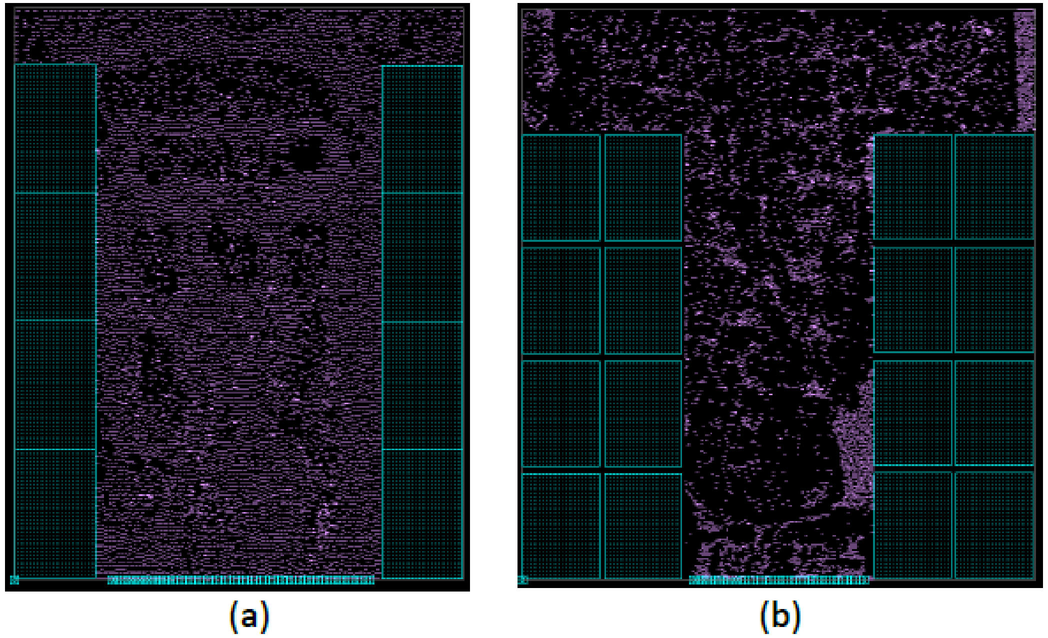
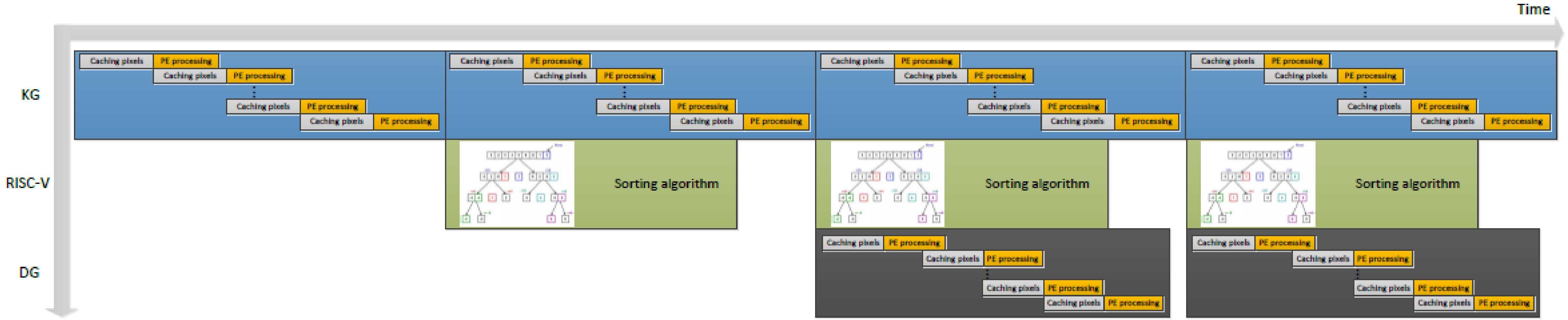


Fig. 5: (a) The layout design of KG; (b) The layout design of DG.

Experimental Results



- KG, DG, and RISC-V core
- Top-k ranking problem
- The pipelined working mode gives the best results.
- Total runtime results of case1, case2 and case3 are 5ms, 3.7ms and 2.3 ms respectively.

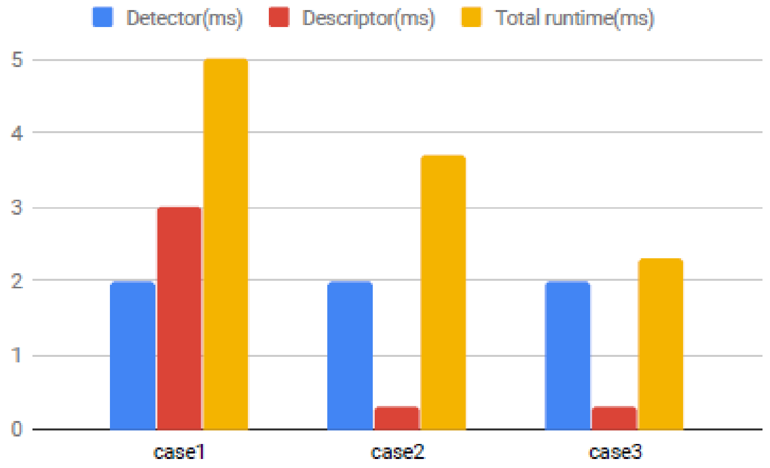


Fig. 6: The runtime comparisons of case1, case2 and case3. The total runtime colored with yellow includes the combined runtime of Detector(KG), Descriptor(DG), and the RISC-V kernel.

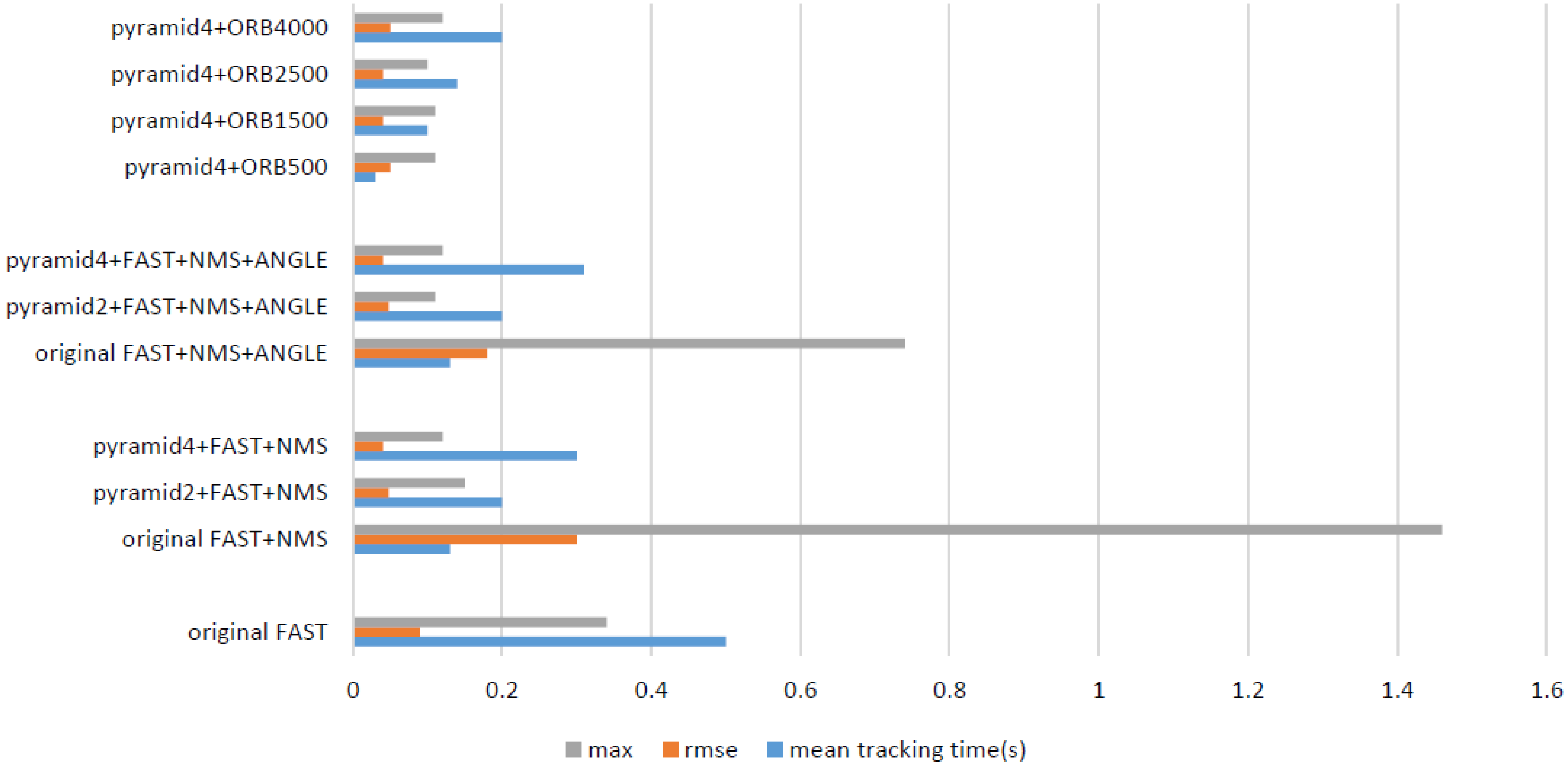


Figure 12: Comparison results of different approaches for KITTI EuRoC MH_03_medium.

激

Experimental Results

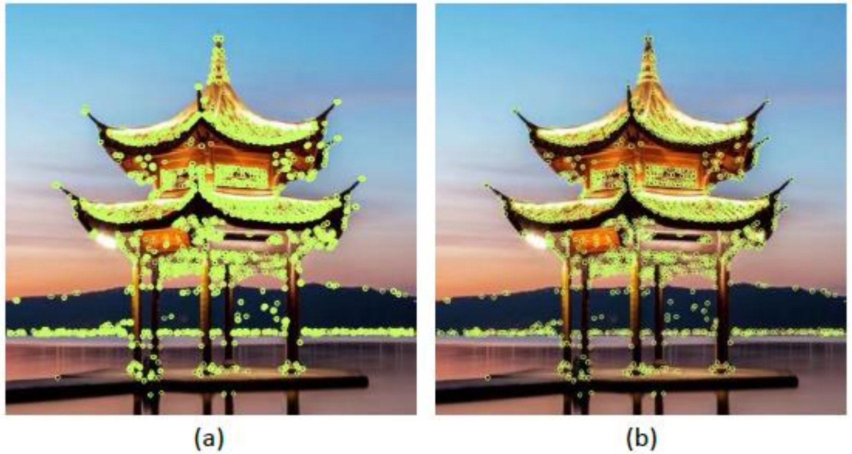


Fig. 7: The comparison of feature detection with or without NMS module. All of key-points are marked with green color.

- NMS mode can help control the feature size
- One-degree discretization
- 3X of 12 degrees discretization

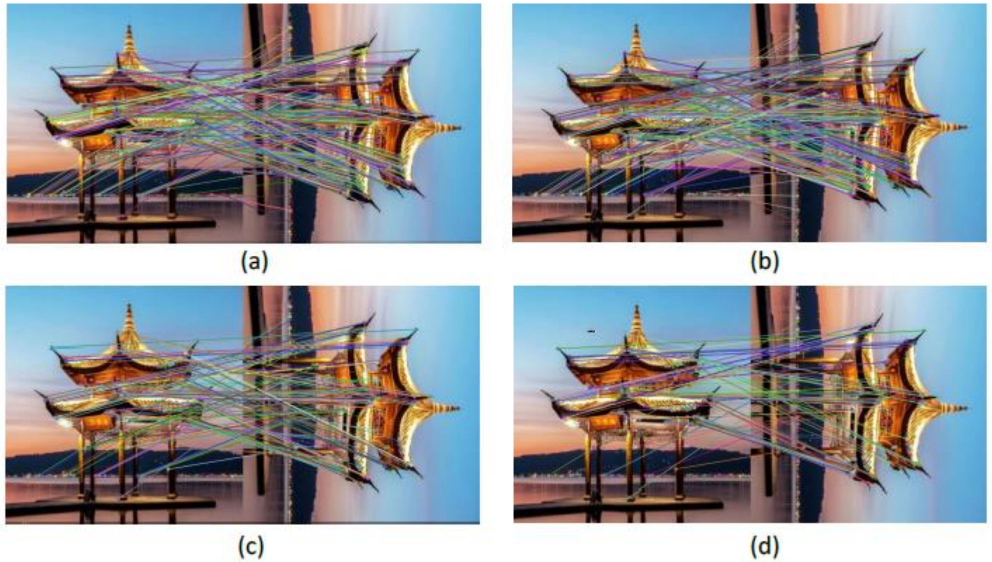


Fig. 8: The matching graphs with using different precision of angle calculation.

TABLE III: The comparison of hardware utilization.

Angle Discretization	*Max Distance*	*Total*	*Matched*	*Error Rate*
20	76	500	149	70.20%
12	60	500	278	44.40%
5	5	500	493	1.40%
1	4	500	498	0.40%

Experimental Results

- ORB-SLAM accuracy
- The EuRoC and TUM dataset
- Absolute translation root-mean-square error (RMSE) metric
- The ground truth trajectory and the estimated trajectory
- Results of RMSE are acceptable and the max error can be controlled within 1m (4.9% worst)

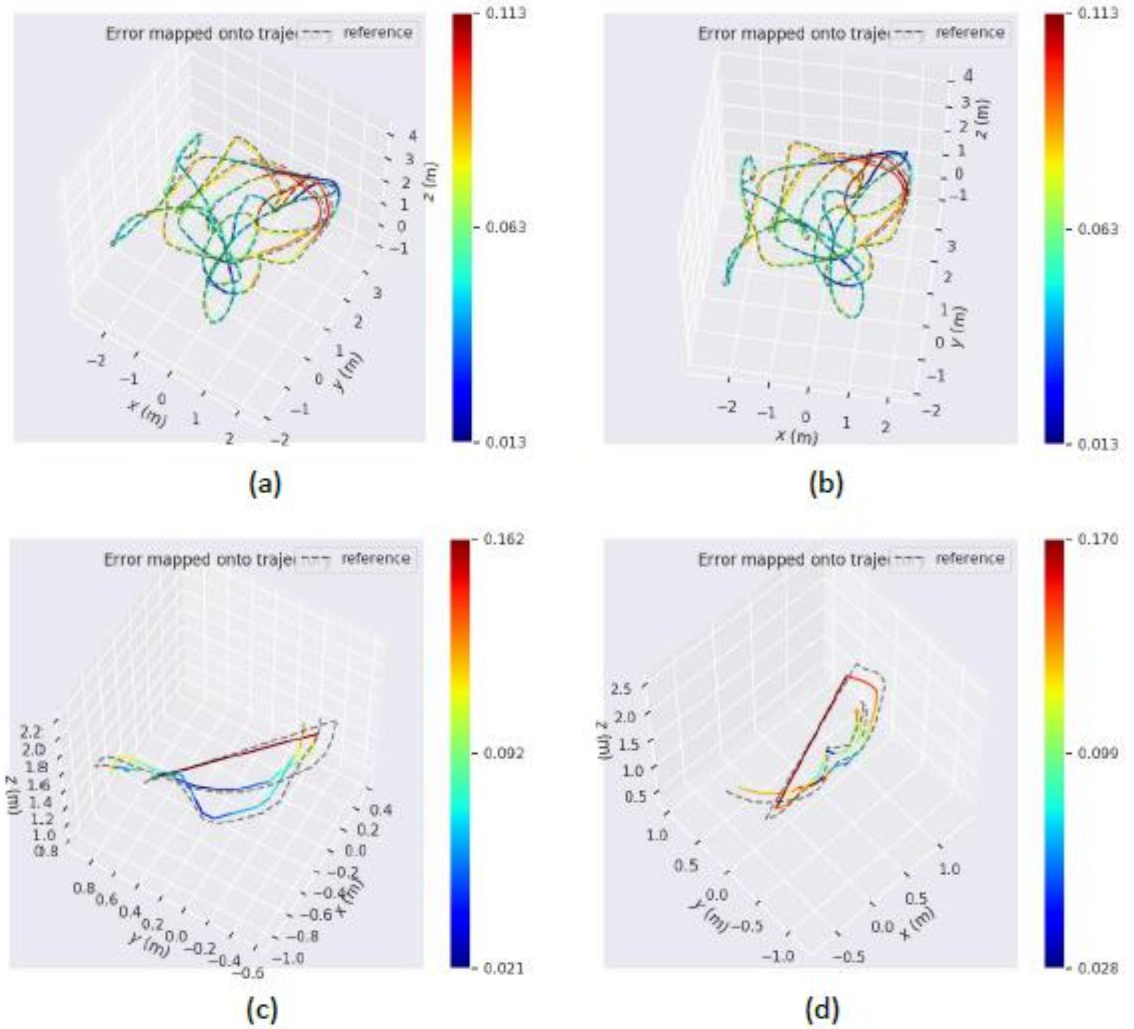


Fig. 9: (a) The original ORB-SLAM trajectory of EuRoC V1_02_medium ($RMSE = 0.064, max = 0.113(m)$) ; (b) The modified ORB-SLAM trajectory of EuRoC V1_02_medium ($RMSE = 0.069, max = 0.498(m)$) ; (c) The original ORB-SLAM trajectory of TUM fr1/room ($RMSE = 0.090, max = 0.162(m)$) ; (d) The modified ORB-SLAM trajectory of TUM fr1/room ($RMSE = 0.118, max = 0.170(m)$).

中国科学院自动化研究所
INSTITUTE OF AUTOMATION CHINESE ACADEMY OF SCIENCES

中科昊芯 HAAWKING

Thank you!

liumeng2013@ia.ac.cn