Imperial College London

Auto-tuning for algorithmic and implementation design choices in GPU-accelerated 3D robot vision

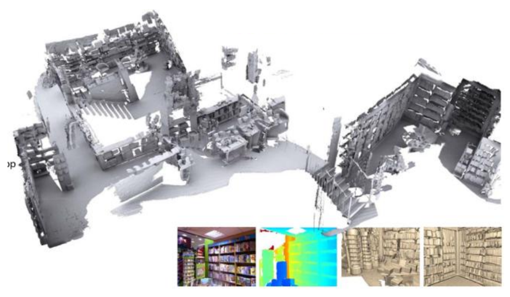
Luigi Nardi (l.nardi@imperial.ac.uk) and collaborators on the PAMELA project

http://apt.cs.manchester.ac.uk/projects/PAMELA/

1. INTRODUCTION: THE THREE R's OF COMPUTER VISION

Reconstruction

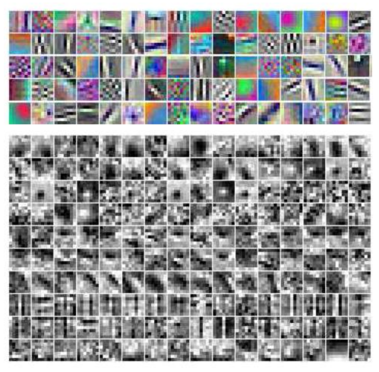
Scalable Kinect Fusion (2013)

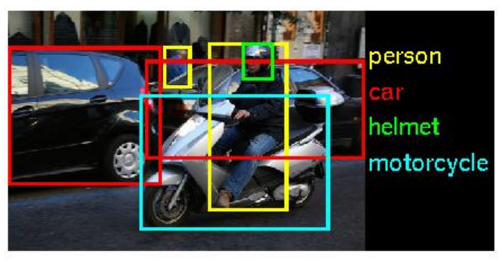


Building Rome on a cloudless day (2010)

Recognition

Deep learning for scalable Object class detection (2014)



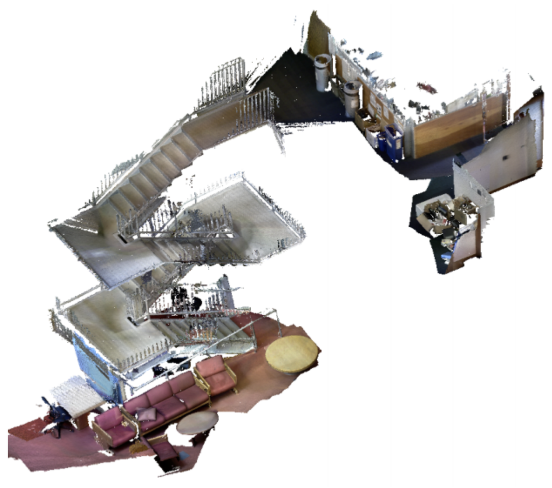


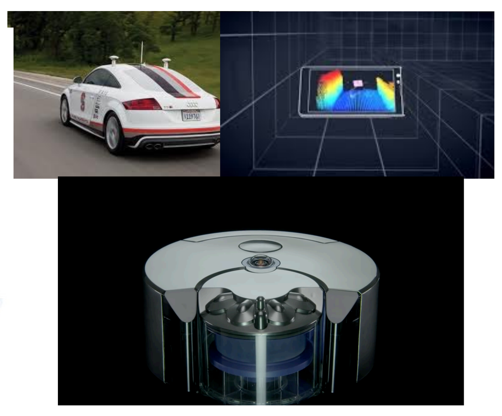
Reorganisation or Grouping

Contour detection and segmentation (2011)

2. BACKGROUND: SIMULTANEOUS LOCALISATION AND MAPPING (SLAM)

We focus on the first "R": Reconstruction; more precisely in SLAM. SLAM builds a coherent world representation and localises the camera in real-time.





SLAMBench FRAMEWORK

Holistic approach to SLAM "performance": SLAMBench. SLAMBench is a publicly-available benchmarking framework for quantitative, comparable and validatable experimental research to investigate trade-offs in performance, accuracy and energy consumption of a SLAM system.

The accuracy is measured as Absolute Trajectory Error (ATE) in cm.

SLAM benchmarks
- Dense SLAM: KinectFusion …
- Semi-dense SLAM: LSD-SLAM …
- Sparse SLAM: ORB-SLAM …

Implementation languages
C++ | OpenMP | OpenCL | CUDA | SYCL | PENCIL | …

Desktop to embedded platforms
ARM | Intel | NVIDIA | …

Datasets
ICL-NUIM | TUM RGB-D | …

Performance evaluation
Frame rate | Energy | Accuracy

3. GOAL: MULTI-OBJECTIVE CO-DESIGN SPACE EXPLORATION

Co-design space

Space 1 — Algorithmic:
- Application-specific parameters
- Minimisation methods
- Early exit condition values

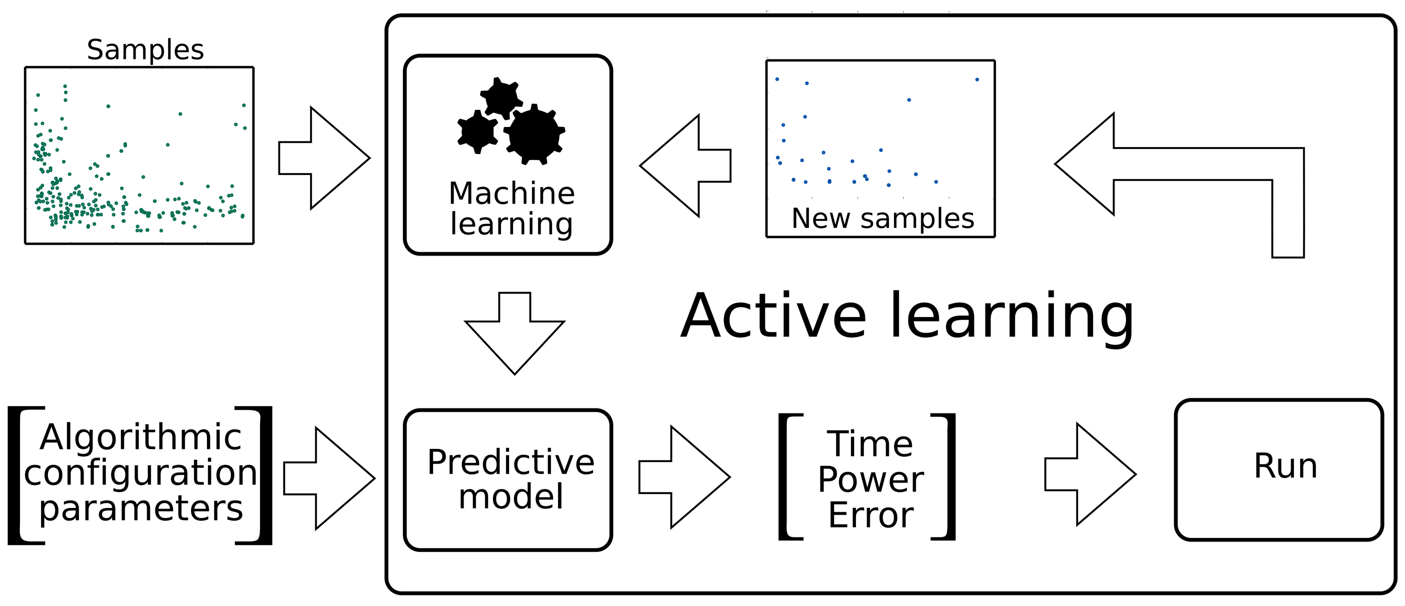
Space 2 — Compilation:
- opencl-params: -cl-mad-enable,-cl-fast-relaxed-math, etc.
- LLVM flags: O1, O2, O3, vectorize-slp-aggressive, etc.
- Local work group size: 16/32/64/96/112/128/256
- Vectorisation: width (1/2/4/8), direction (x/y)

Space 3 — Architecture:
- GPU frequency: 177/266/350/420/480/543/600/DVFS
- # of active big cores: 0/1/2/3/4
- # of active LITTLE cores: 1/2/3/4

ACKNOWLEDGEMENTS

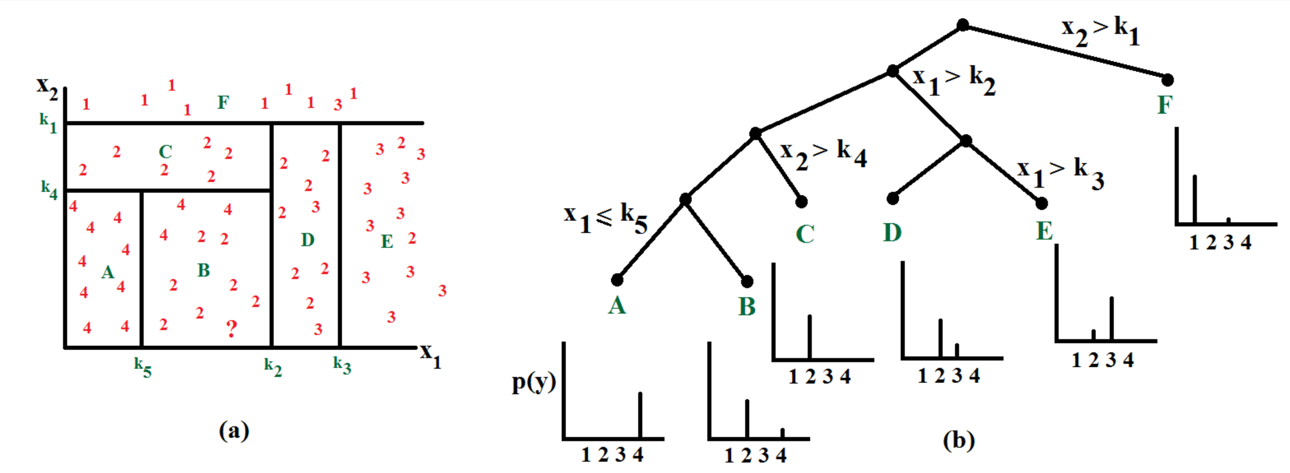
We acknowledge funding by the EPSRC grant PAMELA EP/K008730/1 and thank the PAMELA Steering Group for discussions.

4. METHOD: RANDOM FOREST AND ACTIVE LEARNING REGRESSION

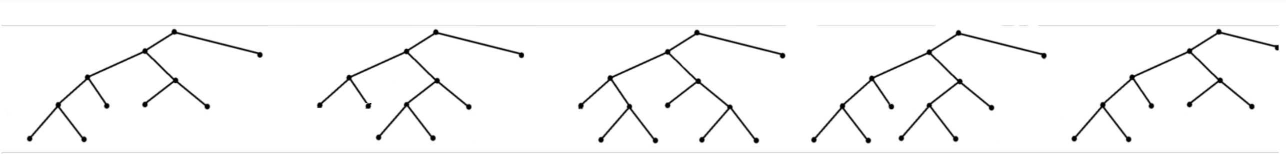


PREDICTIVE MODEL

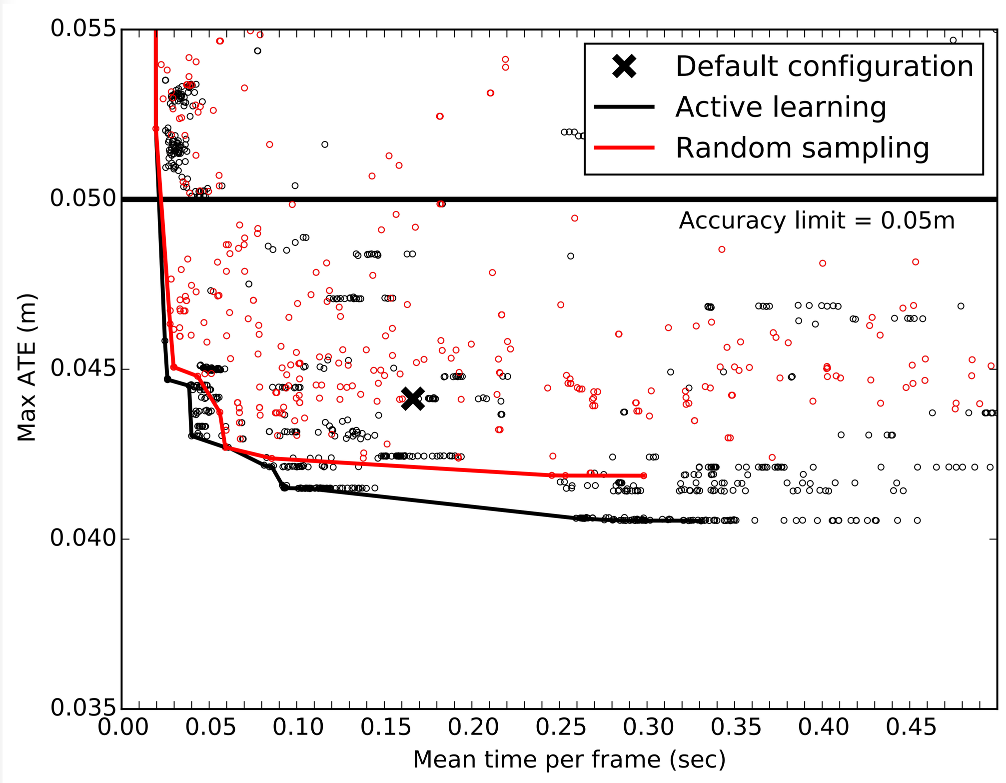
Decision tree

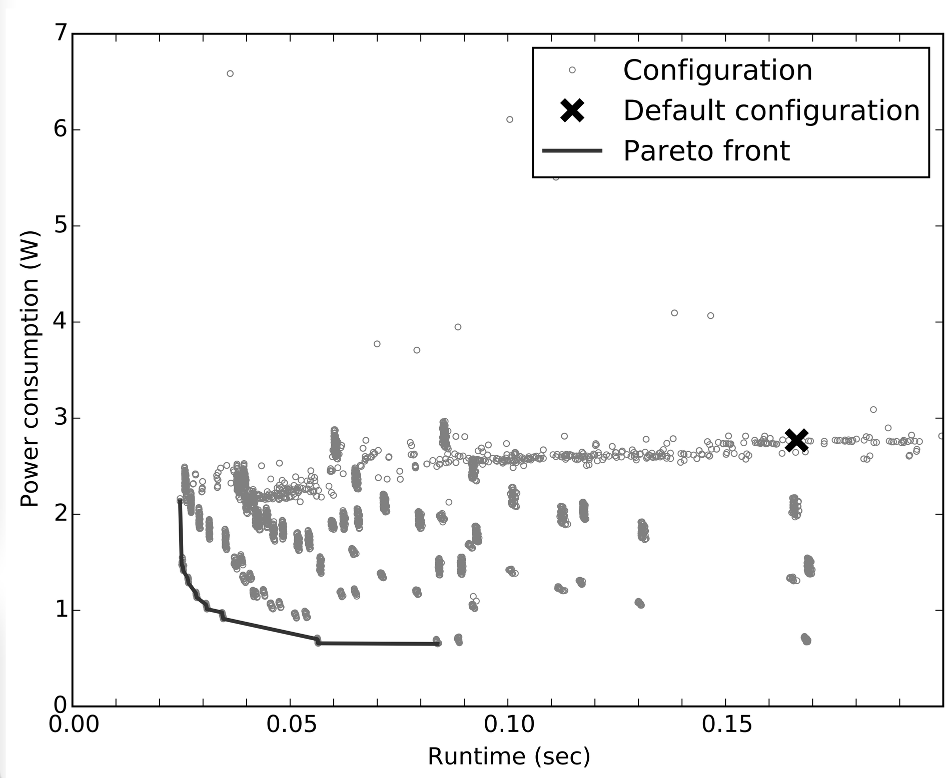


Random forest



5. RESULTS: PARETO SURFACE CONTAINING ALL OPTIMA FOUND





Embedded system								
Machine	CPU	CPU name	CPU GFLOPS	CPU cores	GPU	GPU name	GPU GFLOPS	TDP Watts
Hardkernel ODROID-XU3	ARM A15 + A7	Exynos 5422	80	4 + 4	ARM	Mali-T628	60 + 30	10

SELECTED PARETO POINTS

Constraint	Runtime (FPS)	Max ATE (cm)	Power (Watts)
Default	6.03	4.41	2.77
Best runtime	39.85	4.47	1.47
Best accuracy	1.51	3.30	2.38
Best power	11.92	4.45	0.65
Power < 1W	29.09	4.47	0.98
Power < 2W	39.85	4.47	1.47
FPS > 10	11.92	4.45	0.65
FPS > 20	28.87	4.47	0.91
FPS > 30	32.38	4.47	1.01

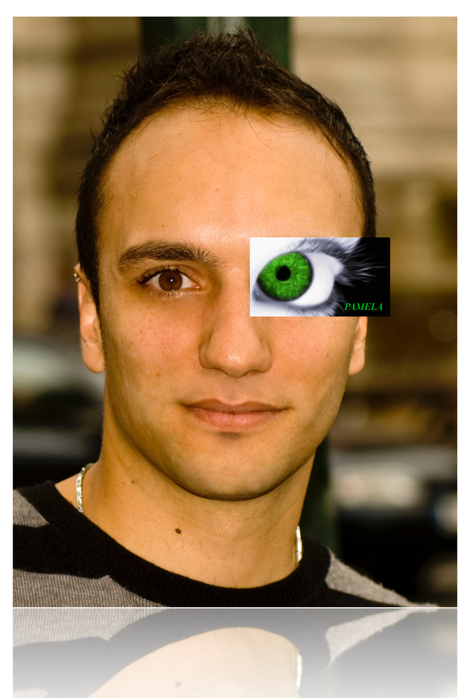
6. CONCLUSION

- Multi-objective machine learning driven optimisation framework on frame rate/power/accuracy brings us to find much better configurations than the default configuration.
- Application accuracy check very powerful: non bit-wise and scope for aggressive approximate computing.

REFERENCES

- M. Z. Zia, et al. Comparative Design Space Exploration of Dense and Semi-Dense SLAM. In IEEE ICRA, Stockholm, Sweden, May 2016.
- L. Nardi, et al. Introducing SLAMBench, a performance and accuracy benchmarking methodology for SLAM. In IEEE ICRA, Seattle, Washington USA, May 2015.

BIO: Luigi Nardi, PhD



Luigi Nardi is a Post-Doctoral Research Associate at Imperial College London in the Software Performance Optimisation group. Luigi's primary role is to work in the co-design of high-performance low-power computer vision systems where performance, power and accuracy are part of the same optimisation space. Prior to his current position, Luigi earned his Ph.D. in computer science at Pierre et Marie Curie University and was a permanent researcher leading the high-performance computing effort at Murex S.A.S.