Map Slammer:

Densifiying Scattered KSLAM 3D maps with Deep Learning

Author: José Miguel Torres-Cámara
Mentors: Miguel Ángel Cazorla & Francisco Gómez

End-of-degree thesis
Robotics Engineering, University of Alicante

July 2019

RoViT

Motivation

- Localisation, mapping, navigation and recognition → 3D representations.
- 3D sensors are big, heavy, power-hungry and sensitive to high distances and sun-light.

Source: Wikipedia and Reality Technologies

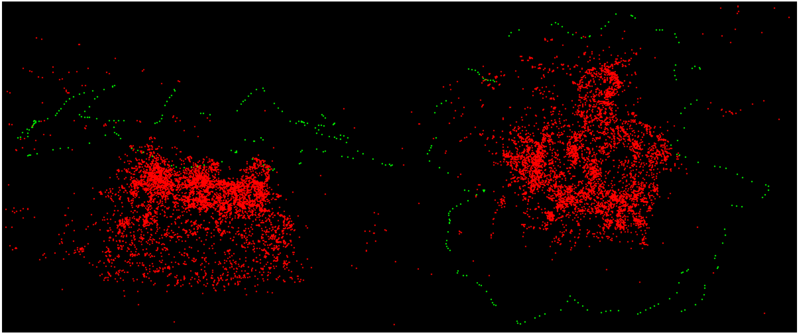
Proposal

- Generate dense 3D representations using monocular data.
- Fuse KSLAM scattered output with DL-estimated depth.
- Demonstrate the performance with qualitative and quantitative results.
  - Localisation accuracy
  - Densification capabilities
  - Accuracy of the resulting map

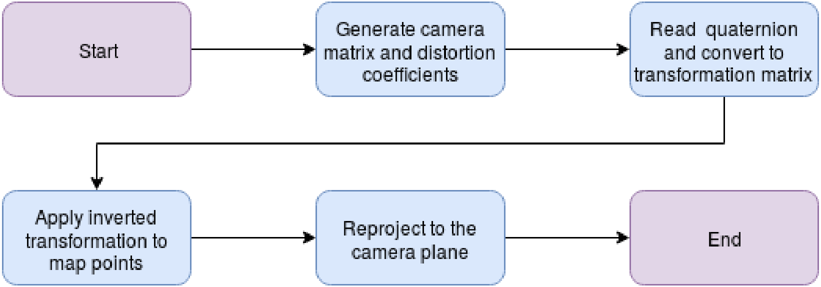
General pipeline

- Generate 3D precise sparse points (persistent scale).
- Reproject the map points (extracted keypoints $\neq$ map points).
- Depthmaps estimation (dense and sparse sub-set).
- Transformation computing and applying.
- Refinement (Voxel Grid + SOR).

Support points obtaining - ORB-SLAM2 output

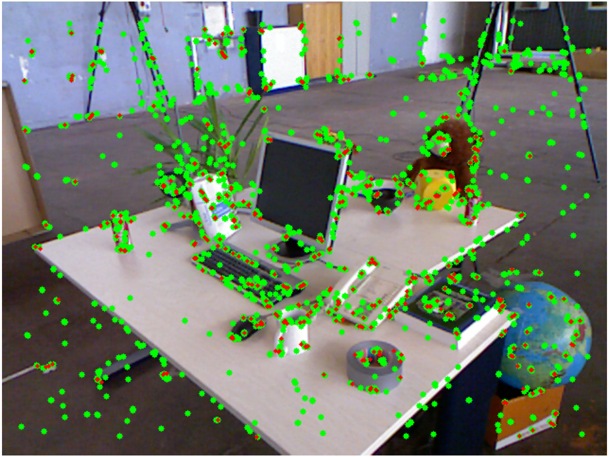


Keypoints reprojection - Flow chart

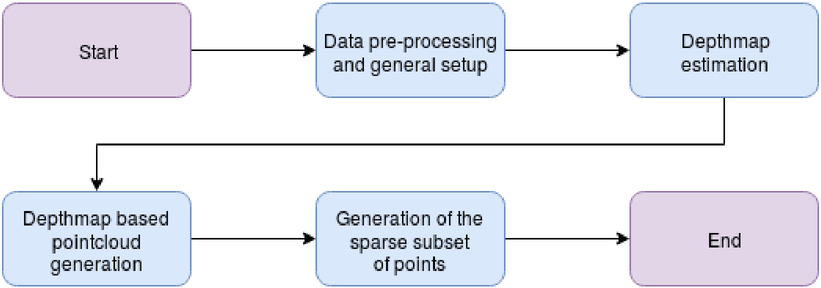


RoViT

Keypoints reprojection - Visualisation

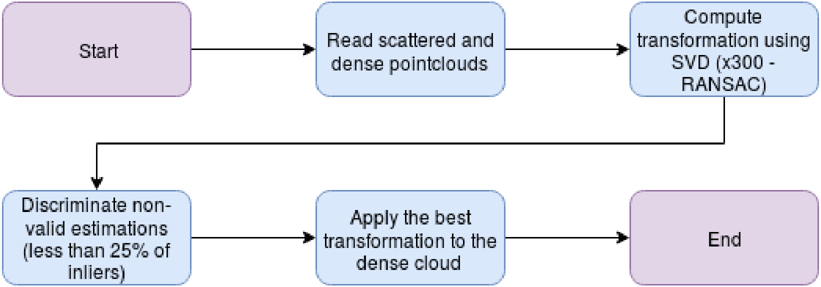


Depth estimation - Flow chart

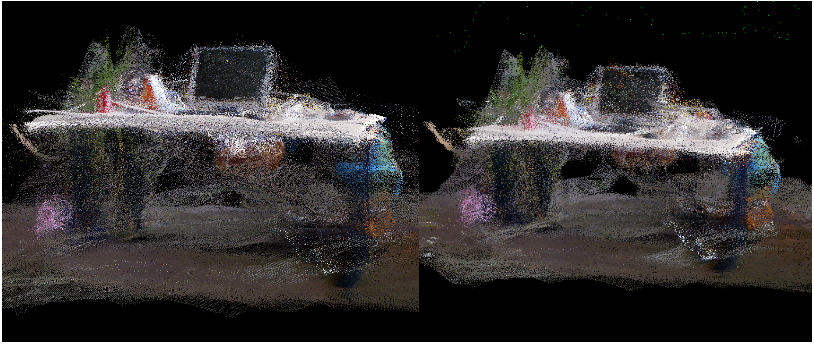


Support points obtaining - DeMoN output

Transformation obtaining and applying - Flow chart



Refinement - Before and after



Localisation accuracy

	ATE (O-S2)	RPE (O-S2)	ATE (Map Slammer)	RPE (Map Slammer)
fr1/xyz	0.00813	0.01636	0.03230	0.03230
fr1/desk	0.01596	0.01896	0.10480	0.21978
fr1/desk2	0.02228	0.02844	0.11920	0.29097
fr2/xyz	0.00244	0.00143	0.07777	0.04896
fr2/desk	0.00913	0.00621	0.11522	0.16180
Average	0.01159	0.01428	0.08986	0.15076
Variance	0.00006	0.00011	0.00130	0.01224

ORB-SLAM2 locates properly, but extra transformations are needed.

Densification capabilites

	ORB-SLAM2	Map Slammer
fr1/xyz	315.156250	49152
fr1/desk	265.337838	49152
fr1/desk2	202.824324	49152
fr2/xyz	284.484848	49152
fr2/desk	270.320988	49152
Average	250.537375	49152
Variance	3102.91105	0

Densification of almost 200 times is performed.

Accuracy of the output map

	ORB-SLAM2	Map Slammer
fr1/xyz	0.086565	0.0310728
fr1/desk	0.040083	0.0464012
fr1/desk2	0.086565	0.0626974
fr2/xyz	0.066125	0.0557921
fr2/desk	0.004647	0.0035932
Average	0.065162	0.039911
Variance	0.000474	0.000553

35% of error reduction.

Conclusions

- It is possible to build dense and precise 3D maps using monocular data.
- 200 times more data from each frame with a 35% of error reduction.
- Limitations:
  - Constant but unreal scale.
  - Vulnerable to KSLAM failures.
- Part of this project has been submitted to the ROBOT'19 conference (O Porto, Portugal), organised by the SEIDROB and SPR societies.

Future work

- Try other VSLAM and depth estimation approaches.
- Improve ORB-SLAM2 performance using another vocabularies or tuning its parameters.
- Improve DeMoN performance providing the sparse map as an input and performing Transfer Learning.
- Include pixel-level semantic information to improve localisation.

End of presentation

Map Slammer:

Densifiying Scattered KSLAM 3D maps with Deep Learning

Q & A

Torres-Cámara, José Miguel