Limited-Scope Localization with RGB-D Features

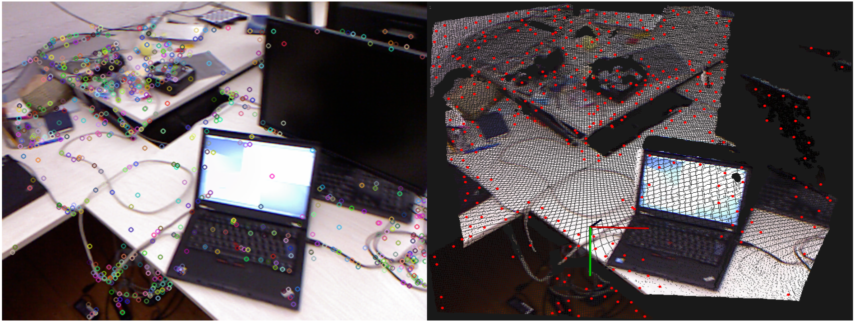

Figure: SIFT RGB and NARF depth detected key points.

Objective

- Match image queries against an representative image database of an indoor environment;
- Set-up and use the BRAND[1] RGB-D descriptor, comparing results against a well-known RGB descriptor;
- Start by applying RGB key point detectors following BRAND paper suggestions, and then investigate possible combinations with depth key point detectors.

[1] E. R. Nascimento, G. L. Oliveira, M. F. M. Campos, A. W. Vieira an W. R. Schwartz, "BRAND: A robust appearance and depth descriptor for RGB-D images," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, 2012, pp. 1720-1726.

Proposal

- Database and query images selected from TUM *fr1_room* RGB-D indoor dataset:
  - Whole environment was arbitrarily broke up in 16 representative images and loaded in a database;
  - Four of the temporal sequence neighbors of each image were selected as queries, presenting slight perspective changes of the same scene elements;
- SIFT descriptors as an alternative against BRAND;
- RGB Key points detected using OpenCV STAR, SIFT and combinations with NARF[2] and PCL SIFT for depth key points.

[2] B. Steder, R. B. Rusu, K. Konolige, and W. Burgard., "NARF: 3D Range Image Features for Object Recognition," In Workshop on Defining and Solving Realistic Perception Problems in Personal Robotics at the IEEE/RSJ Int. Conf. on Intelligent Robots and Systems (IROS). 2010.

Development

- Conversion of RGB and depth images into a point cloud:
  - Camera intrinsic parameters provided by TUM;
  - TUM provided script outputs only PLY files with unorganized clouds;
  - Alternative conversion was developed, receiving original dataset images and providing organized point cloud instances.

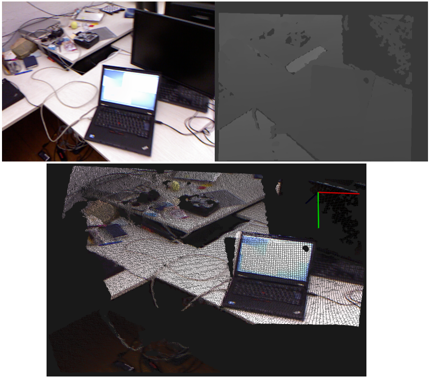

Figure: RGB, depth and point cloud.

Development

- Detection of SIFT and STAR RGB key points using OpenCV, with default parameters values:
  - STAR is derived from CenSurE detector, both originally designed for real-time tasks as lightweight alternatives.

Figure: SIFT and STAR detected key points.

Development

- SIFT matching using L2 metric for descriptor distances, with outlier rejection via cross-checking, Lowe's distance ratio threshold at 75%, and OpenCV RANSAC;
- BRAND matching using Hamming distance, with outlier rejection via cross-checking and OpenCV RANSAC;
- RANSAC first parameter increased from default value due to actual inliers rejection, while maintaining confidence parameter at default 99%.

Results - SIFT only

- Queries for images 2, 4, 7, 11 and 15 have insufficient distinguishable matchings, results were very similar among different database images - particularly for image 2;
- Better candidate for a target 2 query with wrong result had 183 matchings (correct image had 151).

	Database Images															
Query Images	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4															
2		3	1													
3			4													
4			1	3												
5					4											
6						4										
7		1					3									
8								4								
9		1							3							
10										4						
11		2									2					
12												4				
13													4			
14														4		
15		2													2	
16																4

Figure: SIFT key points and descriptor matching results.

Results - STAR + BRAND

- Queries for images 2 and 6 have insufficient distinguishable matchings in relation to their next image of the sequence;
- Queries for image 11 have very few inliers, results were similar among different database images.

	Database Images															
Query Images	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4															
2		2	2													
3			4													
4				4												
5					4											
6						3	1									
7							4									
8								4								
9									4							
10										4						
11		1		2						1						
12												4				
13													4			
14														4		
15		1													3	
16																4

Figure: STAR key points + BRAND matching results.

Results - SIFT + BRAND

- Queries for images 2, 11 have the lowest differences between better candidates and wrong alternatives (101-85 and 82-80 respectively).

	Database Images															
Query Images	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4															
2		4														
3			4													
4				4												
5					4											
6						4										
7							4									
8								4								
9									4							
10										4						
11											4					
12												4				
13													4			
14														4		
15															4	
16																4

Figure: SIFT key points + BRAND matching results.

Development

- Detection of NARF depth key points using PCL, with default parameters values;
- NARF uses extracted borders from range images converted from point clouds, detecting key points on surface changes:
  - Detection behaves poorly near depth shadows and sensor limits.

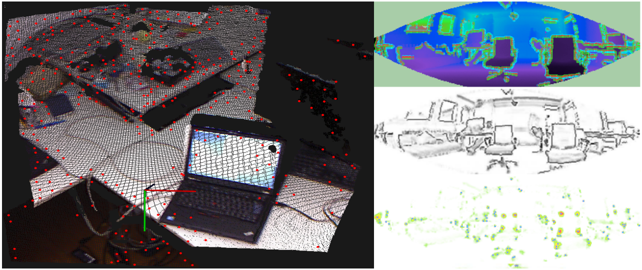

Figure: NARF key point detection.

Results - STAR + NARF + BRAND

- BRAND matchings of detected NARF key points outnumbered STAR key points, alleviating its problems - particulary for image 11 queries;
- Unmatched descriptors and wrong matches have grown due to detection near depth shadows.

	Database Images															
Query Images	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4															
2		4														
3			4													
4				4												
5					4											
6						4										
7							4									
8								4								
9									4							
10										4						
11											4					
12												4				
13													4			
14														4		
15															4	
16																4

Figure: STAR/NARF key points + BRAND matching results.

Development

- Detection of SIFT depth key points using PCL, with default parameters values;
- Key points estimation was changed, using z gradient instead of intensity, following PCL documentation.

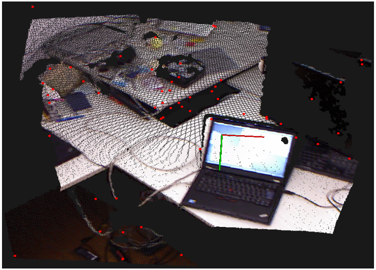

Figure: PCL SIFT key point detection.

Results - SIFT RGB and depth + BRAND

- Queries for images 2, 11 have the lowest differences between better candidates and wrong alternatives (123-93 and 97-95 respectively).

	Database Images															
Query Images	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	4															
2		4														
3			4													
4				4												
5					4											
6						4										
7							4									
8								4								
9									4							
10										4						
11										1	3					
12												4				
13													4			
14														4		
15															4	
16																4

Figure: SIFT RGB and depth key points + BRAND matching results.

Conclusion

- BRAND RGB-D descriptor matching improved with better key point detection (SIFT or STAR + NARF);
- SIFT RGB key points solved some issues of STAR use, but presented new problems;
- NARF depth key points helped alleviating some STAR issues, but behave poorly with depth shadows and sensor limits;
- SIFT depth key points failed to improve results, and have very slow detection process;
- Next steps:
  - Investigate recent depth key point detectors and RGB-D descriptors (e.g. Serafin et al work, LOIND, RISAS) - some of them promise improvements upon NARF and BRAND.