Dense 3D Visual Mapping via Semantic Simplification

Introduce

dense slam : high accuracy, but redundant data(planar regions)

sparse slam : insufficient data for 3d reconstruction

tradeoff methods

- regions with less details contain fewer vertices, but only applied to the mesh refinement stage
- n layers map / pieces-wise planar surface
- man-made structure
- leverage semantic information

Framework

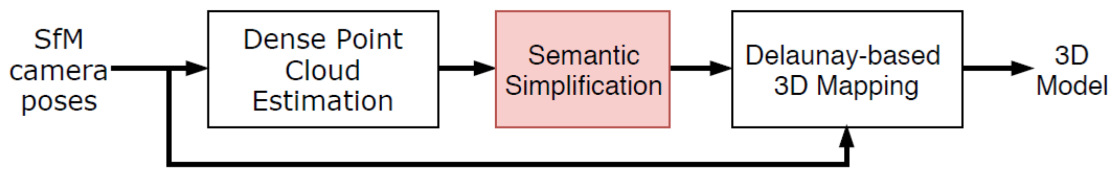


Fig. 1. Dense 3D Mapping pipeline.

in delaunay stage u_i is

$$1 - e^{\frac{d^2}{2\sigma_{\text{free}}}},$$

d is the distance between point to its closest facet

- project each point to segmented image to get class, also can use most frequently class on more image or compose weight
- reduce p belong to a class(compose a planar)
- region-wise method, use 半球面搜索/KNN : from p_i get the region ρ_i

  compute the factor c :

  $$c(\mathcal{D}, \bar{\mathcal{D}}, \tilde{p}) \in [0; 1],$$

  D is ρ_i density, $\overline{D}$ is average, $\hat{p}$ is the percent of different semantic

  discard last $n = (1 - c) * |p_i|$ points

- assume ground and wall is planar (to simplification), and identify a plane by RANSAC

  $$r(p_i) = (n_i - n_l)^T (n_i - n_l).$$

  expect the norm after ransac n_l is close to n_i

  use score to choose subset and compute the factor c

  - Linear Simplification

    $$c(\mathcal{D}, \bar{\mathcal{D}}) = -\frac{\bar{\mathcal{D}}}{\bar{c}} * (\mathcal{D} - \bar{\mathcal{D}}) + \bar{c},$$

  - Adaptive Simplification

    $$f(x) = 0.5 + \frac{wx}{2(1 + w|x|)},$$

    $$c(\mathcal{D}, \bar{\mathcal{D}}) = f(\mathcal{D} - \tau(\bar{\mathcal{D}}, \bar{r})).$$

  a small w corresponds to larger difference between two similar densities

  - Adaptive Class Simplification

    $$c(\mathcal{D}, \bar{\mathcal{D}}, \tilde{p}) = (1 + \tilde{p}) f(\mathcal{D} - \tau(\bar{\mathcal{D}}, \bar{r})),$$

when ρ_i contains few points of other classes, it maybe boundary, limit simplification

This is the function plot image

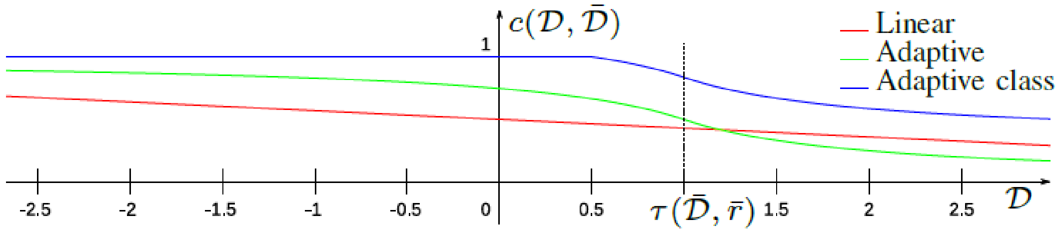


Fig. 3. The conservation factor trend for Linear simplification (with $\bar{c} = 0.5$ and $\tau(\bar{\mathcal{D}}, \bar{r})) = 1.0$), Adaptive Simplification ($w = 1$) and (with $\tilde{p} = 0.5$ and $\tau(\bar{\mathcal{D}}, \bar{r})) = 1.0$)

- Probabilistic simplification

in previous methods, extracted a region for each point. to limit the computational burden. Now predefined number of regions

$$P_c(p|\tilde{p}, \mathcal{D}, \bar{\mathcal{D}}, \rho_i) = \\ i(\tilde{p}) \cdot P_I(p|\mathcal{D}, \bar{\mathcal{D}}) + (1 - i(\tilde{p})) \cdot P_B(p|\mathcal{D}, \bar{\mathcal{D}}, \rho_i),$$

when percentage of points with a different semantics is less then 10%, i = 1.

in a flat surface :

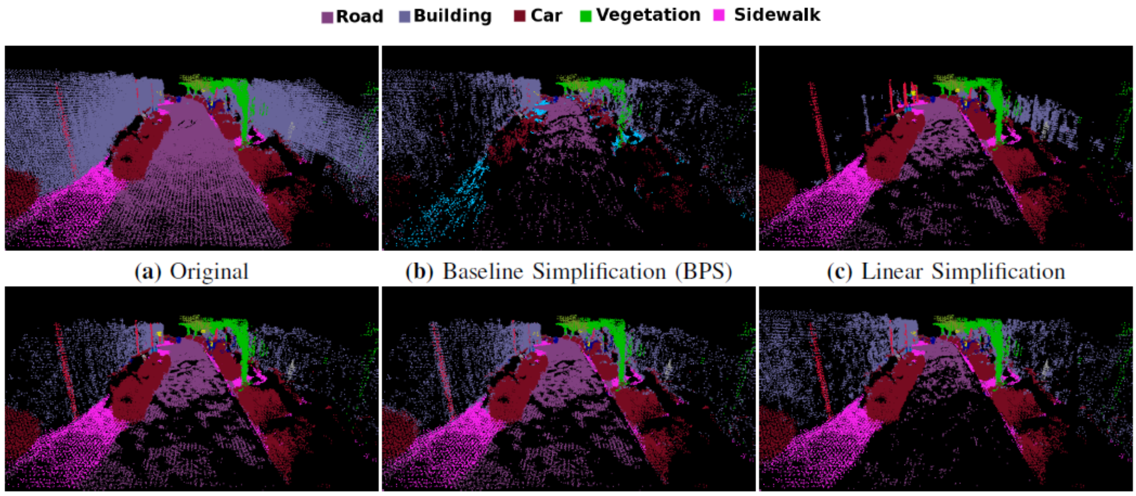
$$P_I(p|\mathcal{D}, \bar{\mathcal{D}}) = 1 - \mathcal{N}(\mu, \Sigma),$$

in a boundary : detect point far from boundary and remove it

$$P_B(p) = P_B(d_{\mathbf{pb}}) \sim \mathcal{N}(0, \sigma^2),$$

Result

in kitti dataset use different simplification method result

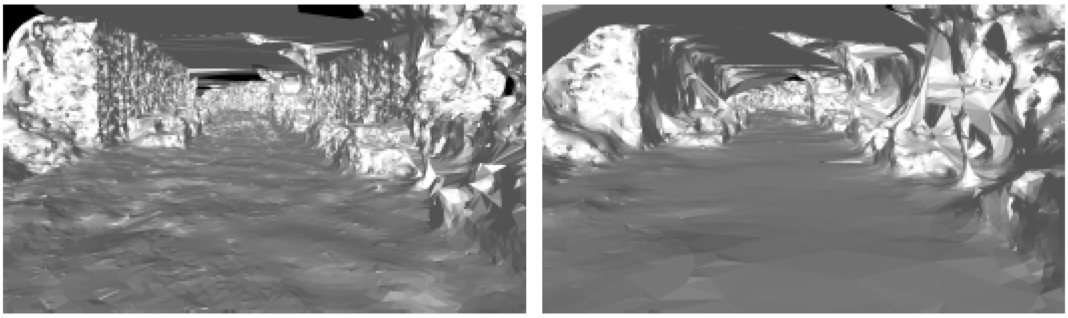


(a) Original (b) Baseline Simplification (BPS) (c) Linear Simplification

(d) Adaptive Simplification (e) Adaptive Class Simplification (f) Probabilistic Simplification

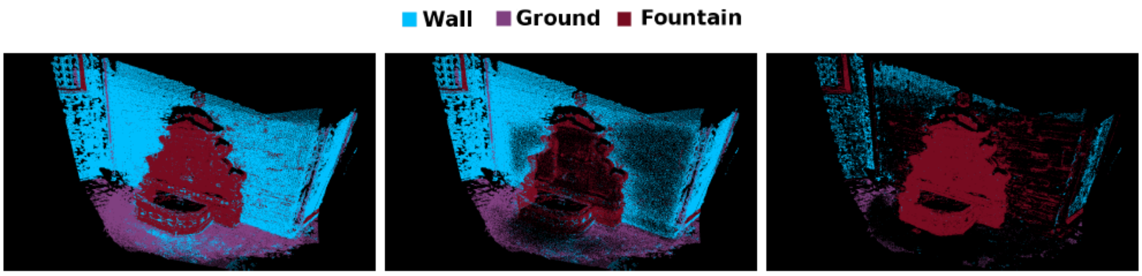
Fig. 5. Street perspective of the KITTI point clouds.

discard redundant data and the pcl is more smooth



(a) Original (b) Simplified

simplification background compared to other method



(a) Original (b) Baseline Simplification (BPS) (c) Probabilistic Simplification

Fig. 8. Frontal perspective of the fountain-P11 point clouds.

time and accuracy

Sequence	Original		Incremental Simplification	
	Accuracy RMSE(m)	Frame rate (f/s)	Accuracy RMSE(m)	Frame rate (f/s)
95	**0.7135**	1.0082	0.7173	**1.5434**
104	**0.7775**	0.4708	0.7846	**0.6397**