Autonomous Reconstruction of Unknown Indoor Scenes Guided by Time-varying Tensor Fields

KAI XU*, Shenzhen University and National University of Defense Technology
LINTAO ZHENG*, National University of Defense Technology
ZIHAO YAN, Shenzhen University
GUOHANG YAN, Shenzhen University
EUGENE ZHANG, Oregon State University
MATTHIAS NIESSNER, Technical University of Munich and Stanford University
OLIVER DEUSSEN, University of Konstanz and Shenzhen SIAT
DANIEL COHEN-OR, Shenzhen University and Tel-Aviv University
HUI HUANG†, Shenzhen University

ACM Reference format:
Kai Xu*, Lintao Zheng*, Zihao Yan, Guohang Yan, Eugene Zhang, Matthias Niessner, Oliver Deussen, Daniel Cohen-Or, and Hui Huang†. 2017. Autonomous Reconstruction of Unknown Indoor Scenes Guided by Time-varying Tensor Fields. *ACM Trans. Graph.* 36, 6, Article 202 (November 2017), 1 pages.
https://doi.org/10.1145/3130800.3130812

APPENDIX: RECONSTRUCTION ENTROPY AND VIEW QUALITY MEASURE

We measure the quality of a camera view by how much the uncertainty of reconstruction would be reduced when scanning from that view. Since we adopt the volumetric representation (TSDF) for depth fusion, we measure the uncertainty of reconstruction based on the probabilistic model for TSDF described in [Curless 1997]. Given a view v_i and the corresponding depth image d_i, we assume that the depth of a pixel (j,k) in the depth image can be modeled by a 1D Gaussian distribution: $d_i[j,k] \sim N(S_i[j,k], \sigma_i^2[j,k])$, where $\sigma_i^2[j,k]$ is the variance of depth measurement along the line of sight passing through pixel (j,k) from view v_i. $S_i[j,k]$ is the intersection of that line of sight with the reconstructed surface S.

In depth fusion based reconstruction [Curless and Levoy 1996], a TSDF is constructed to store sums of weighted signed distances, as well as sums of weights. The weights across the zero level set correspond to the reciprocals of accumulated variances, σ_t^2, which provides a measure of reconstruction uncertainty at time t. For example, the update of the accumulated variance along the line of sight of pixel (j,k) from t to $t+1$, given the depth measurement of a new view v_n can be written as:

$$(\sigma_{t+1}^2[j,k])^{-1} = (\sigma_t^2[j,k])^{-1} + (\sigma_n^2[j,k])^{-1}. \tag{1}$$

Given that the entropy of a 1D Gaussian is $\frac{1}{2}\ln(2\pi e\sigma^2)$, we can compute the information gain from t to $t+1$, given view v_n, as:

$$I_{t+1} = \frac{1}{2}\ln(\frac{\sigma_t^2}{\sigma_{t+1}^2}) = \frac{1}{2}\ln(1+\frac{\sigma_t^2}{\sigma_n^2}). \tag{2}$$

Writing it in terms of pixels of the depth image corresponding to view v_n, we have:

$$I_{t+1}(v_n) = \frac{1}{2}\sum_{j,k}\ln\left(1+\frac{\sigma_t^2[j,k]}{\sigma_n^2[j,k]}\right)\frac{d_n^2[j,k]}{f^2|\vec{n}_z[j,k]|}, \tag{3}$$

where $d_n[j,k]$ is the depth value for pixel (j,k), in the virtual depth image from view v_n, generated by sampling the current reconstruction S_t. f is the focal length of the depth camera. $\vec{n}_z[j,k]$ is the z-component of the surface normal at pixel (j,k). $\frac{d_n^2[j,k]}{f^2|\vec{n}_z[j,k]|}$ is the differential area on the reconstructed surface corresponding to pixel (j,k) (see [Krainin et al. 2011]).

The key to computing the information gain in Equation (3) is the estimation of variance σ_n^2 for each pixel. For a pixel whose line of sight intersects with the reconstructed surface (at a known voxel), the estimation of the standard deviation is given by the following experimental equation (measured in mm) found by Fankhauser et al. [2015]:

$$\sigma_n(d,\theta) = 1.5 - 0.5d + 0.3d^2 + 0.1d^{\frac{3}{2}}\theta^2/(90^\circ - \theta),$$

with d being the depth value for a pixel in the virtual depth image from view v_n, generated by sampling the current reconstruction S_t. θ is the angle between the line of sight and the surface normal at the intersection. For a pixel whose line of sight shoots into an unknown voxel, we directly compute the information gain as the maximum variance, with a standard deviation of 5cm.

REFERENCES

Brian Curless and Marc Levoy. 1996. A volumetric method for building complex models from range images. In *Proc. of SIGGRAPH*. 303–312.

Brian Lee Curless. 1997. *New methods for surface reconstruction from range images.* Ph.D. Dissertation. Stanford University.

*Joint first authors.
†Corresponding author (hhzhiyan@gmail.com).

0730-0301/2017/11-ART202 $15.00
https://doi.org/10.1145/3130800.3130812

Péter Fankhauser, Michael Bloesch, Diego Rodriguez, Ralf Kaestner, Marco Hutter, and Roland Siegwart. 2015. Kinect v2 for mobile robot navigation: evaluation and modeling. In *Proc. ICAR*. 388–394.

Michael Krainin, Brian Curless, and Dieter Fox. 2011. Autonomous Generation of Complete 3D Object Models Using Next Best View Manipulation Planning. In *Proc. ICRA*. 5031–5037.