Continuous Aerial Path Planning for 3D Urban Scene Reconstruction (Supplementary Material)

HAN ZHANG, Shenzhen University, China
YUCONG YAO, Shenzhen University, China
KE XIE, Shenzhen University, China
CHI-WING FU, The Chinese University of Hong Kong, China
HAO ZHANG, Simon Fraser University, Canada
HUI HUANG*, Shenzhen University, China

ACM Reference Format:
Han Zhang, Yucong Yao, Ke Xie, Chi-Wing Fu, Hao Zhang, and Hui Huang. 2021. Continuous Aerial Path Planning for 3D Urban Scene Reconstruction (Supplementary Material). *ACM Trans. Graph.* 40, 6, Article 1 (December 2021), 4 pages. https://doi.org/10.1145/3478513.3480483

A. CONVERGENCE ANALYSIS

This section presents an empirical analysis on how the following four measures vary as the algorithm progresses: (i) scene coverage, i.e., percentage of sampled surface points whose reconstructability (Eq. (1)) exceeds threshold h; (ii) objective value (Eq. (6)); (iii) trajectory length (Eq. (4)); and (iv) total cost for turning angles (Eq. (5)). Please note that Eqs. (1-6) in this supplementary material are referred to the corresponding equations in the main paper.

Fig. 1 shows the plots when running our algorithm on the UK-1 dataset. From the plots, we can see that all quantities increase initially, when the algorithm starts to build up the random tree with more nodes. Later, when the scene coverage improves to a level close to and beyond threshold percentage G, the optimization model starts to minimize the trajectory length and total cost for turning angles, as shown by the drops near the end of the bottom two plots. Yet, our algorithm can converge in just 150 iterations.

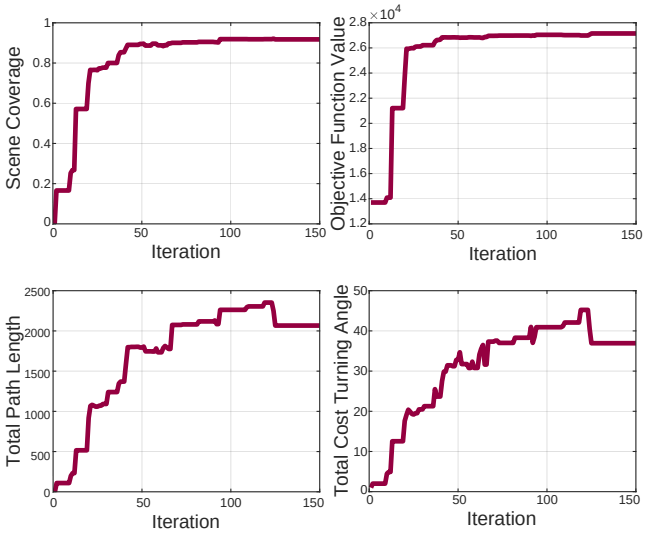

Fig. 1. Plots that show how the various measures (see above) vary as the algorithm progresses on the UK-1 dataset.

B. RECONSTRUCTION QUALITY VS. PATH LENGTH

This section presents an additional experiment that explores the tradeoff between the 3D scene reconstruction quality and the length of the aerial path. To explore their relationship, we make use of threshold h (see Algorithm 2 in the main paper), which is the minimum reconstructability value required for capturing each surface point on the scene proxy. Generally, a larger h requires more view coverage on the surface points, thus demanding for a longer aerial path and longer flight time for capturing more details. Hence, we can adjust h to tradeoff between the scene reconstruction quality and the path length in this experiment.

In more detail, we perform this experiment on the NY-1 benchmark dataset using a five-camera drone and plan aerial paths with different values of h. Table 1 reports the results, showing that longer aerial paths do not necessarily help to increase the accuracy of the 3D reconstruction, as revealed by the "error" measure. Yet, a larger h certainly increases the scene coverage and leads to a more complete 3D reconstruction, as revealed by the "completeness" measure. Importantly, it is worth to note that longer aerial paths may not necessarily capture more useful information.

C. IMPLEMENTATION DETAIL

Next, we provide further implementation details of our method.

(i) Details on D_i^j and $\omega_i(h)$. First, we give details of D_i^j and $\omega_i(h)$ in Eq. (1), which are from [Roberts et al. 2017], for self-containedness. For disk D_i^j, its radius decays exponentially as the associated

*Corresponding author: Hui Huang (hhzhiyan@gmail.com)

Authors' addresses: Han Zhang, Shenzhen University, China; Yucong Yao, Shenzhen University, China; Ke Xie, Shenzhen University, China; Chi-Wing Fu, The Chinese University of Hong Kong, China; Hao Zhang, Simon Fraser University, Canada; Hui Huang, College of Computer Science & Software Engineering, Shenzhen University, China.

0730-0301/2021/12-ART1 $15.00
https://doi.org/10.1145/3478513.3480483

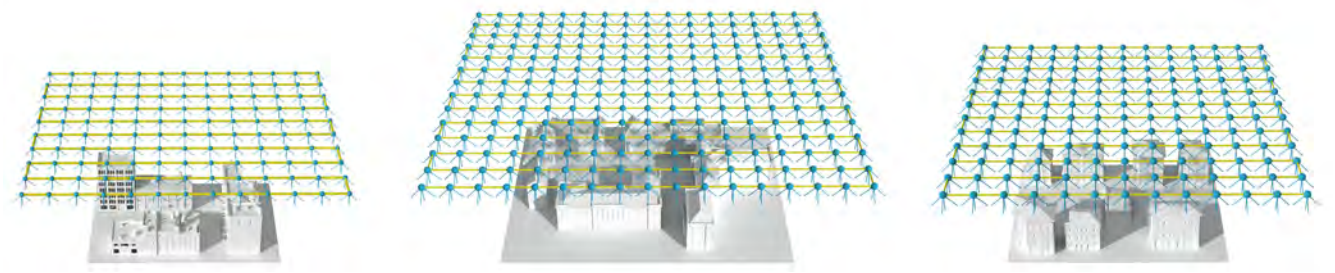

Fig. 2. The aerial paths of oblique photography for benchmarks NY-1, UK-1 and Bridge-1 from left to right.

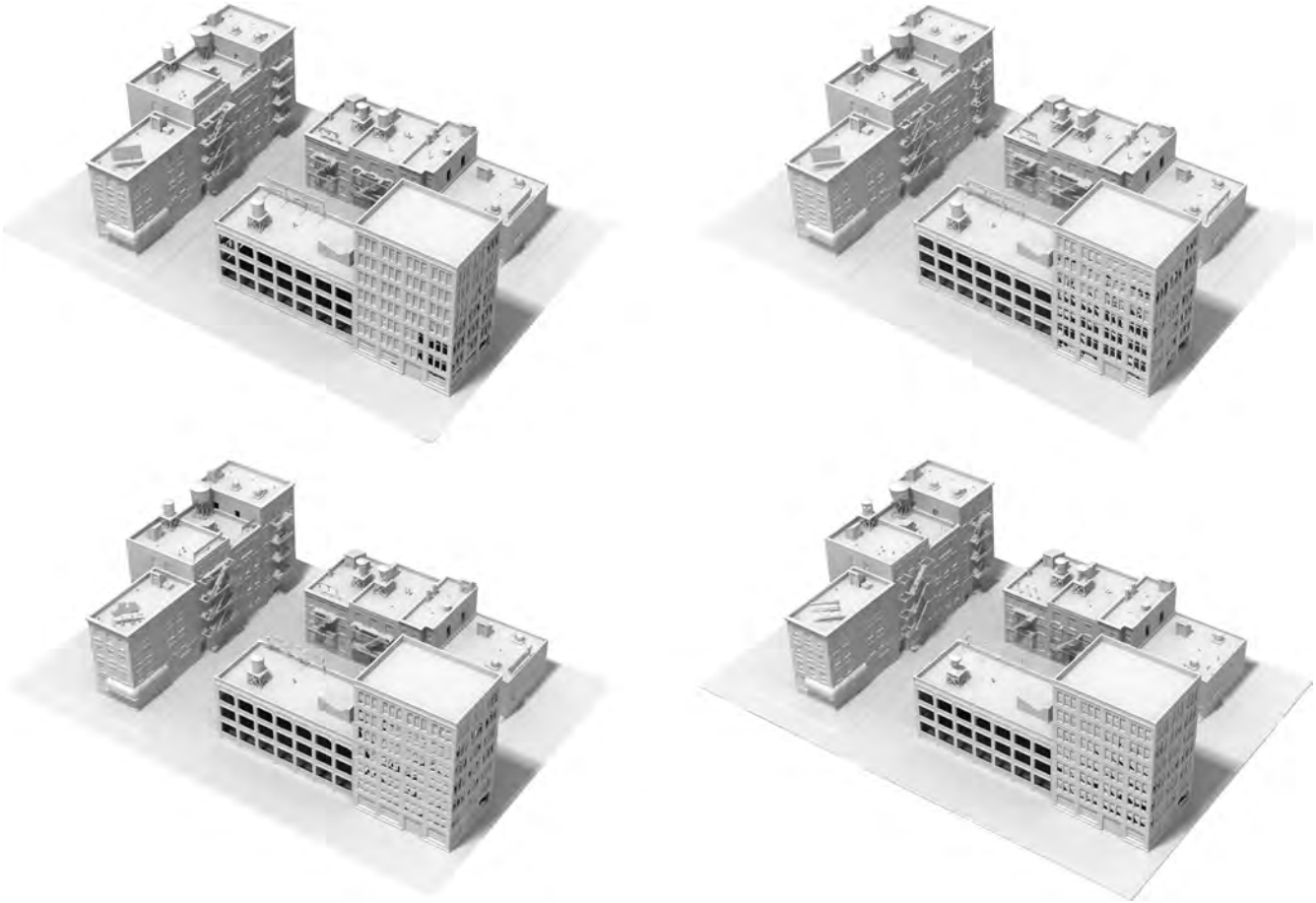

Fig. 3. Large views of the 3D reconstructions on NY-1 by our method (top-left), [Zhou et al. 2020] (top-right), [Smith et al. 2018] (bottom-left), and oblique photography (bottom-right).

camera moves away from surface point s_i. Here, a cone is formed with D_i^j as its base and s_i as its tip, and D_i^j's radius is adjusted by the cone's apex half angle:

$$\theta_i^j = \theta_{\max} \, 2^{\frac{-\max(t_i^j - t_0, 0)}{t_{\text{half}}}}, \tag{7}$$

where $\theta_{\max}$ is the maximum apex half angle in radians; t_i^j is the distance from camera to s_i in meters; t_0 is a threshold distance in meters, below which θ_i^j does not change; and t_{half} is the decay half-life of the half angle in meters. In the implementation, we set $\theta_{max} = \frac{\pi}{90}$ radians, $t_0 = 20m$, and $t_{\text{half}} = 25m$. For weight function $\omega_i(h)$, it is set to have a cosine-weighted falloff, i.e., $\omega_j(h) = cos(\alpha_h)$, where α_h is the angle between the hemisphere pole and the vector from the hemisphere origin (s_i) to the hemisphere location h.

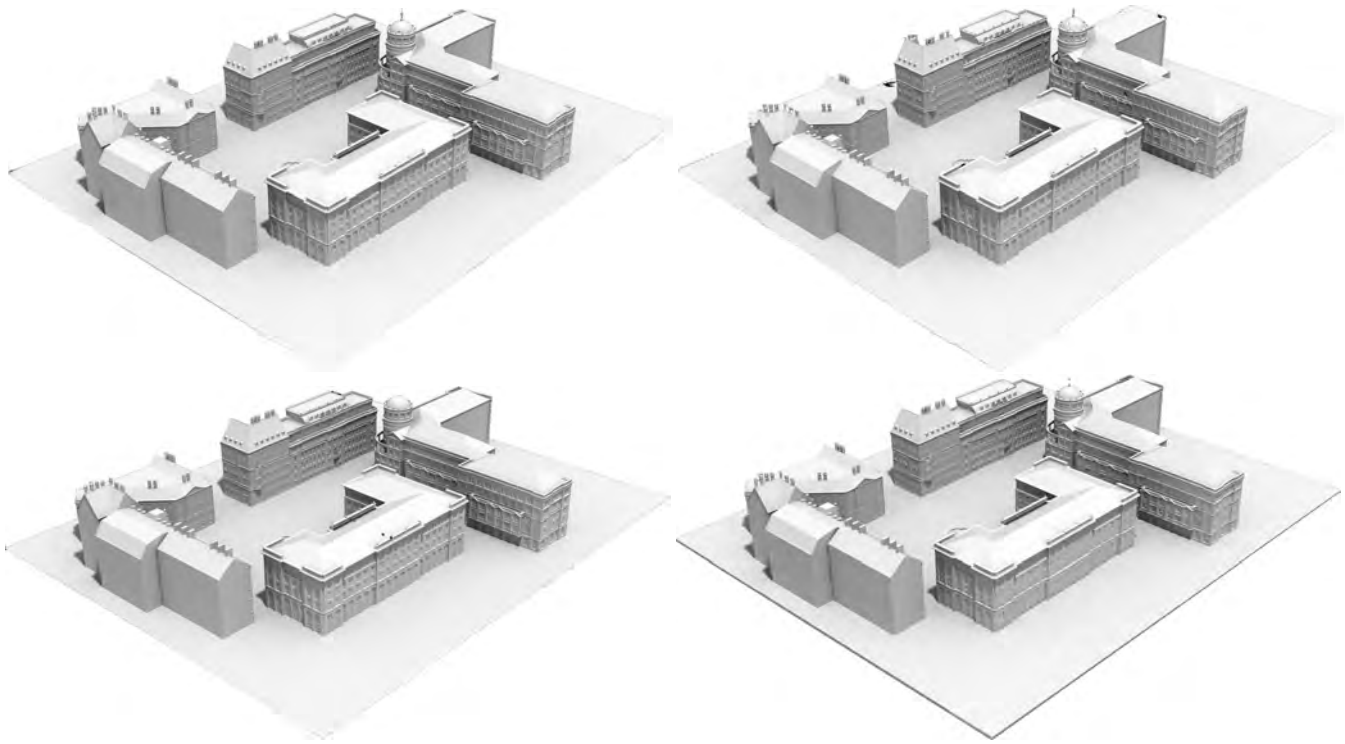

Fig. 4. Large views of the 3D reconstructions on UK-1 by our method (top-left), [Zhou et al. 2020] (top-right), [Smith et al. 2018] (bottom-left), and oblique photography (bottom-right).

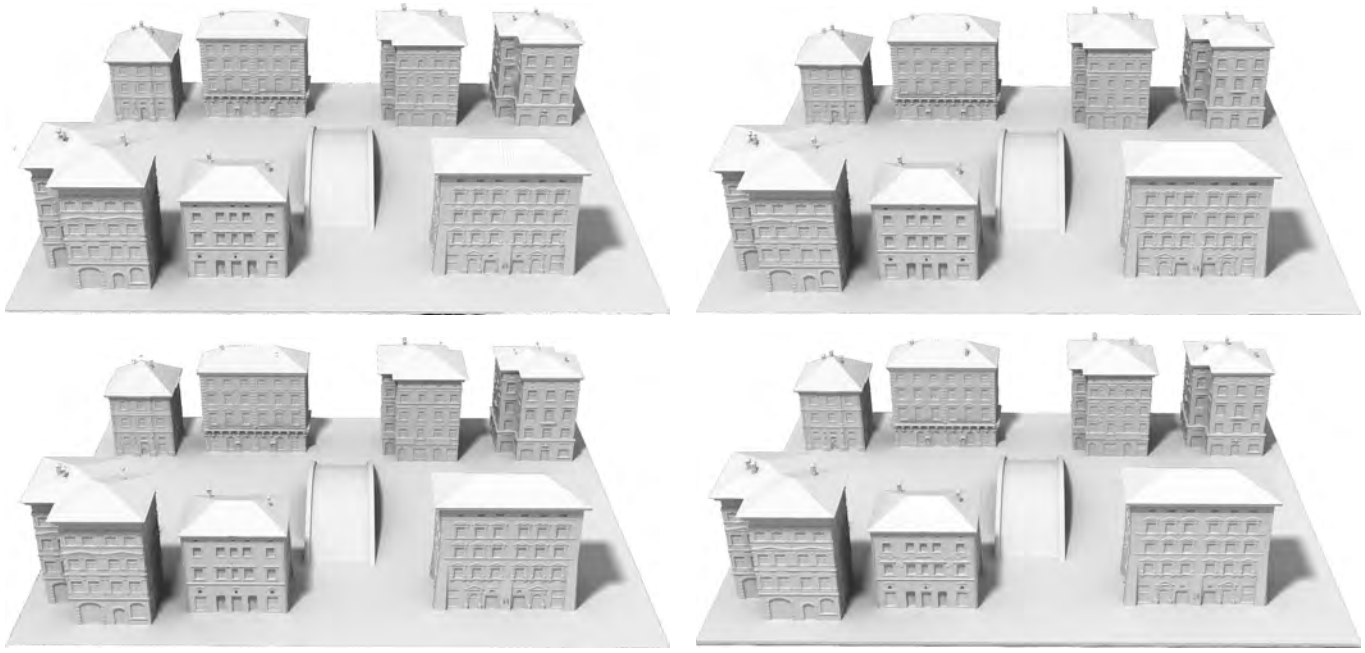

Fig. 5. Large views of the 3D reconstructions on Bridge-1 by our method (top-left), [Zhou et al. 2020] (top-right), [Smith et al. 2018] (bottom-left), and oblique photography (bottom-right).

Table 1. Exploring the tradeoff between 3D reconstruction quality and path length. The results reveal that longer aerial paths generally improves the completeness but it may not improve the reconstruction accuracy.

Threshold	Num. Images	Length [m]	Error [m] 90%	Error [m] 95%	Completeness [%] 0.020m	Completeness [%] 0.050m	Completeness [%] 0.075m
0.05	176	814	0.029	0.162	36.37	42.48	46.67
0.10	284	1,779	0.033	0.351	36.79	43.89	47.48
0.15	428	1,870	0.021	0.087	40.92	47.61	51.97
0.20	768	2,378	0.028	0.062	41.64	46.76	51.84
0.25	1,057	3,368	0.020	0.079	41.28	47.33	52.37

(ii) Hemisphere discretization. To efficiently store and compute the coverage disks, we discretize hemisphere H_i as N_h points evenly on H_i and represent a coverage disk as an N_h-bit vector; each bit marks if the associated point on H_i is inside or outside D_i^j. In our implementation, N_h is 256, so a coverage disk consumes only 16 bytes. Also, this representation allows us to implement the disk union and disk subtraction efficiently as bit-wise operations.

(iii) Coverage disk storage. A crucial element in the computation of reconstructability and information gain is the union of coverage disks that accumulates along the trajectory from the root of the tree. Here, we propose to store $\cup_j D_i^j$ per surface point per node in tree as an N_h-bit vector. This can speed up various computations, e.g., when we rewire a node to a new parent, we can update its $\cup_j D_i^j$ based on the $\cup_j D_i^j$ of its parent, rather than recomputing its $\cup_j D_i^j$ all the way from the root. Also, this strategy enables us to readily compute the reconstructability of any s_i at any node.

Note that after we finish this research work, we found that the supplementary material of [Roberts et al. 2017] describes a approach called coverage indicator matrix, which is similar to what we presented above. Yet, the overall approach is different. In our method, our VIF focuses on defining a continuous field for representing the information distribution in the free space so that we can compute the information gain *efficiently* and *continuously in space*.

(iv) View vector assignment for single-camera drone. For efficient computation, we need to discretize the space of view vectors, which is the lower half of $\mathbb{S}^2$ (since the camera needs not look upward). In detail, we discrete the longitude range in $\mathbb{S}^2$ by eight directions: $\{0, \frac{\pi}{4}, \frac{\pi}{2}, \ldots, \frac{7\pi}{4}\}$ and the latitude range in $\mathbb{S}^2$ by three directions: $\{0, \frac{\pi}{4}, \frac{\pi}{2}\}$ (in the lower hemisphere). In this way, the number of samples in the view-vector space is 17 (which is V in the main paper). During the view vector assignment for single-camera drone systems, we traverse each possible view vector and select the one that maximizes Eq. (3). We constrain the view vector not to deviate too much from the one in previous viewpoint by constraining the maximum angular difference between subsequent view vectors. Note that the maximum allowed angular difference depends on the distance between the two adjacent nodes in the random tree; if the distance is short, the maximum allowed angular difference should be small, such that the drone does not need to make a large turn for short-distance travel.

D. OBLIQUE PHOTOGRAPHY PATHS ON BENCHMARKS

To supplement Fig. 11 in the main paper, we show the paths of oblique photography for the three synthetic benchmark datasets, NY-1, UK-1, and Bridge-1, in Fig. 2 of this supplementary material.

E. LARGE VIEWS ON BENCHMARK RESULTS

Due to space limit in the main paper, we present Figs. 3, 4, and 5 in this supplementary material to show large views of the 3D reconstruction results on the three benchmark datasets.

ACKNOWLEDGMENTS

We sincerely thank the reviewers for their valuable comments. This work was supported in parts by NSFC (U2001206), GD Talent Program (2019JC05X328), GD Science and Technology Program (2020A0505100064), DEGP Key Project (2018KZDXM058), Shenzhen Science and Technology Program (RCJC20200714114435012, JCYJ20210324120213036), National Engineering Laboratory for Big Data System Computing Technology, and Guangdong Laboratory of Artificial Intelligence and Digital Economy (SZ).

REFERENCES

Mike Roberts, Debadeepta Dey, Anh Truong, Sudipta Sinha, Shital Shah, Ashish Kapoor, Pat Hanrahan, and Neel Joshi. 2017. Submodular Trajectory Optimization for Aerial 3D Scanning. In *Proc. of Int. Conf. on Computer Vision*. 5324–5333.

Neil Smith, Nils Moehrle, Michael Goesele, and Wolfgang Heidrich. 2018. Aerial Path Planning for Urban Scene Reconstruction: A Continuous Optimization Method and Benchmark. *ACM Trans. on Graphics (Proc. of SIGGRAPH Asia)* 37, 6 (2018), 183:1–183:15.

Xiaohui Zhou, Ke Xie, Kai Huang, Yilin Liu, Yang Zhou, Minglun Gong, and Hui Huang. 2020. Offsite Aerial Path Planning for Efficient Urban Scene Reconstruction. *ACM Trans. on Graphics (Proc. of SIGGRAPH Asia)* 39, 6 (2020), 192:1–192:16.