SCIENCE CHINA
Information Sciences

• **Supplementary File** •

Effectively Modeling Piecewise Planar Urban Scenes Based on Structure Priors and CNN

Wei WANG[1*], Wei GAO[2,3] & Zhanyi HU[2,3]

[1] *School of Network Engineering, Zhoukou Normal University, Zhoukou 466001, China;*
[2] *National Laboratory of Pattern Recognition, Institute of Automation, Chinese Academy of Sciences, Beijing 100190, China;*
[3] *University of Chinese Academy of Sciences, Beijing 100049, China*

Appendix A Data description

To evaluate the performance of the proposed method, we conducted experiments on several data sets of urban scenes where planar structures dominate:

(1) CASIA data sets [1]: Life science building (LSB, 4368×2912), Tsinghua school (TS, 2184×1456).

(2) Oxford VGG data sets [2]: Valbonne (512×768), Wadham (1024×768).

For the current image in each experiment, only its left and right neighboring images were used.

All the experiments were conducted on a desktop PC with Intel Core 4 Duo 4.0 GHz CPU and 32 G RAM. Our method in all experiments was implemented in parallel C++.

Appendix B Parameter settings

The proposed method appeared to be less sensitive to parameter settings; the majority of the parameters were fixed. The parameter settings are summarized in Table 1.

Table 1 Parameter settings

ID	Name	Default value	Function
1	γ	0.6	Weight of regularization term in Eq.1
2	λ_{occ}	2	Occlusion penalty
3	λ_{err}	4	Free-space violation penalty
4	λ_{dis}	2	Plane discontinuity penalty
5	μ	0.6	Relaxation parameter of structure priors
6	ρ	0.2	Weight of high-level image features
7	δ	0.5	Truncation threshold of color difference
8	ω	0.5	Weight of regularization term in Eq.7

Appendix C Initializations

The proposed method focuses on how to jointly optimize superpixels and their associated planes by incorporating scene structure priors. The initializations for different data sets are presented in Table 2.

Table 2 Initializations

Data set	3D point	Superpixel	Plane
LSB	6636	3788	28
TS	9265	3706	102
Valbonne	561	360	17
Wadham	2120	1243	38

* Corresponding author (email: wangwei@zknu.cn)

Appendix D Evaluation criteria

In this letter, we adopt the following criteria to evaluate the reliabilities of the reconstructed 3D points and planes.

(1)Reliable 3D points: the 3D points P_m and P_n corresponding to pixel $m \in I_r$ and $n \in N_i\ (i = 1, 2)$ respectively, are considered the same as the 3D point that is reliable to pixel $m \in I_r$ only when the difference (i.e., $(d\,(P_m) - d\,(P_n))\,/d\,(P_m)$)between the depths $d(P_m)$ and $d(P_n)$ with respect to the image I_r is less than a prespecified threshold (set to 0.2 in this letter).

(2)Reliable planes: for the reconstructed 3D points of all pixels in superpixel $s \in I_r$, the plane associated with superpixel s is considered reliable only when the percentage of reliable 3D points is greater than a prespecified threshold (set to 0.8 in this letter).

Based on the above definitions, we adopt the point accuracy M_1 and plane accuracy M_2 to comprehensively measure the accuracy of the scene reconstruction. Here, M_1 denotes the ratio of reliable reconstructed 3D points to all reconstructed 3D points and M_2 denotes the number of reliable planes.

Appendix E Quantitative results

In Table 3, $M_1(Ini)$ and $M_1(Opt)$ denote the accuracy of the initial scene structures produced in Step 2 and the scene structures optimized under the MRF framework, respectively. SRP denotes the number of initial superpixels with a reliable plane, and SP and PL denote the number of superpixels (including superpixels and sub-superpixels) and planes produced in Step 2, respectively.

Table 3 Quantitative results

Data set	SRP	SP	PL	$M_1(Ini)$	$M_1(Opt)$	M_2
LSB	292	6592	1107	0.6541	0.8835	18
TS	182	9896	2112	0.6073	0.7809	26
Valbonne	30	1940	156	0.5598	0.8184	9
Wadham	85	7113	409	0.7555	0.8701	12

For the efficiency of each method, as listed in Table 4, the proposed method performed relatively quickly at the multi-plane fitting, but consumed considerable time in extracting the high-level image features and resegmenting the superpixels. However, the plane assignment process was faster because of reliable candidate planes generated by the guidance of the angle priors. Further, the optimization process required less time owing to the improved initialization derived from the initial plane assignment produced in Step 2.

Table 4 Computation time (seconds)

Data set	Initial superpixel	Initial plane	Initial structure	Global optimization	Total time
LSB	4.7	12.7	56.8	2.9	77.1
TS	3.9	22.1	68.1	3.7	97.8
Valbonne	1.1	4.9	17.8	0.8	24.6
Wadham	2.4	7.6	34.2	1.1	45.3

References

1 [Online]: http://vision.ia.ac.cn/data/index.html
2 [Online]: http://www.robots.ox.ac.uk/ vgg/data/data-mview.html