Supplementary Note for
"Detecting Changes in 3D Structure of a Scene from Multi-view Images Captured by a Vehicle-mounted Camera"

Ken Sakurada   Takayuki Okatani   Koichiro Deguchi
Tohoku University, Japan
{sakurada,okatani}@vision.is.tohoku.ac.jp

Modeling $p(s_d)$ for correctly matched points

As described in Section 4, the proposed method uses models of $p(s_d)$, the density of the patch similarity s_d of the correctly matched pair of points, and also $p(s'_d)$ that is similarly defined for s'_d. Behind this, there is a fact that even for correctly matched points, s_d will not be 0 due to image noises and shape changes of the patches. As mentioned in Section 4.3 and 4.4, we chose $p(s_d) = \exp(-s_d/\sigma)/\sigma$ and the same model for $p(s'_d)$; for the parameter σ, we chose $\sigma = 1.5$ throughout our experiments.

These choices are made based on the following analysis of real images. As we do not know correct matches of points among the images and thus the true density $p(s_d)$ is difficult to obtain, we instead computes $\tilde{s}_d = \min_d s_d$, the minimum similarity over possible depth d; we generate its frequency histogram for images of scenes without severe occlusion and specular reflections. By excluding scenes with occlusions etc., $\tilde{s}_d$ should be a good substitute for s_d for correctly matched points. Figure 9(a) shows the histogram of $\tilde{s}_d$ for about 5 million points from 30 image pairs of such scenes. Figure 9(b) shows our model $p(s_d) = \exp(-s_d/\sigma)/\sigma$ with $\sigma = 1.5$. It is seen that the shape of the histogram is well approximated by our model of a half Laplace distribution. We manually chose the parameter as $\sigma = 1.5$ by considering a few differences between the ideal s_d and $\tilde{s}_d$, such as, that the histogram does not have the maximum peak at $s_d = 0$, whereas it ideally should have.
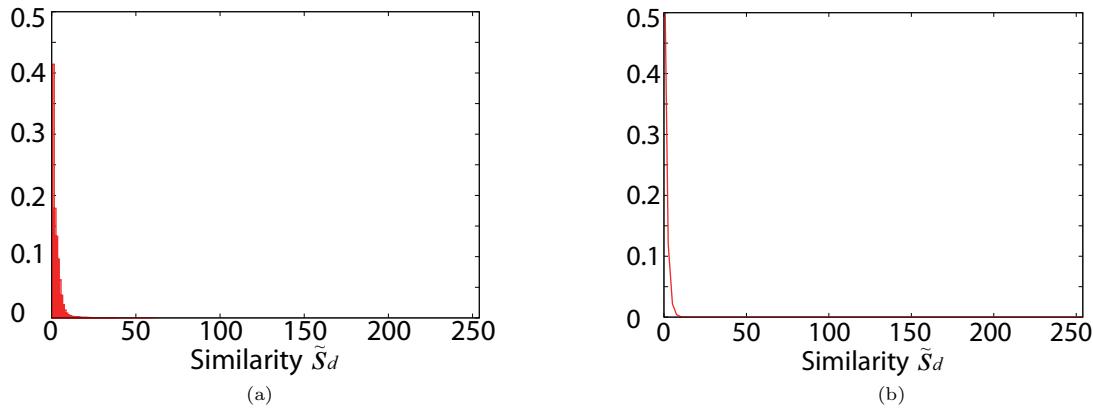
Figure 9. (a) Frequency histogram of $\tilde{s}_d$ for 5 million points from 30 pairs of images. (b) Our model of $p(s_d)$ for correctly matched points: a half Laplace distribution $\exp(-s_d/\sigma)/\sigma$ with $\sigma = 1.5$.

Prior on the probability of scene changes

In the proposed method, $p(c = 1)$, the prior on the probability of scene changes, needs to be specified. As mentioned in Section 4.4, we set $p(c = 1) = 0.5$ for all the experiments. We show here that the choice does not affect the results much. Figure 10 shows the results obtained when different values of $p(c = 1)$ are used. Table 3 shows the accuracy of change detection. It is seen from these that the results tend to be worse only for small $p(c = 1)$, i.e., $p(c = 1) \leq 0.3$.
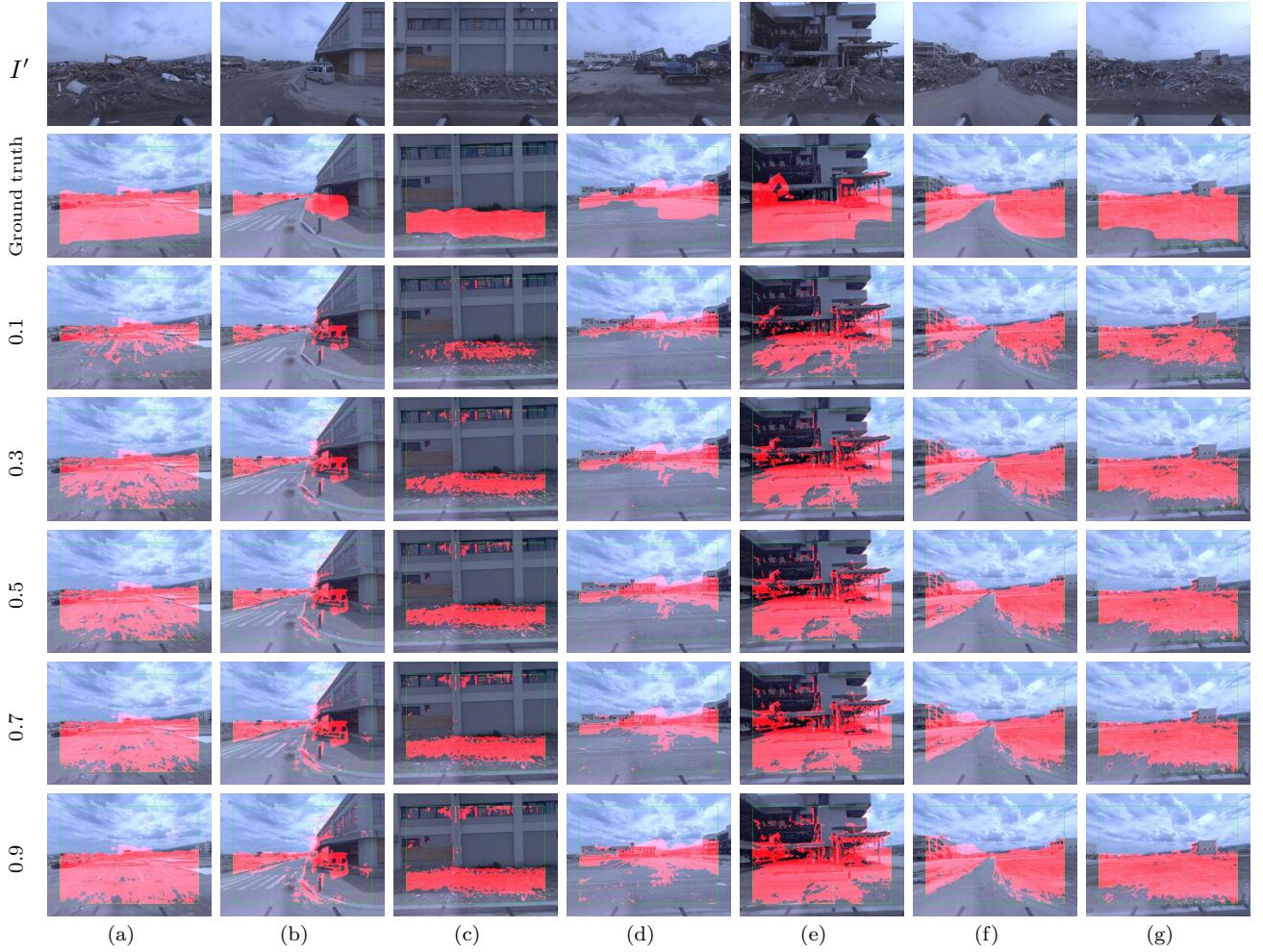
Figure 10. Results of the proposed method for different $p(c = 1)$ values.

More results of PMVS2 when applied to our image data

In the comparative experiments, we used PMVS2 as one of the methods that explicitly reconstruct scene structures. Although PMVS2 is known as one of the state-of-the-art methods for dense reconstruction from multi-view images, it does not produce good results for the images of urban areas captured by a camera mounted on a vehicle running in streets, as is mentioned in Section 1. We show here additional results demonstrating this.

As is described in our main paper, we input distortion-corrected versions of the six images captured by the six cameras comprising our omni-directional camera to PMVS2. Figure 11 shows these input images. Figure 12 shows the results of PMVS2 obtained from the images of two streets. The top row shows the overviews of the reconstructed scene structures, and the middle row shows their magnified portions. Comparing the latter with those of the input images shown in the bottom row of the figure, it is observed that there are many missing and erroneous parts, particularly where there is only limited texture.

Figure 11. An example of the set of six distortion-corrected images that are input to PMVS2 for each viewpoint.
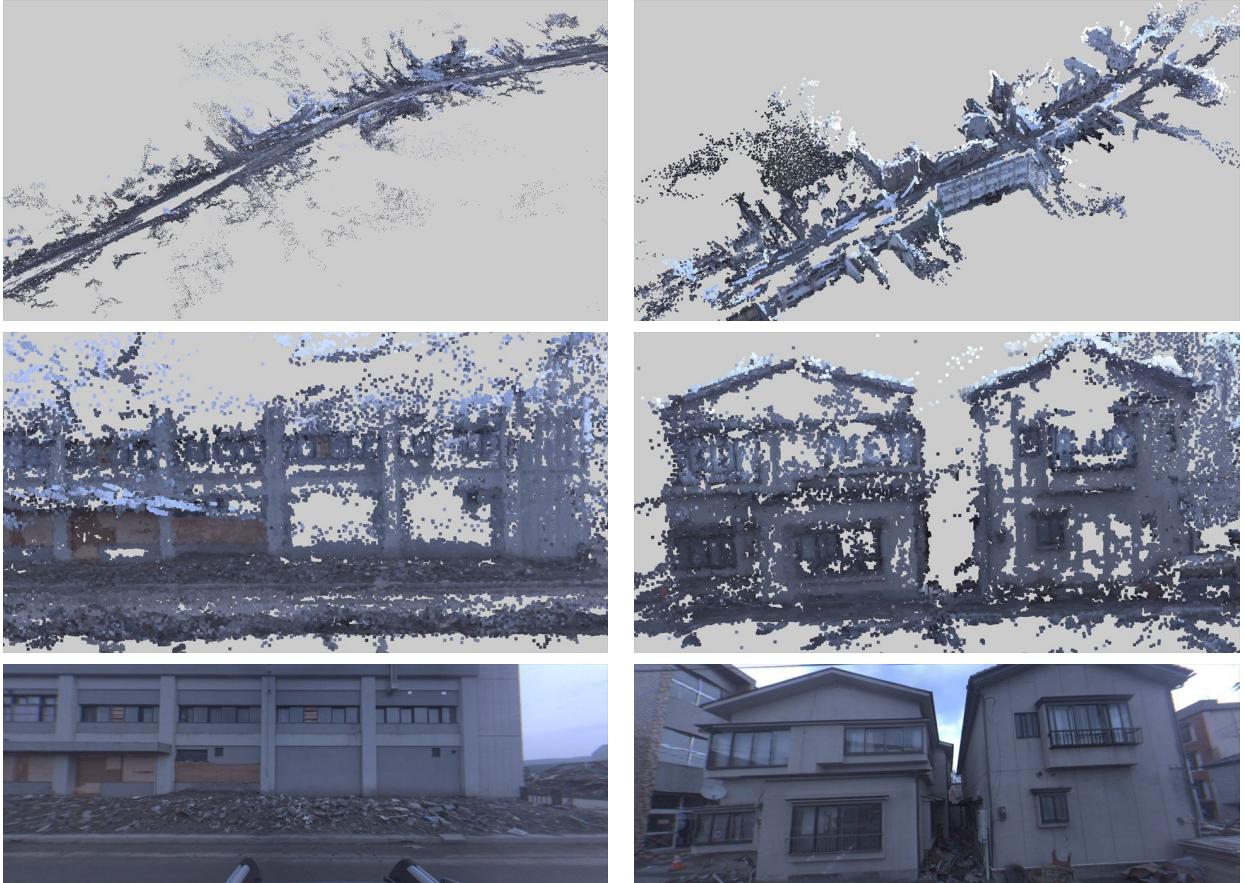
Figure 12. Results of PMVS2 when applied to our images. Top rows: The reconstructed structures. Middle rows: Their magnified portions. Bottom rows: One of the input images captured from similar viewpoints.

Additional results

We show here additional experimental results. Figure 13 shows an extended version of Fig.8, and Figs.14 and 15 show results for two different scenes. In these figures, the results obtained by the proposed method, PMVS2, Patch-MVS, and SIFT-MVS are shown, along with the depth maps obtained by PMVS2. Similar to the results shown in our main paper, it can be observed that the proposed method performs better than any of the other MVS-based methods.
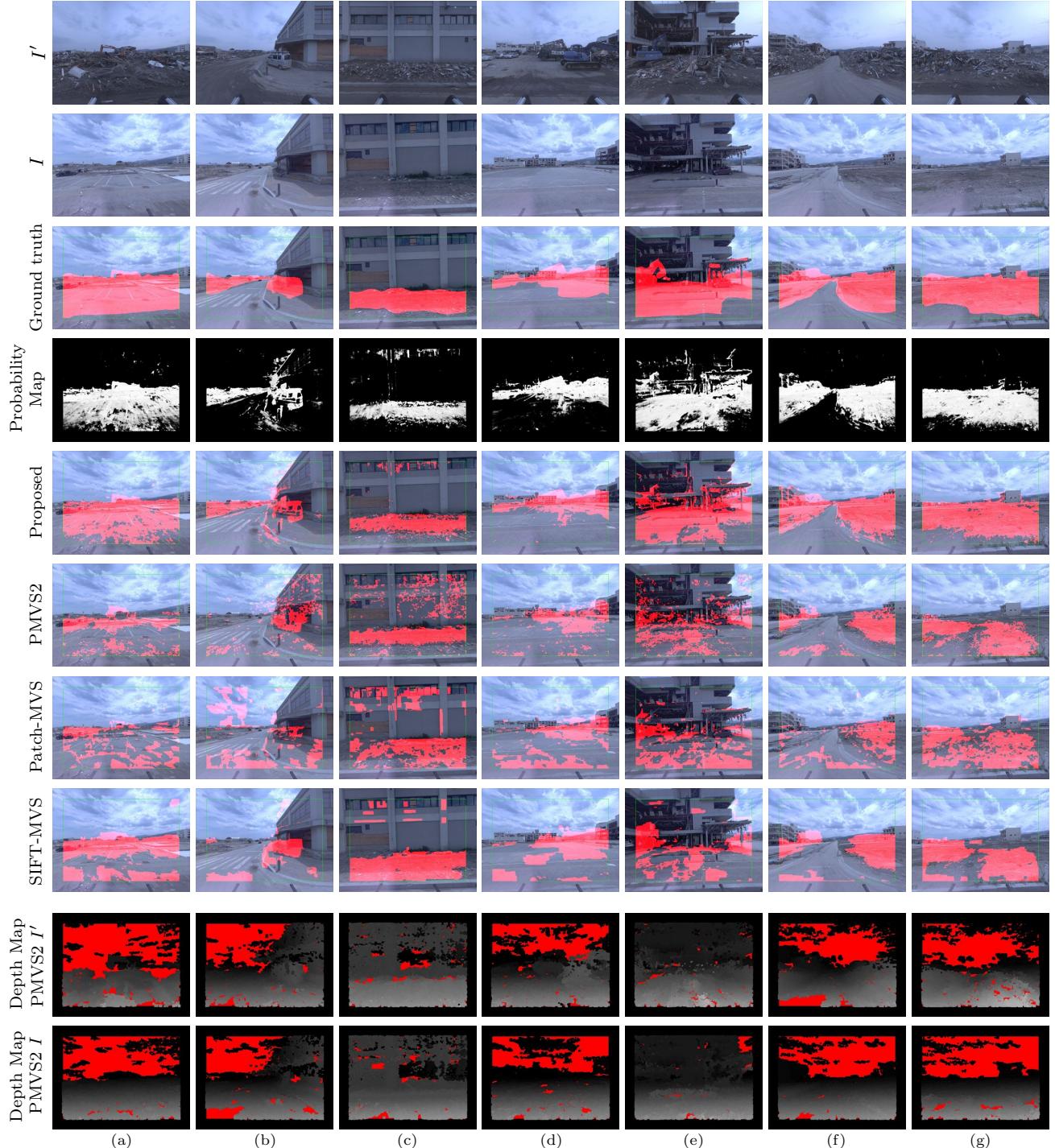
Figure 13. Extended results for the scene of Fig.8. From top to bottom rows, I', I, the change probability maps, the results of the proposed method, those of PMVS2, Patch-MVS, SIFT-MVS, and the depth maps obtained by PMVS2, respectively.
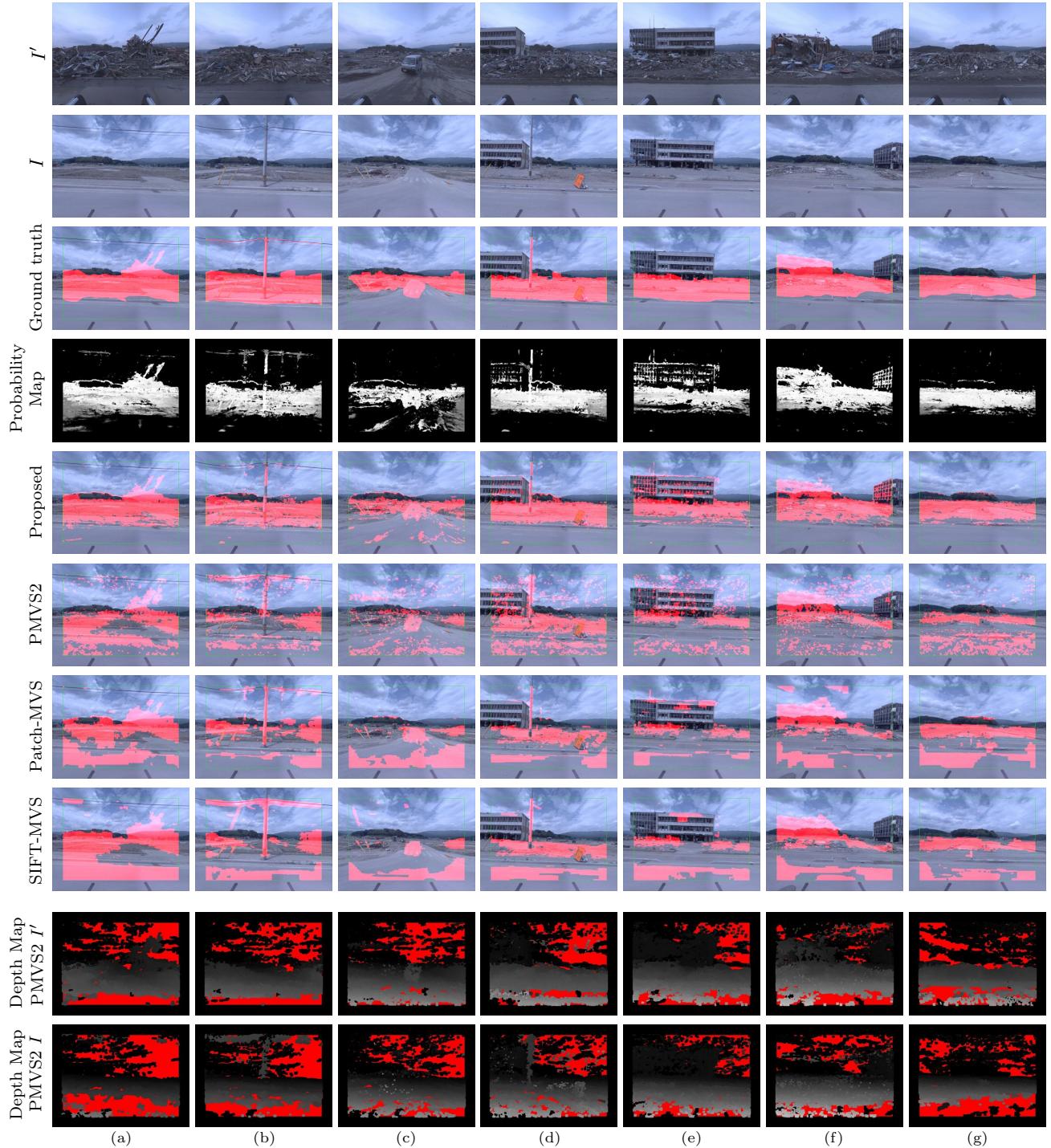
Figure 14. Results for a different scene. From top to bottom rows, I', I, the change probability maps, the results of the proposed method, those of PMVS2, Patch-MVS, SIFT-MVS, and the depth maps obtained by PMVS2, respectively.

5
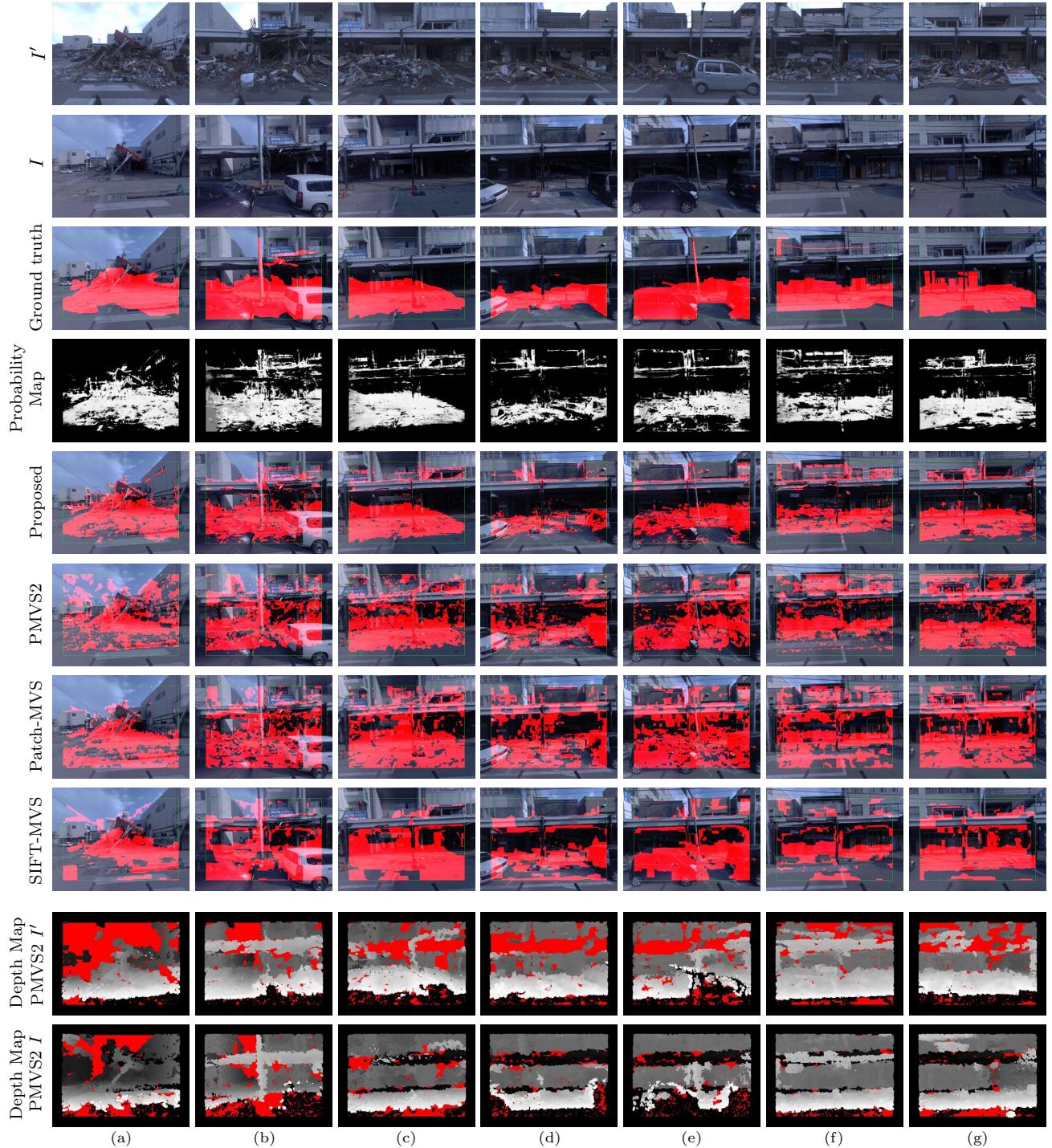
Figure 15. Results for a different scene. From top to bottom rows, I', I, the change probability maps, the results of the proposed method, those of PMVS2, Patch-MVS, SIFT-MVS, and the depth maps obtained by PMVS2, respectively.

Table 3. F_1 scores of the proposed method for different $p(c = 1)$ values for the scene shown in Fig. 10.

$p(c = 1)$	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Average
0.01	0.17	0.28	0.03	0.30	0.13	0.27	0.12	0.18
0.1	0.59	0.65	0.44	0.73	0.70	0.80	0.80	0.67
0.2	0.75	0.67	0.65	0.80	0.78	0.87	0.88	0.77
0.3	0.82	0.68	0.72	0.82	0.80	0.89	0.91	0.81
0.4	0.86	0.67	0.75	0.84	0.81	0.90	0.91	0.82
0.5	0.88	0.67	0.77	0.85	0.82	0.91	0.92	0.83
0.6	0.89	0.66	0.79	0.85	0.82	0.91	0.93	0.84
0.7	0.90	0.65	0.80	0.85	0.82	0.91	0.93	0.84
0.8	0.91	0.64	0.81	0.85	0.82	0.92	0.93	0.84
0.9	0.92	0.62	0.81	0.84	0.82	0.92	0.93	0.84
0.99	0.91	0.55	0.80	0.77	0.80	0.90	0.92	0.81
0.999	0.90	0.48	0.77	0.69	0.77	0.88	0.90	0.77

Table 4. F_1 scores of the detected changes shown in Fig. 14.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Average
Proposed	0.91	0.87	0.72	0.90	0.72	0.89	0.92	0.84
PMVS2	0.59	0.39	0.33	0.55	0.34	0.59	0.36	0.46
Patch-MVS	0.64	0.50	0.28	0.58	0.36	0.55	0.55	0.49
SIFT-MVS	0.58	0.38	0.25	0.52	0.31	0.55	0.53	0.46

Table 5. F_1 scores of the detected changes shown in Fig. 15.

	(a)	(b)	(c)	(d)	(e)	(f)	(g)	Average
Proposed	0.85	0.78	0.88	0.66	0.71	0.72	0.79	0.77
PMVS2	0.58	0.65	0.70	0.60	0.61	0.52	0.61	0.60
Patch-MVS	0.75	0.78	0.68	0.54	0.69	0.58	0.60	0.64
SIFT-MVS	0.71	0.75	0.74	0.64	0.76	0.57	0.68	0.68