Supplemental Document for "Interactive Images: Scene Understanding for Smart Image Manipulation"

In this supplemental document, we provide experimental compar-

isons to the work of Make3D [Saxena et al. 2009] and room-space 2

estimation [Lee et al. 2010; Gupta et al. 2011] framework. We also 3

conducted a user study to evaluate the effectiveness of our system.

Comparison with Other Methods 1 5

Figure 1 shows comparison examples between Make3D, image un-6

derstanding technique [Lee et al. 2010], and our system. Although 7

the Make3D framework can learn orientation and parameters for a 8 set of planar proxies to explain the image, such 3D information is

9 far from enough for geometric edits. 10

In another attempt, Gupta and colleagues [Lee et al. 2010; Gupta 11 et al. 2011] use simple user annotations to model the scene as a col-12 lection of candidate axis-aligned box. We do not need a Manhattan 13 assumption requiring objects to be axis aligned. Such assumption 14 limiting scope of image scenes. We use the images shown in [Lee 15 et al. 2010], since we did not have full access to the code of the 16 same system. 17

2 User Study Design 18

In order to evaluate the effectiveness of our system, we asked users 19

to distinguish between original images as well as our editing results. 20

We prepared 13 image pairs, with each pair composed by an orig-21

inal image as well as its associated editing result (see Figures 2-6 22

and marked as real or fake). Note the fake results are direct output 23

of our system, and were not rendered offline. 24

Each user was shown 13 images, one from each pair selected at ran-25 dom with an equal chance of real/fake. Each image was shown for a 26 maximum of 5 seconds and the user was asked to categorize the im-27

age as REAL or FAKE. The experiment takes about 1 minute/user.

User Study Results and Discussions 3 29

44 volunteers participated in our user study. Most of the users were 30 computer science graduates with many having background in com-31 puter graphics. We denote four types of possible results for each 32 image pair as: RR, RF, FR, and FF, each representing real image 33 recognized as real, real image recognized as fake, fake image rec-34 ognized as real, and fake image recognized as fake. On an average, 35 users recognized 63.2% real images as real, and 44.5% fake images 36 as fake 37

In Table 1, we show statistics results for each user. Results for each 38 image pair and each user is shown in Table 2. 39

References 40

- GUPTA, A., SATKIN, S., EFROS, A., AND HEBERT, M. 2011. 41
- From 3d scene geometry to human workspace. In IEEE CVPR, 42 1961-1968. 43
- LEE, D., GUPTA, A., HEBERT, M., AND KANADE, T. 2010. Esti-44
- mating spatial layout of rooms using volumetric reasoning about 45
- objects and surfaces. NIPS.

User ID	#Real	#Fake	RR	FR	FF	
1	6	7	83.3	42.9	57.1	
2	6	7	66.7	71.4	28.6	
3	6	7	50.0	57.1	42.9	
4	7	6	71.4	50.0	50.0	
5	6	7	50.0	28.6	71.4	
6	6	7	83.3	42.9	57.1	
7	4	9	100.0	66.7	33.3	
8	2	11	100.0	63.6	36.4	
9	5	8	60.0	62.5	37.5	
10	9	4	55.6	75.0	25.0	
11	6	7	66.7	42.9	57.1	
12	8	5	62.5	60.0	40.0	
13	8	5	75.0	60.0	40.0	
14	8	5	87.5	20.0	80.0	
15	6	7	66.7	57.1	42.9	
16	6	7	66.7	57.1	42.9	
17	10	3	70.0	100.0	0.0	
18	7	6	71.4	83.3	16.7	
19	7	6	57.1	66.7	33.3	
20	9	4	66.7	50.0	50.0	
21	7	6	42.9	83.3	16.7	
22	6	7	50.0	57.1	42.9	
23	6	7	83.3	85.7	14.3	
24	6	7	16.7	28.6	71.4	
25	8	5	37.5	40.0	60.0	
26	6	7	83.3	57.1	42.9	
27	6	7	66.7	57.1	42.9	
28	8	5	25.0	60.0	40.0	
29	6	7	83.3	28.6	71.4	
30	6	7	33.3	57.1	42.9	
31	6	7	33.3	71.4	28.6	
32	8	5	50.0	60.0	40.0	
33	11	2	72.7	50.0	50.0	
34	7	6	85.7	50.0	50.0	
35	7	6	57.1	50.0	50.0	
36	7	6	57.1	33.3	66.7	
37	5	8	60.0	50.0	50.0	
38	6	7	83.3	42.9	57.1	
39	6	7	66.7	42.9	57.1	
40	8	5	62.5	40.0	60.0	
41	6	7	83.3	71.4	28.6	
42	6	7	33.3	57.1	42.9	
43	8	5	37.5	40.0	60.0	
44	6	7	66.7	71.4	28.6	

Table 1: Summary statics for each user.

SAXENA, A., SUN, M., AND NG, A. 2009. Make3D: Learning 47 3d scene structure from a single still image. IEEE TPAMI 31, 5, 824-840.

48



Figure 1: Comparisons with Make3D [Saxena et al. 2009] and room-space estimation [Lee et al. 2010].

	S 0	S 1	S 2	S 3	S 4	S 5	S 6	S 7	S 8	S 9	S 10	S 11	S 12
User 0	FF	RR	RR	FF	RF	RR	FF	RR	FR	RR	FR	FR	FF
User 1	RR	FR	FR	FF	FR	RF	RR	FR	FF	RR	FR	RR	RF
User 2	FF	FR	RF	FR	FR	FF	RF	RR	RR	RF	FR	FF	RR
User 3	FR	FR	RR	RF	RF	FF	FR	RR	FF	RR	FF	RR	RR
User 4	FF	RF	FF	FF	FF	RR	FR	FR	RF	RF	RR	FF	RR
User 5	RR	FF	FR	FF	FF	RF	RR	FR	FF	RR	FR	RR	RR
User 6	RR	RR	FR	FF	FR	FF	FR	RR	FF	RR	FR	FR	FR
User 7	FF	FR	FF	RR	RR	FR	FR	FF	FR	FR	FR	FR	FF
User 8	RF	FR	FF	FF	FR	FR	RR	FR	FF	RR	FR	RF	RR
User 9	FR	RF	RR	FF	RR	RF	RF	RR	FR	RR	FR	RF	RR
User 10	RR	FR	RR	FF	FF	FR	FF	FR	FF	RF	RR	RR	RF
User 11	FR	RR	FF	RF	FR	FF	RF	RR	FR	RR	RR	RF	RR
User 12	RR	RR	FR	RR	FF	RF	FR	RR	RF	FF	RR	FR	RR
User 13	RR	RR	RR	FF	RF	RR	FF	FR	RR	RR	RR	FF	FF
User 14	FF	RF	FR	FF	RF	FF	FR	RR	RR	FR	RR	FR	RR
User 15	RR	RR	FF	FF	FR	RF	RF	FR	FR	RR	FF	FR	RR
User 16	RR	RR	RF	RF	RR	RR	RF	RR	FR	FR	RR	FR	RR
User 17	RR	FR	FF	RF	RF	RR	FR	FR	FR	RR	FR	RR	RR
User 18	RR	RR	RF	FF	FR	FF	RF	RR	FR	RF	RR	FR	FR
User 19	RR	RF	RR	RF	FF	RR	FF	FR	RR	RR	FR	RF	RR
User 20	RR	FR	RF	FF	RF	RR	FR	FR	RR	FR	RF	RF	FR
User 21	RR	RF	FR	FF	RF	FF	RF	FR	RR	FR	RR	FR	FF
User 22	FF	RR	RR	RF	RR	FR	FR	FR	RR	FR	FR	FR	RR
User 23	RF	FR	FR	RF	RF	FF	RF	FF	RR	FF	FF	RF	FF
User 24	FF	FF	RF	RF	RF	FR	RR	RR	FF	RR	RF	FR	RF
User 25	RR	FR	FR	FF	FF	RR	RR	FF	FR	RF	FR	RR	RR
User 26	RR	FR	FR	FF	FF	RR	RR	FF	FR	RR	FR	RF	RF
User 27	FF	FR	RF	RF	RF	RR	FF	FR	FR	RF	RF	RR	RF
User 28	RR	FR	FR	FF	FF	RF	RR	FF	FF	RR	FF	RR	RR
User 29	RF	FR	FR	FF	FF	RF	RF	FR	FF	RR	FR	RR	RF
User 30	RF	FR	FR	FR	FR	RR	RF	FF	FF	RR	FR	RF	RF
User 31	RR	FF	RF	RF	FR	FF	RR	FR	RF	RF	RR	FR	RR
User 32	RF	FR	RR	FF	RR	RF	RR	RR	RR	RR	RR	RF	RR
User 33	RR	FR	FF	RR	FF	FR	RR	RF	FR	RR	FF	RR	RR
User 34	RF	FF	FF	RF	FR	RF	RR	FR	RR	RR	FR	FF	RR
User 35	FF	RF	RF	FF	RR	RR	RF	FR	FR	FF	RR	FF	RR
User 36	FF	FF	RR	FF	RF	FR	FR	RR	FR	FF	RR	FR	RF
User 37	RR	FF	FF	FF	FR	RF	RR	FR	FR	RR	FF	RR	RR
User 38	FF	FF	RF	RR	FF	FR	FR	FF	RR	RF	RR	FR	RR
User 39	FF	RF	RR	RF	FF	FR	RR	FR	RR	RR	FF	RR	RF
User 40	RR	FF	FR	FF	FR	RF	RR	FR	FR	RR	FR	RR	RR
User 41	RF	FR	FR	FF	RF	RR	RF	FR	FF	FR	RR	RF	FF
User 42	FF	RF	RF	RF	FF	FF	RF	RF	RR	RR	FR	RR	FR
User 43	RR	FR	FF	RF	FR	FF	FR	FR	RF	RR	RR	RR	FR
RR Ratio	54.2	75.8	55.0	38.9	69.2	25.0	86.7	76.5	52.9	84.2	60.0	51.9	72.7
FR Ratio	45.0	63.6	50.0	50.0	27.8	8.3	75.9	59.3	70.4	72.0	73.7	70.6	45.5
FF Ratio	55.0	36.4	50.0	50.0	72.2	91.7	24.1	40.7	29.6	28.0	26.3	29.4	54.5

 Table 2: User study results for each user across all test images.



Figure 2: Examples 1 - 3.

original image

edited result counterpart



Figure 3: Examples 4 - 5.



Figure 4: Examples 6 - 8.

6



Figure 5: Examples 9 - 11.

Online Submission ID: 0401



original image

edited result counterpart

Figure 6: *Examples 12 - 14.*