

Supplementary Material for Holistic⁺⁺ Scene Understanding: Single-view 3D Holistic Scene Parsing and Human Pose Estimation with Human-Object Interaction and Physical Commonsense

Yixin Chen^{*1}, Siyuan Huang^{*1}, Tao Yuan¹, Siyuan Qi^{1,2}, Yixin Zhu^{1,2}, and Song-Chun Zhu^{1,2}

^{*} Equal Contributors

¹ University of California, Los Angeles (UCLA)

² International Center for AI and Robot Autonomy (CARA)

{ethanchen, huangsiyuan, taoyuan, syqi, yixin.zhu}@ucla.edu, sczhu@stat.ucla.edu

1. Parametrization

We represent the objects and room layout for each scene as 3D bounding boxes. Each 3D bounding box is parametrized by its 3D size $S \in \mathbb{R}^3$, center $C \in \mathbb{R}^3$, and orientation $Rot(\theta) \in \mathbb{R}^{3 \times 3}$, all in world coordinates. The 3D boxes can be reconstructed by first computing the 8 bounding box corners with center and size, and then rotate all the corners in x-y plane with θ . Our parametrization is similar to [2, 1].

2. Baseline Model

As mentioned in Section 6.1 in the paper, we design a baseline model for multi-person 3D pose estimation in world coordinate.

We first extract a 2048-D image feature vector using the Global Geometry Network (GGN) [1] to capture the global geometry of the scene. Then we concatenate GGN image feature, 2D pose, 3D pose in the local coordinate, together with the camera intrinsic matrix as a feature vector, which is then fed into a 5-layer fully connected network to predict the global 3D pose. The fully-connected layers are trained using the mean squared error loss. The network structure of the baseline model is shown in Figure 1.

As described in Section 6.2, we augment SUN RGB-D dataset [3] by projecting sampled 3D poses back onto the image plane, which gives us the ground-truth global 3D poses and their corresponding 2D poses. We then train the proposed baseline on the training set of the synthetic SUN RGB-D dataset, which has 21234 pose pairs under 2666 different scenes.

3. Additional Qualitative Results

Figure 2 to Figure 101 show additional qualitative results.

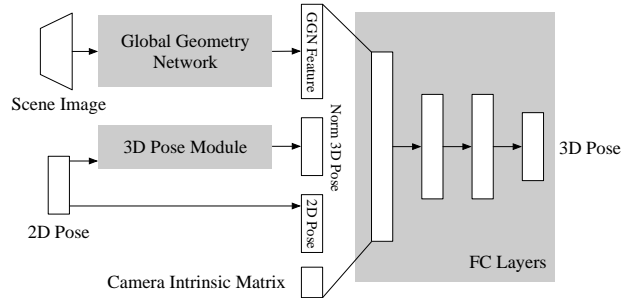


Figure 1. Baseline model for global 3D pose estimation.

References

- [1] Siyuan Huang, Siyuan Qi, Yinxue Xiao, Yixin Zhu, Ying Nian Wu, and Song-Chun Zhu. Cooperative holistic scene understanding: Unifying 3d object, layout and camera pose estimation. In *Conference on Neural Information Processing Systems (NIPS)*, 2018. 1
- [2] Charles R Qi, Wei Liu, Chenxia Wu, Hao Su, and Leonidas J Guibas. Frustum pointnets for 3d object detection from rgb-d data. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018. 1
- [3] Shuran Song, Samuel P Lichtenberg, and Jianxiong Xiao. Sun rgb-d: A rgb-d scene understanding benchmark suite. In *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2015. 1

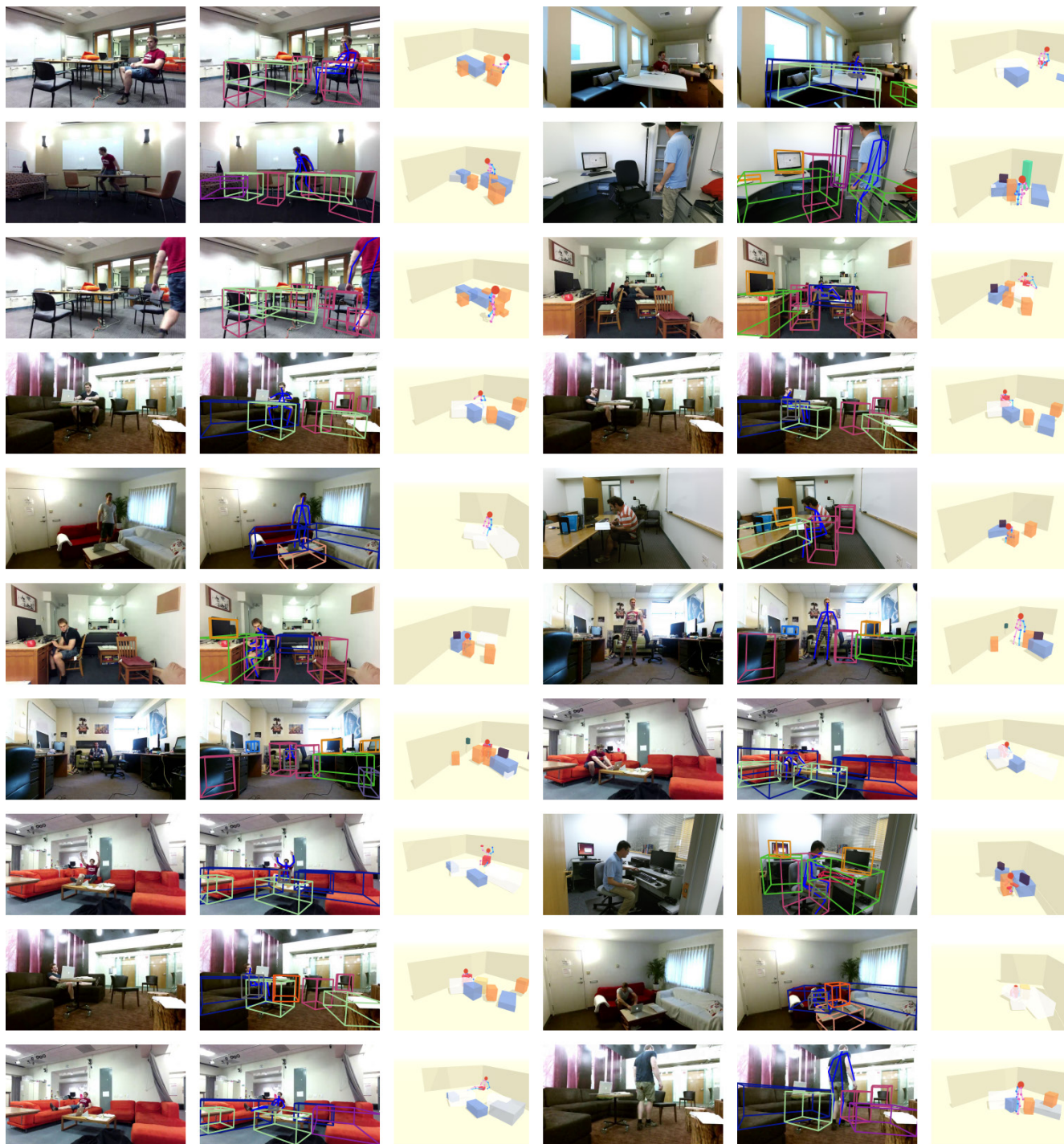


Figure 2. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

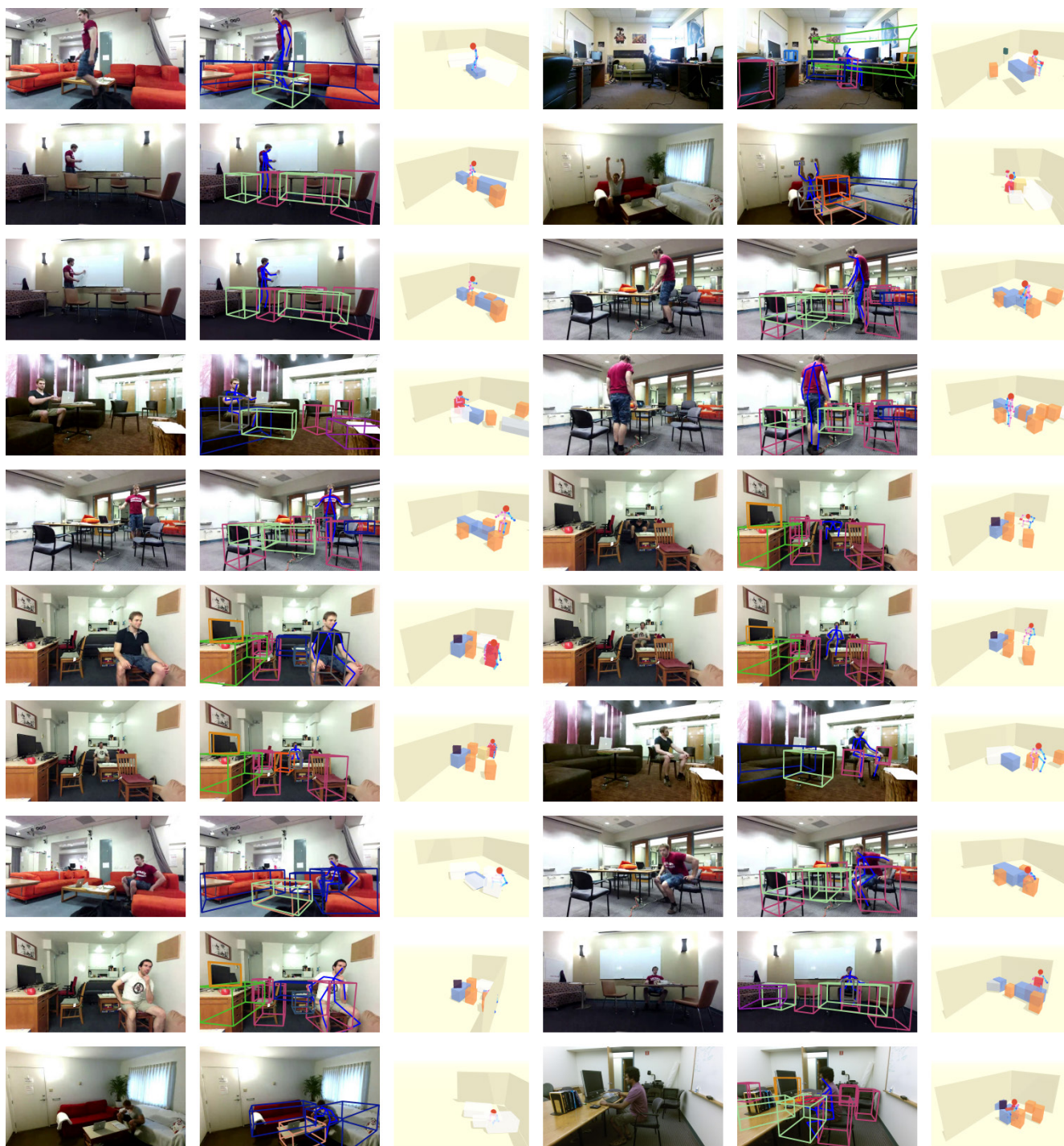


Figure 3. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

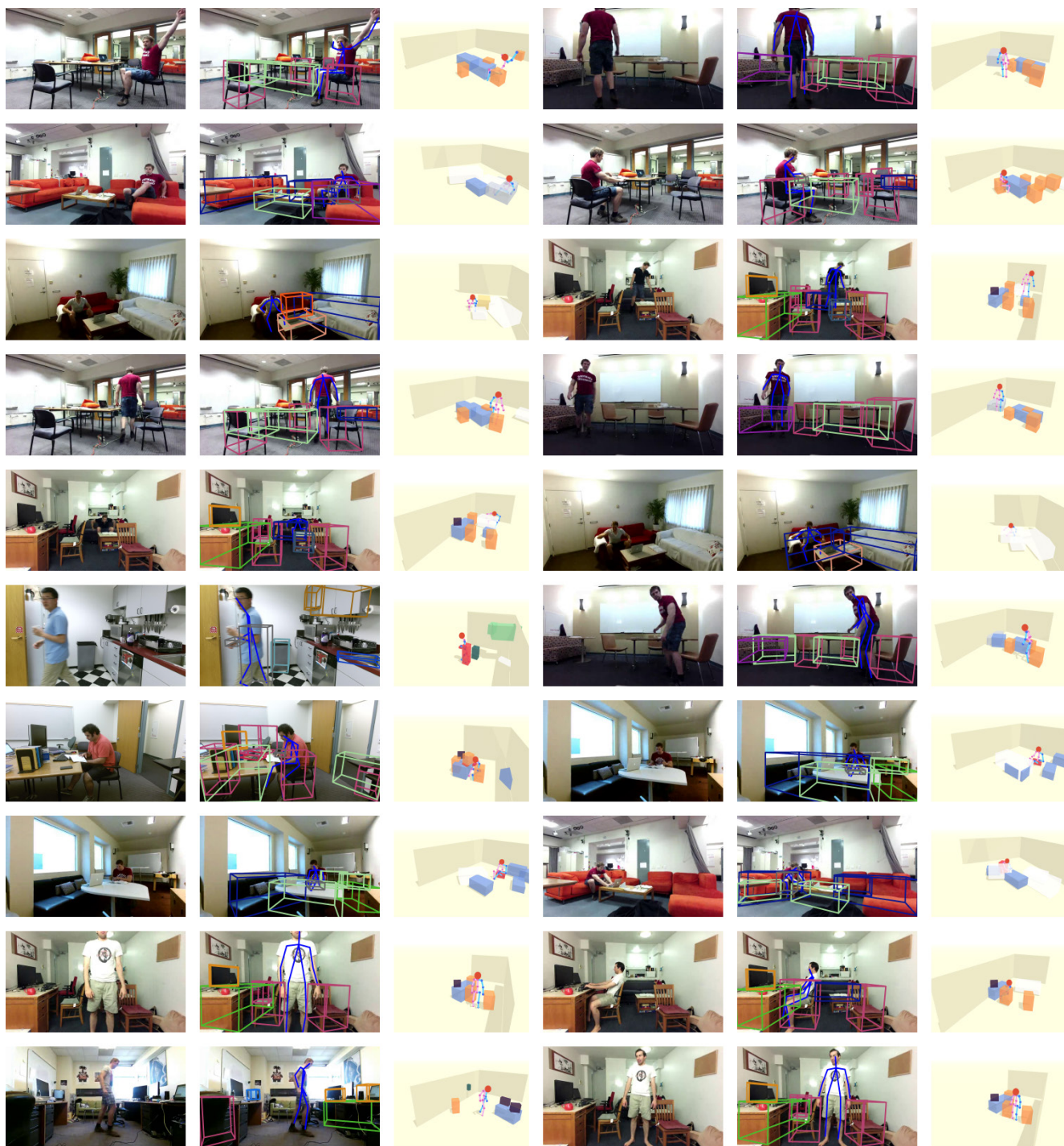


Figure 4. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

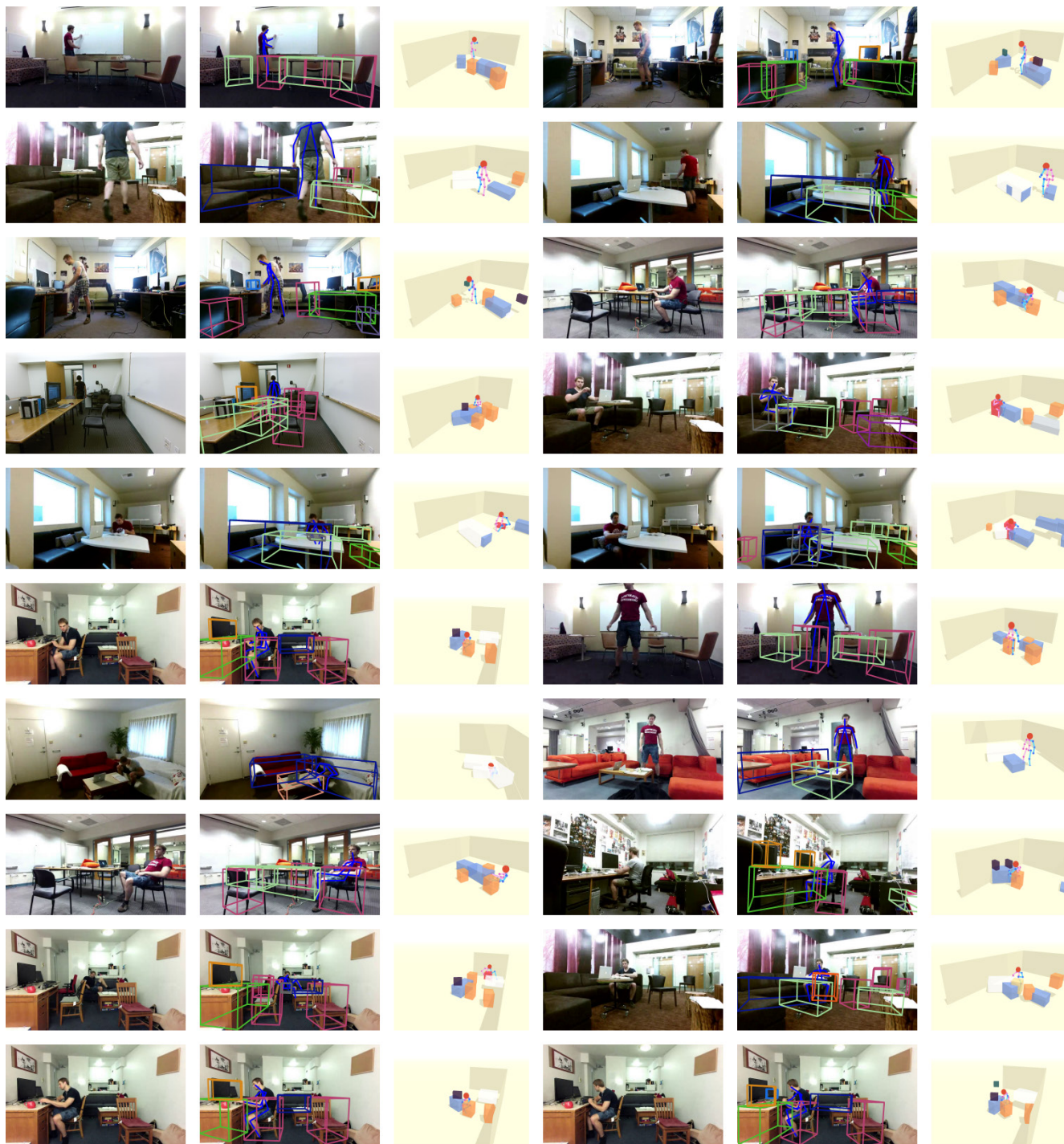


Figure 5. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

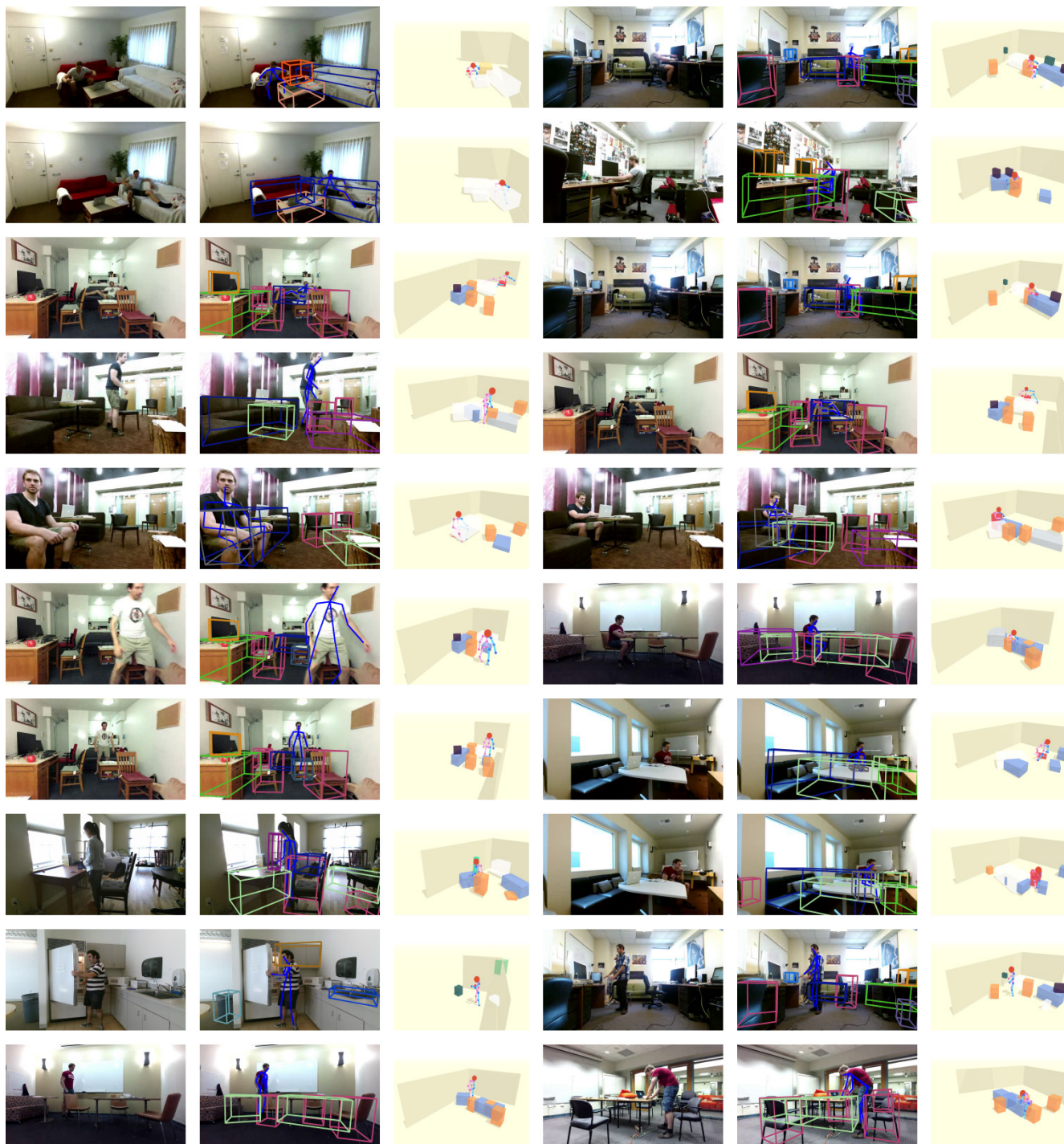


Figure 6. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

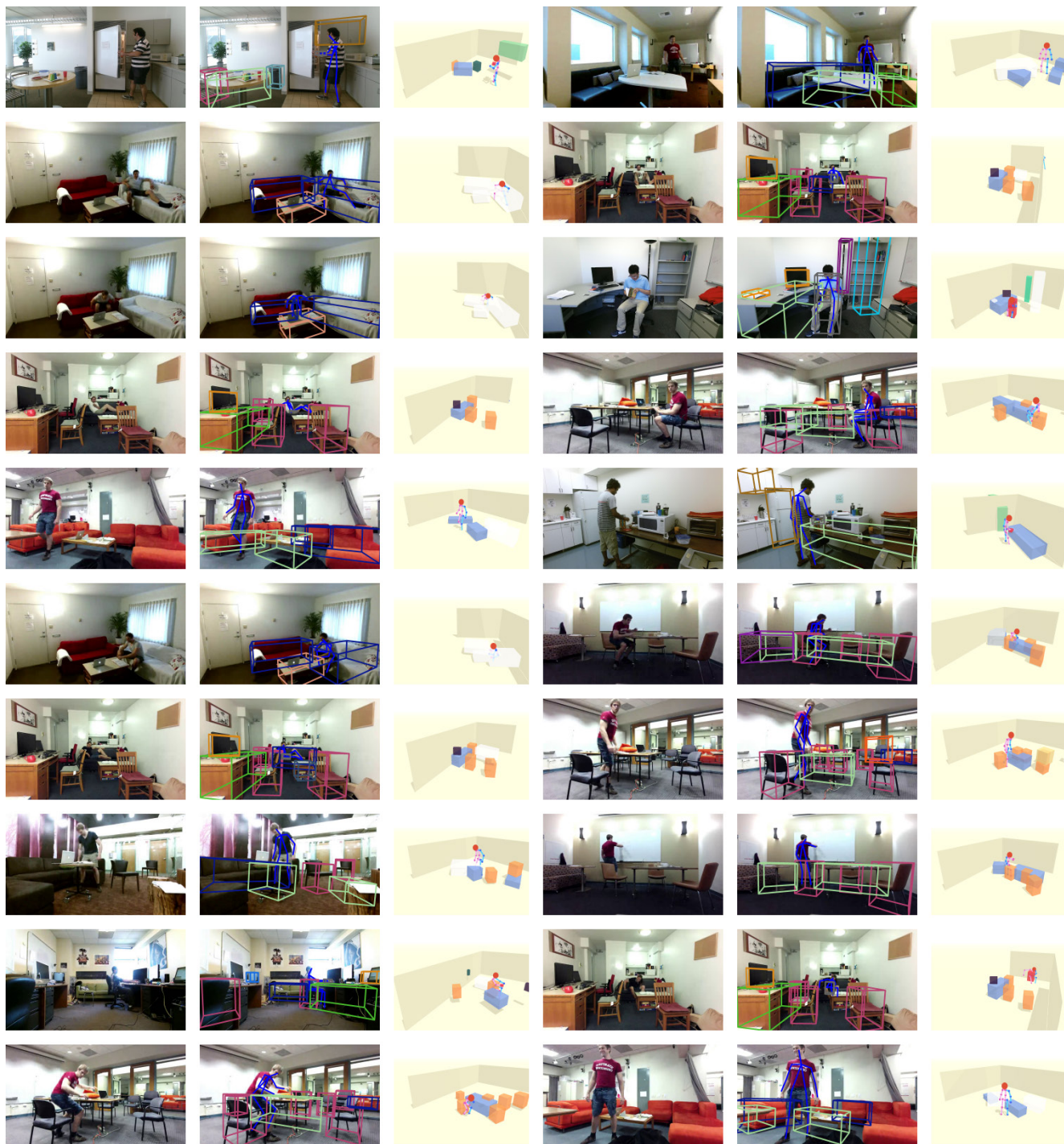


Figure 7. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

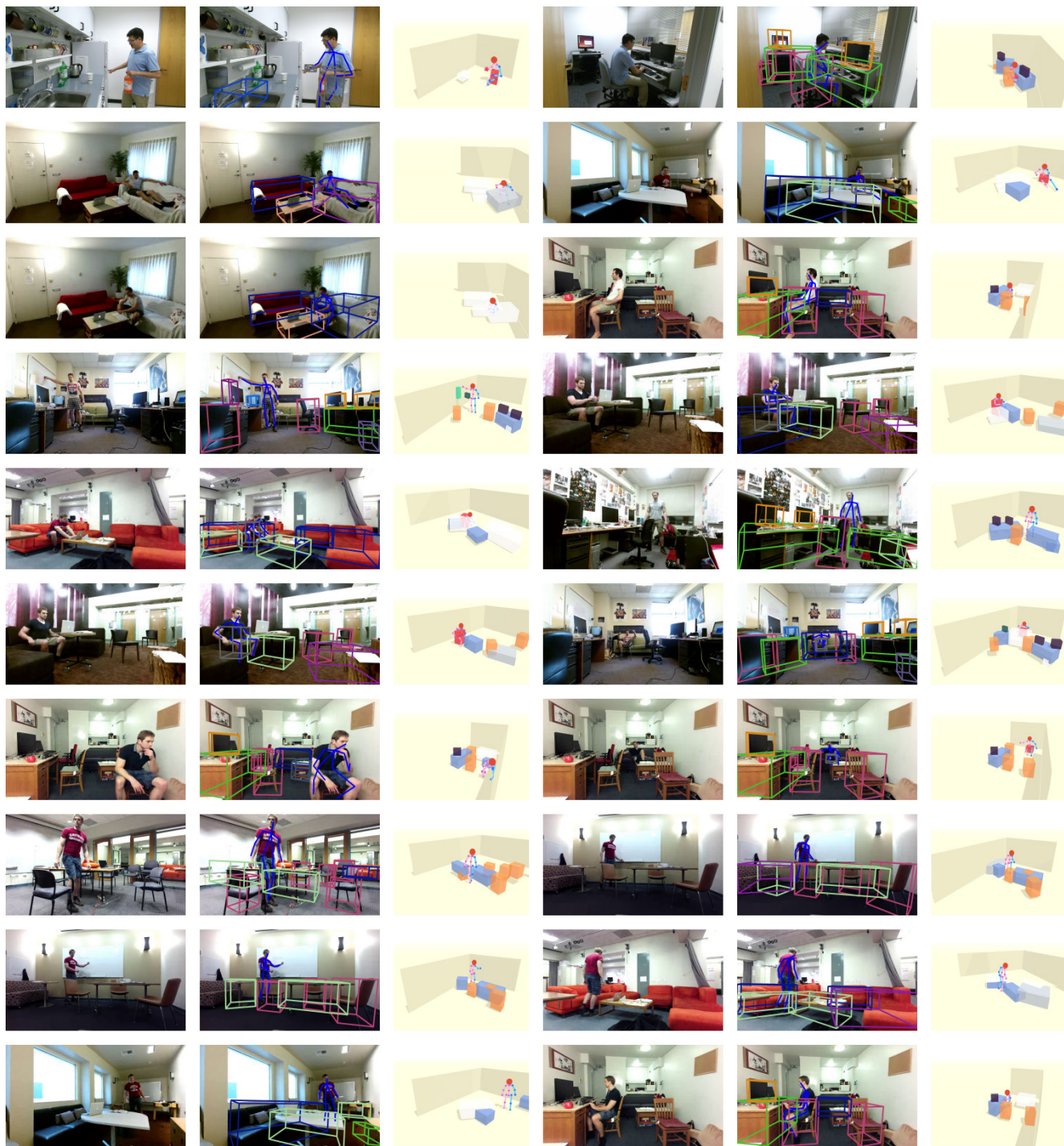


Figure 8. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

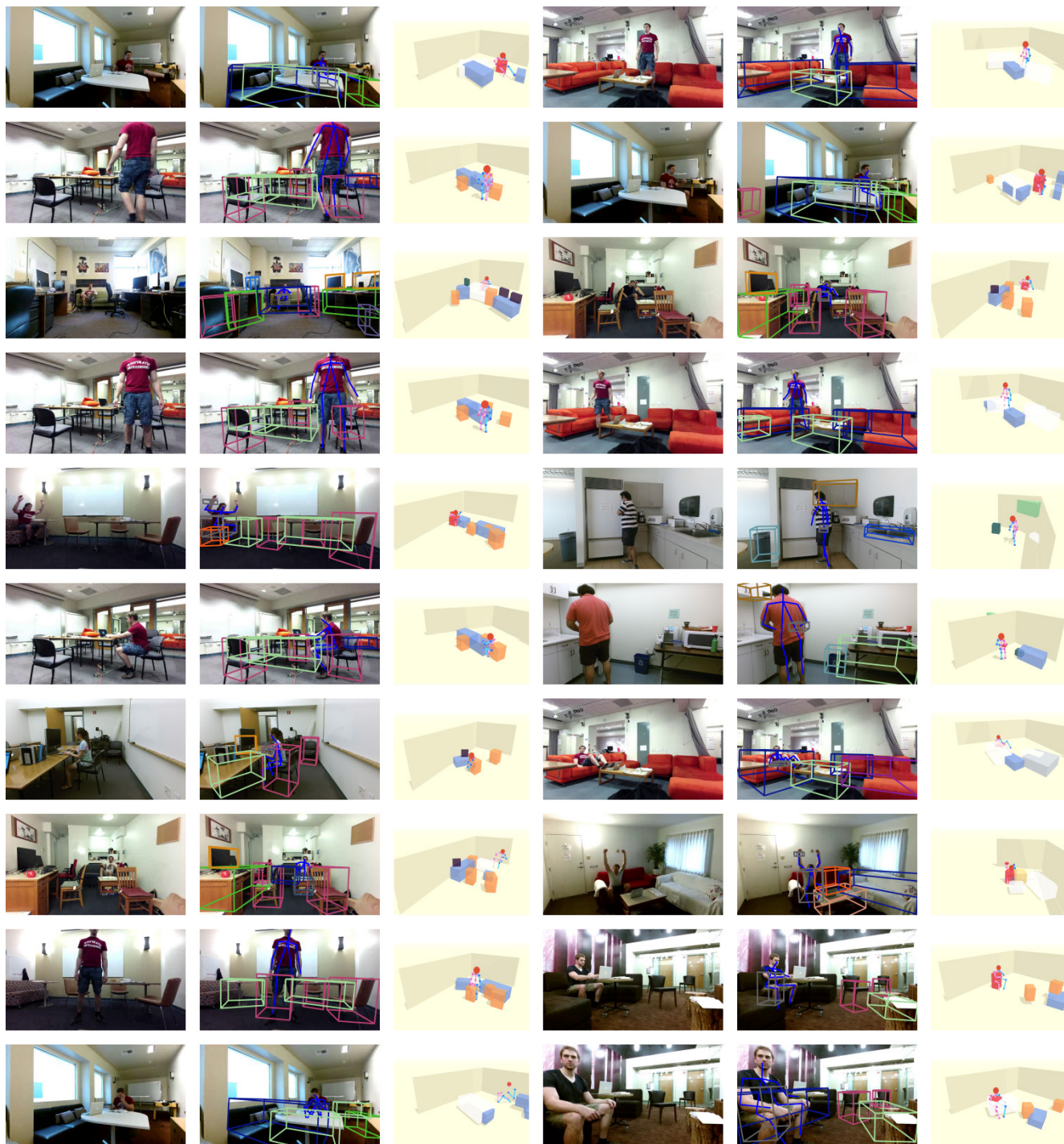


Figure 9. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

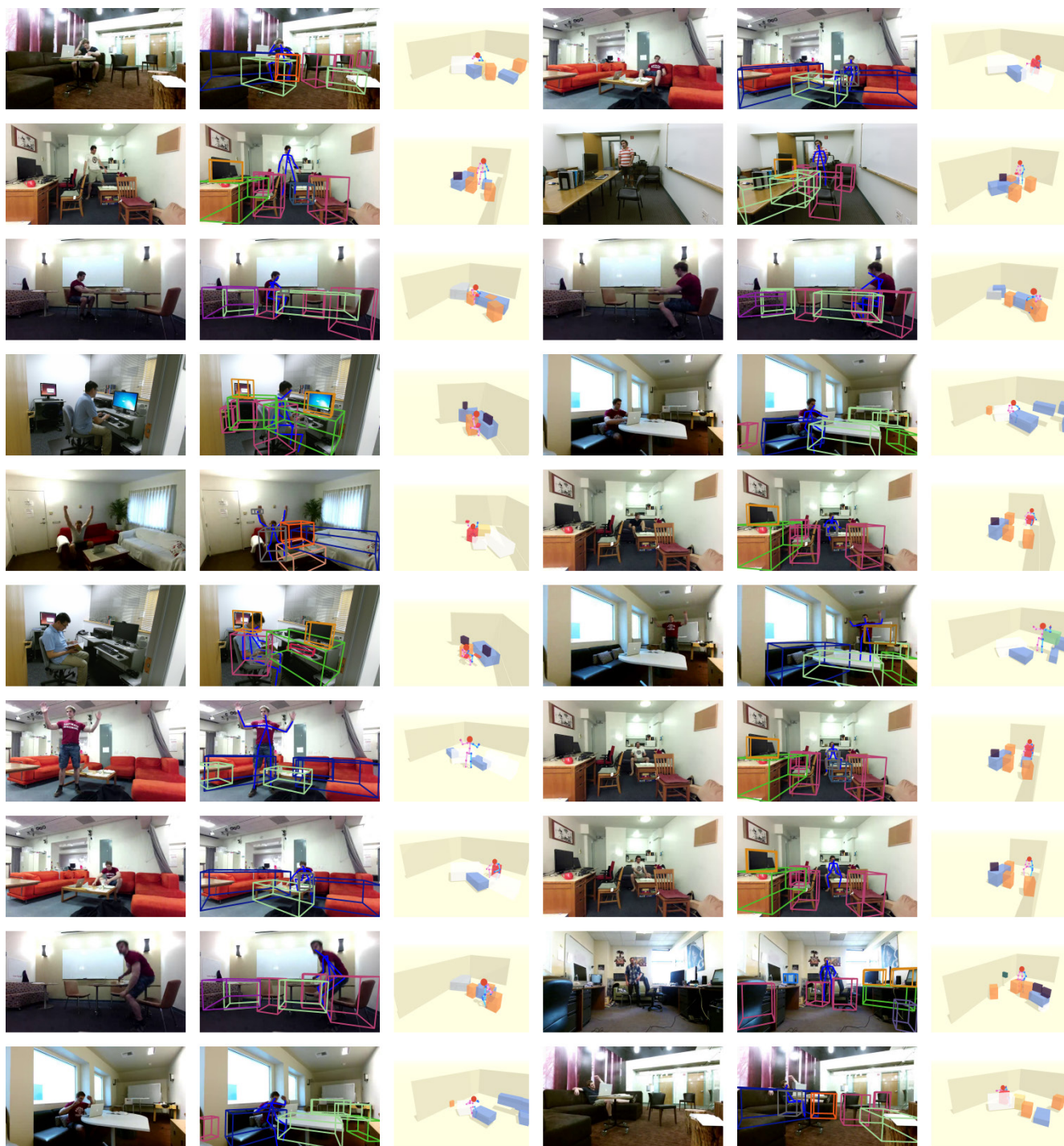


Figure 10. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

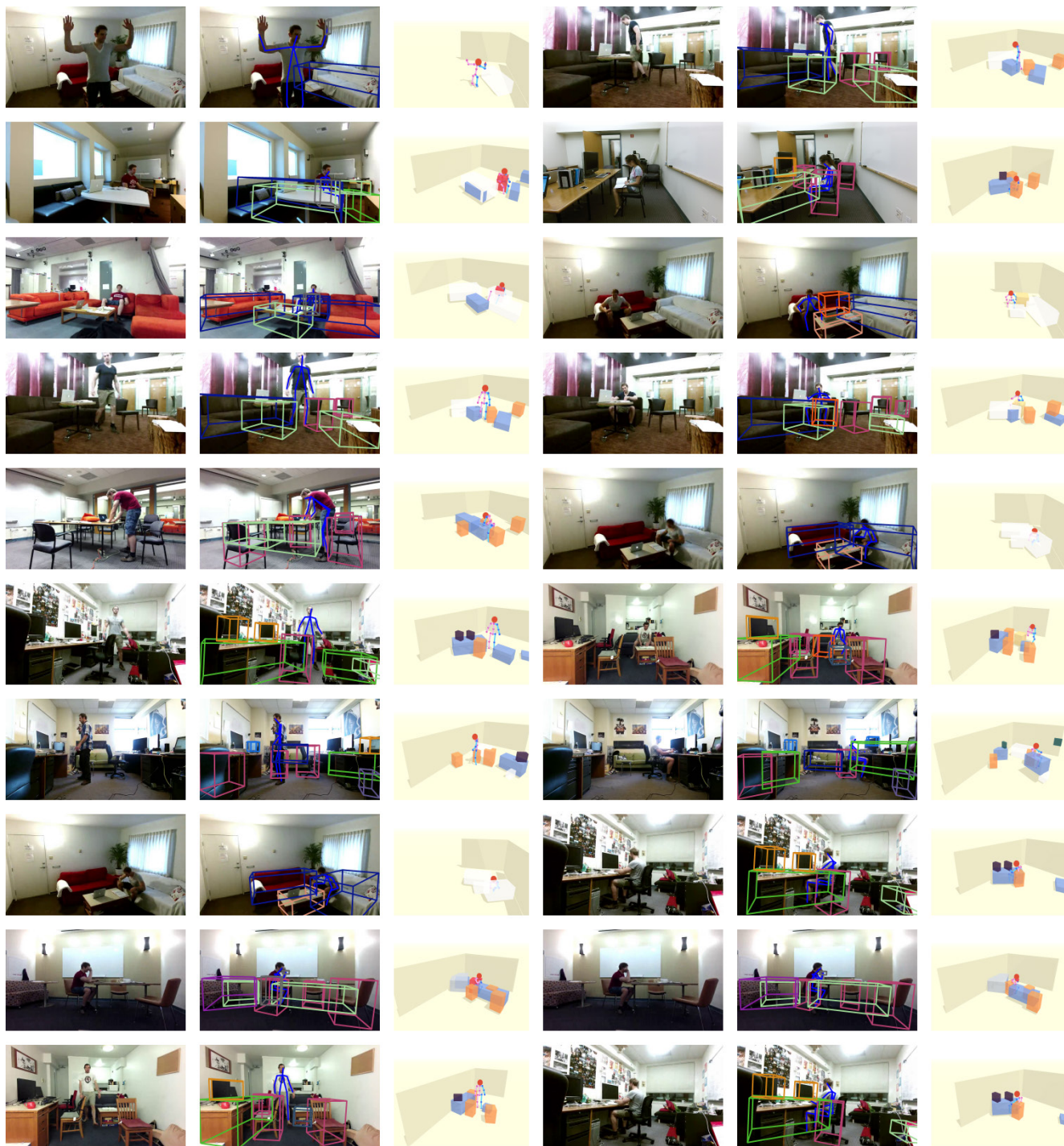


Figure 11. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

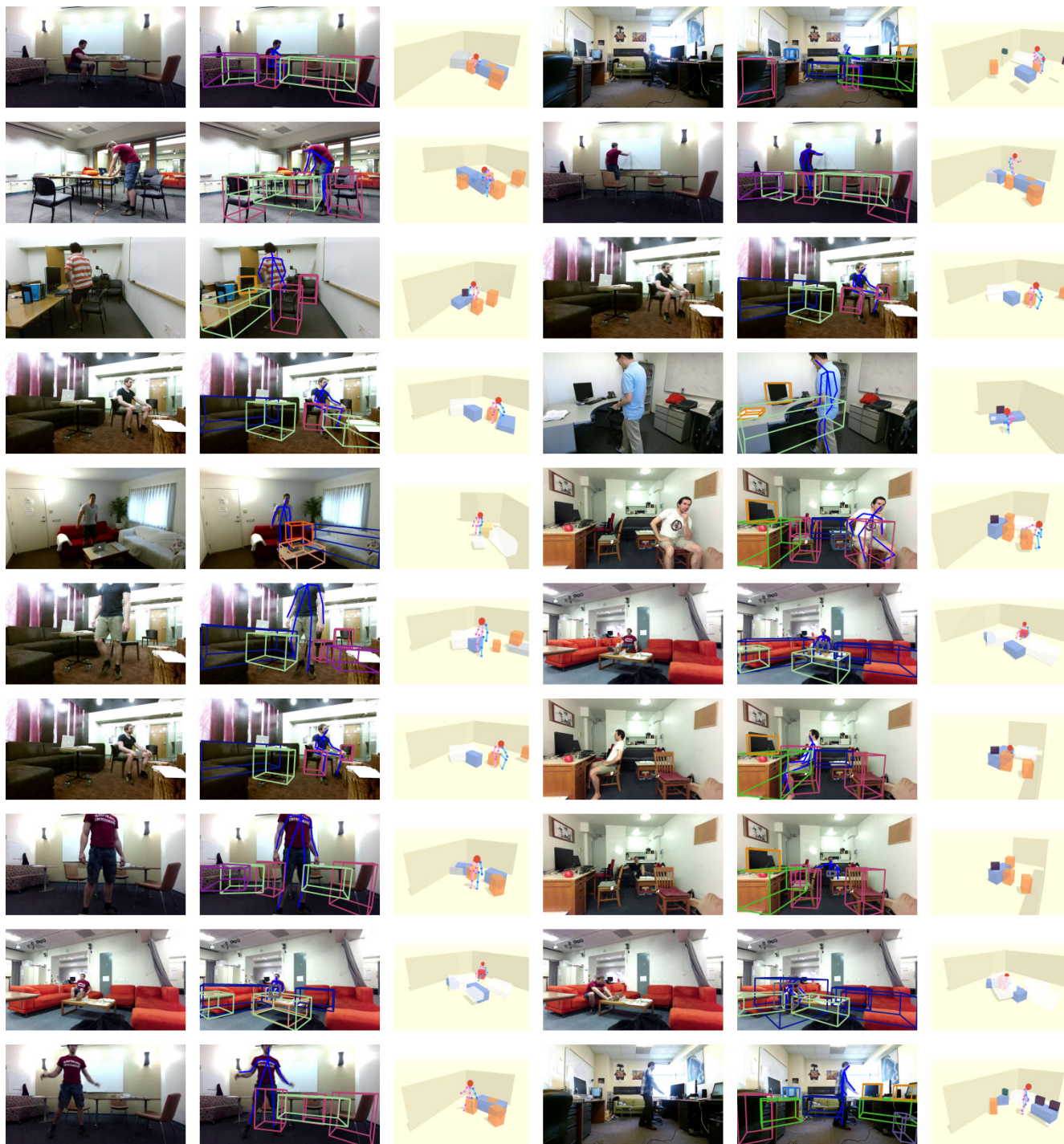


Figure 12. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

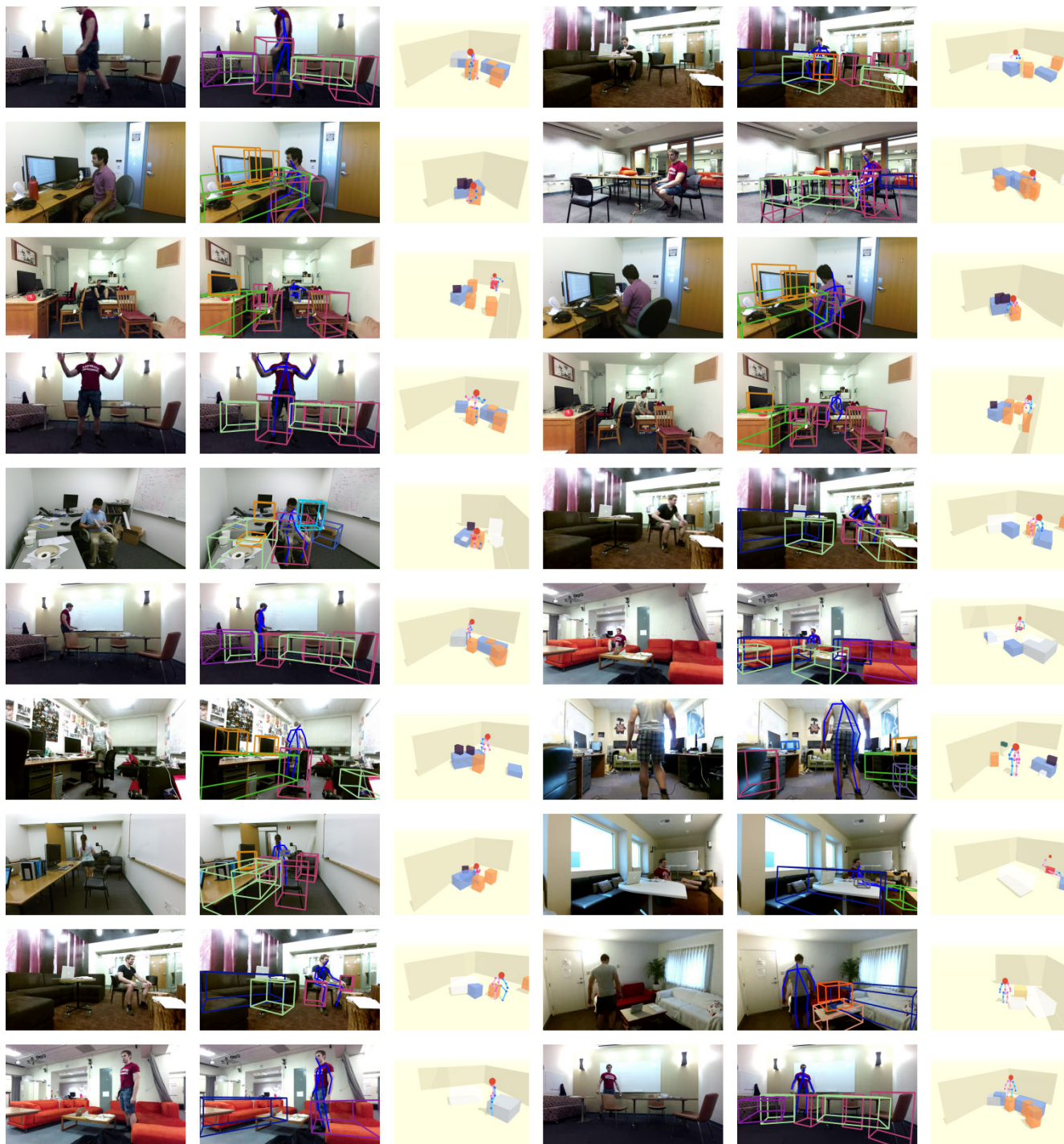


Figure 13. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

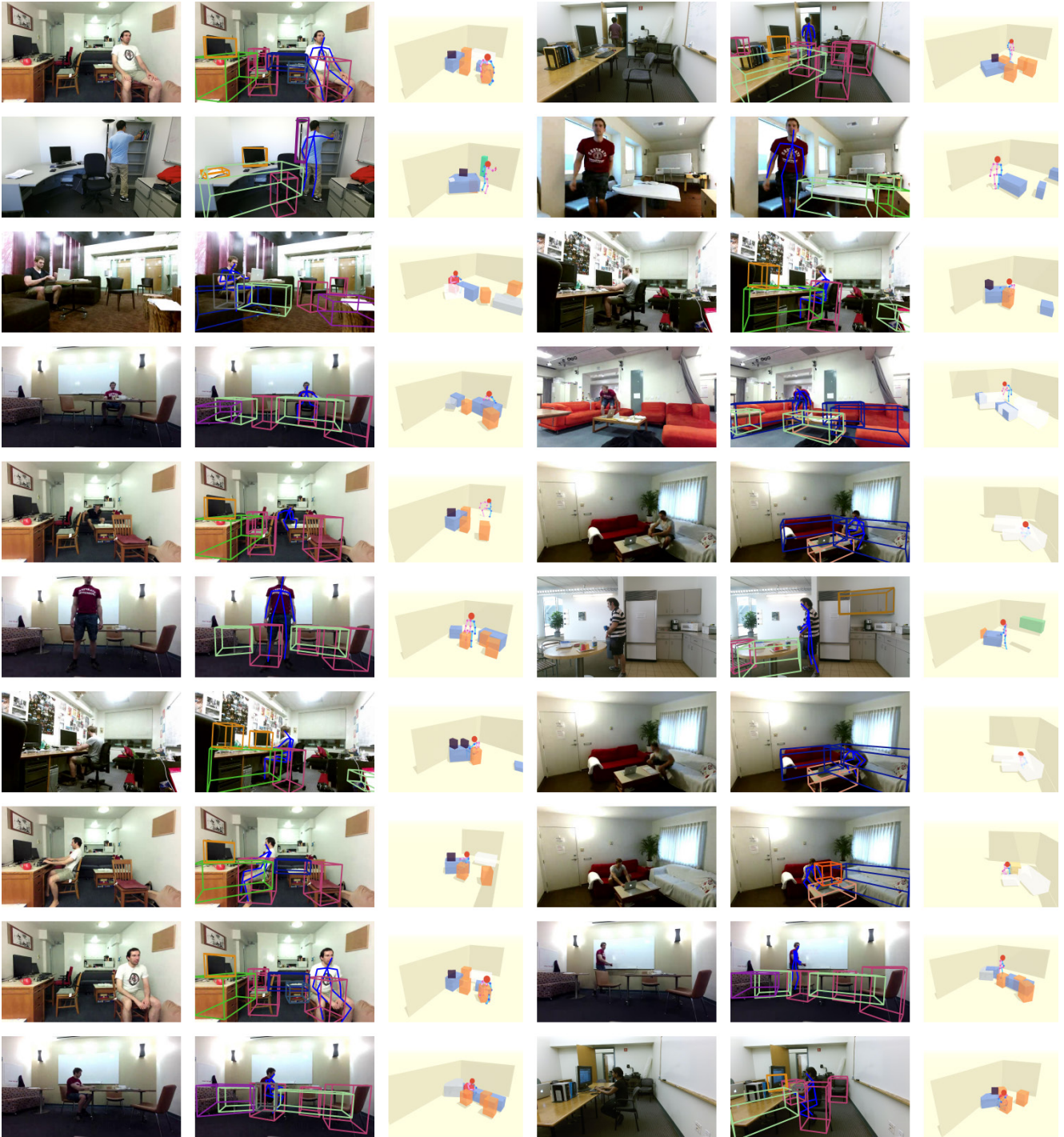


Figure 14. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

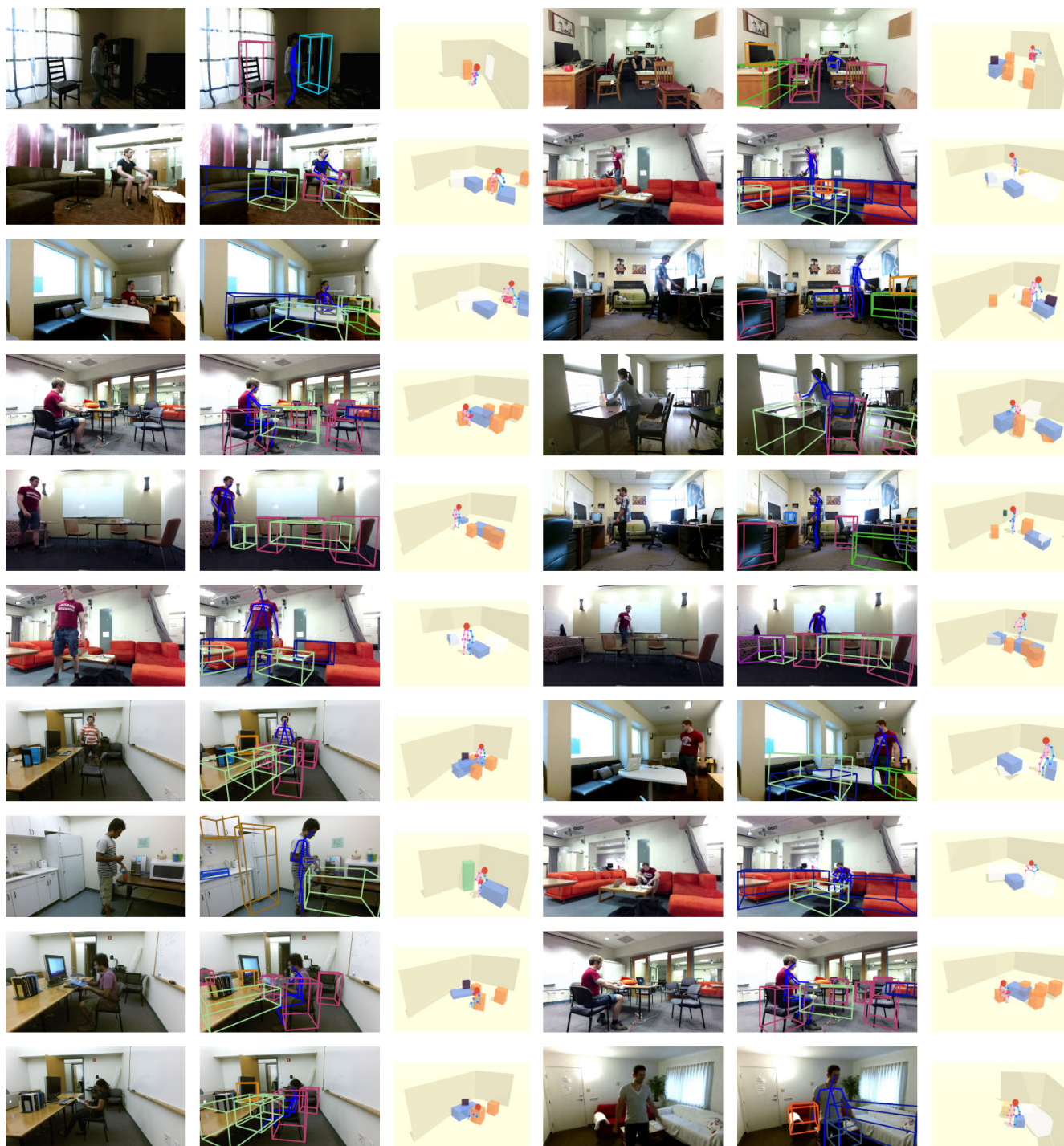


Figure 15. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

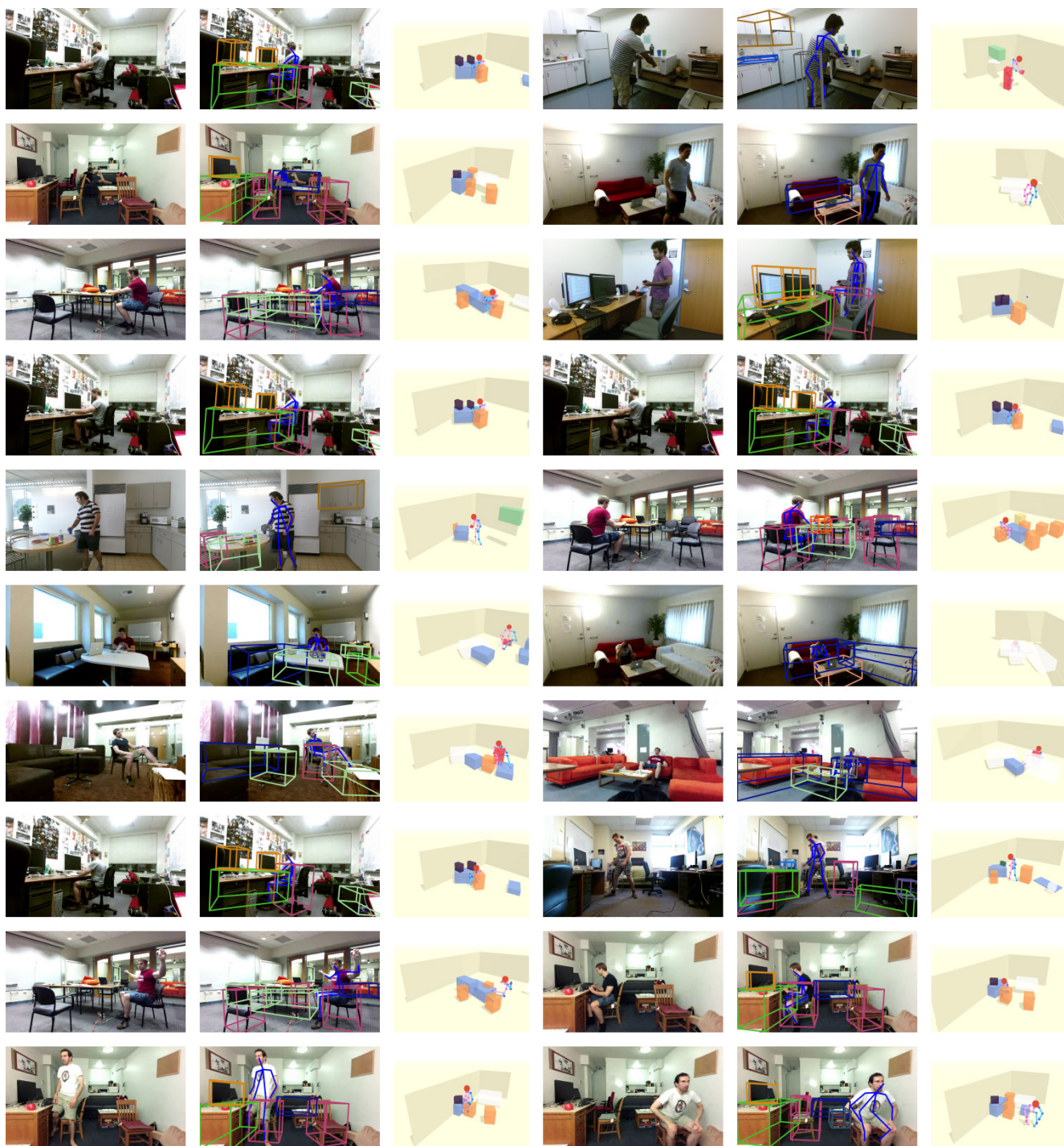


Figure 16. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

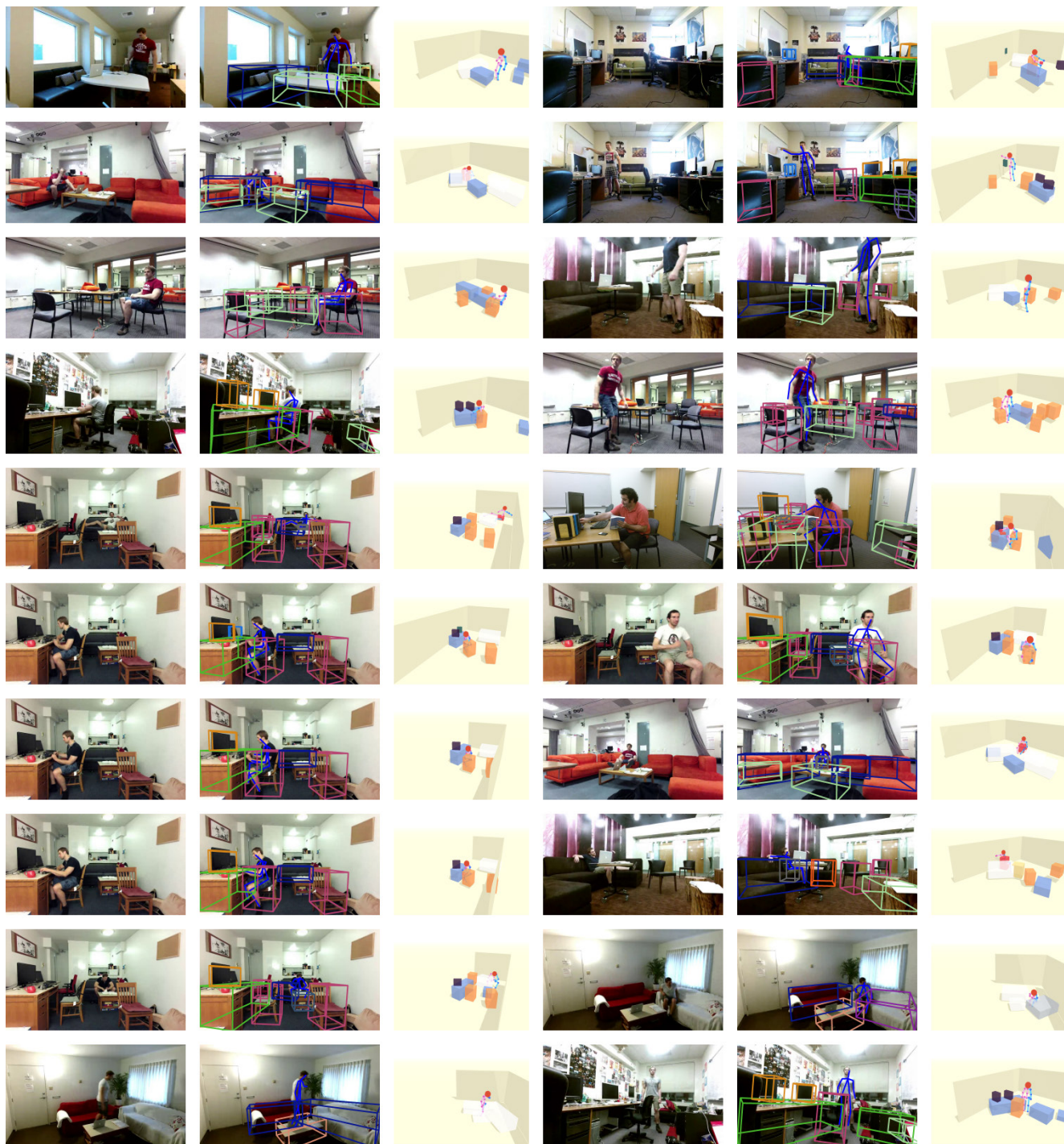


Figure 17. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

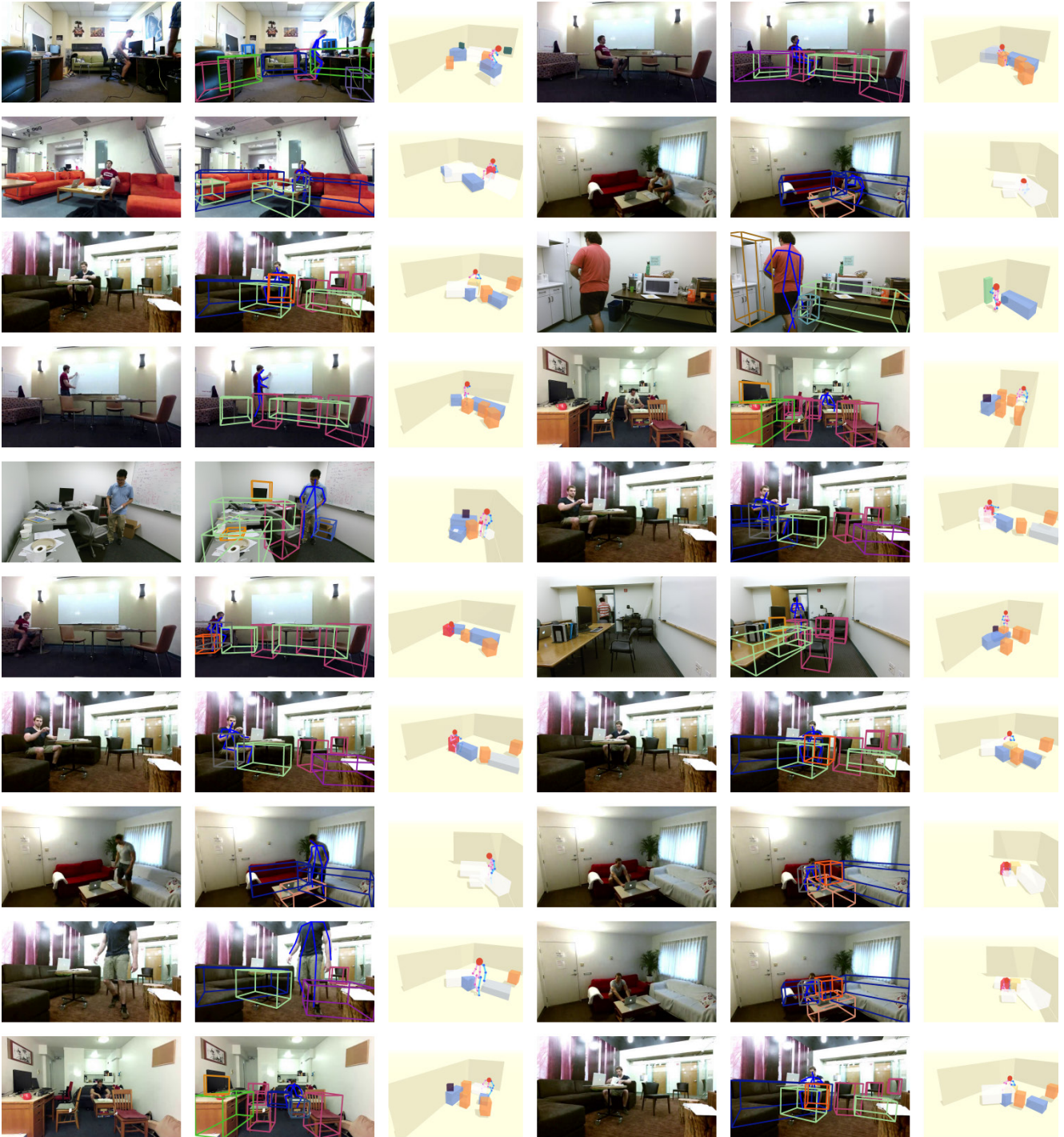


Figure 18. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

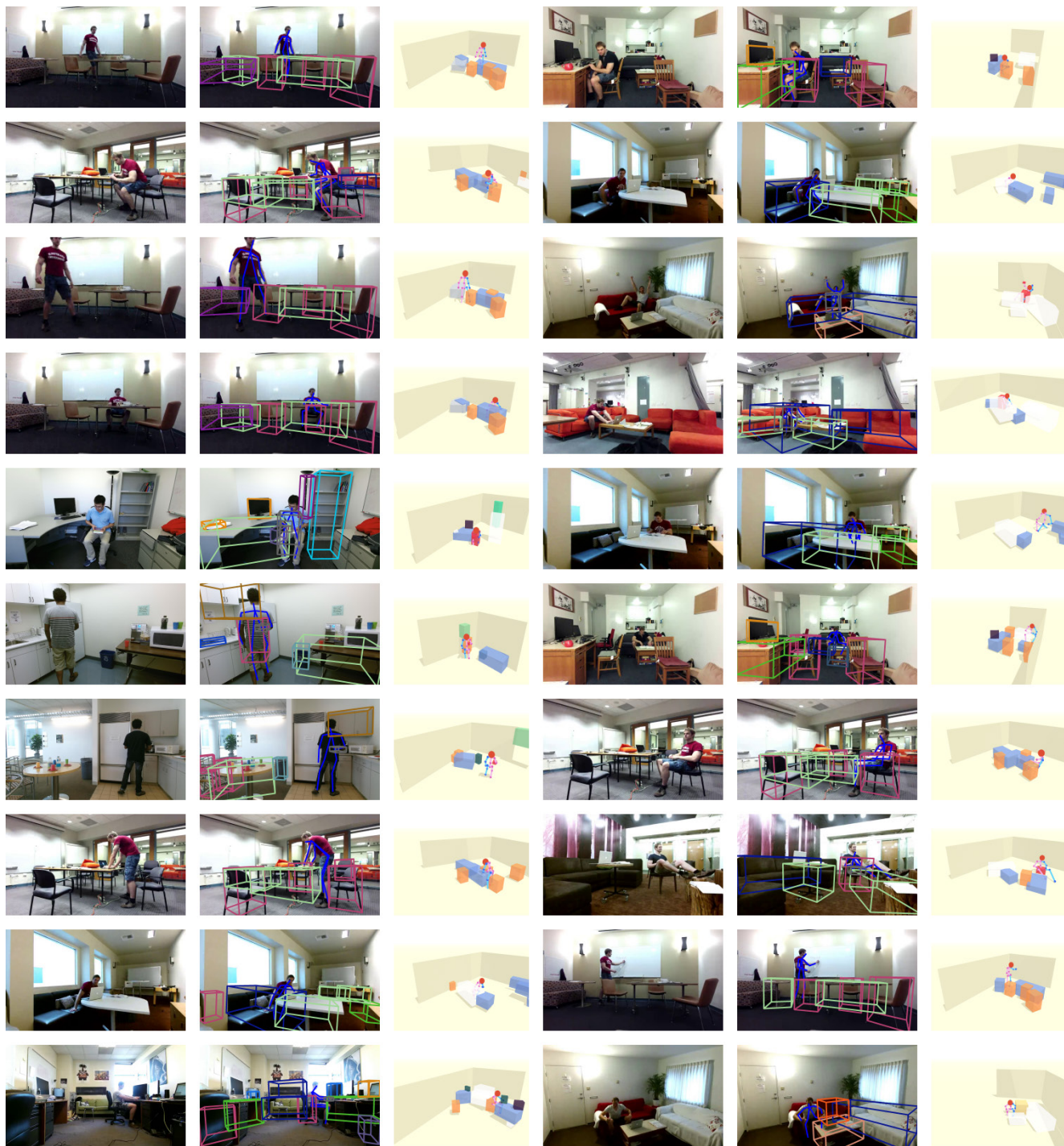


Figure 19. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

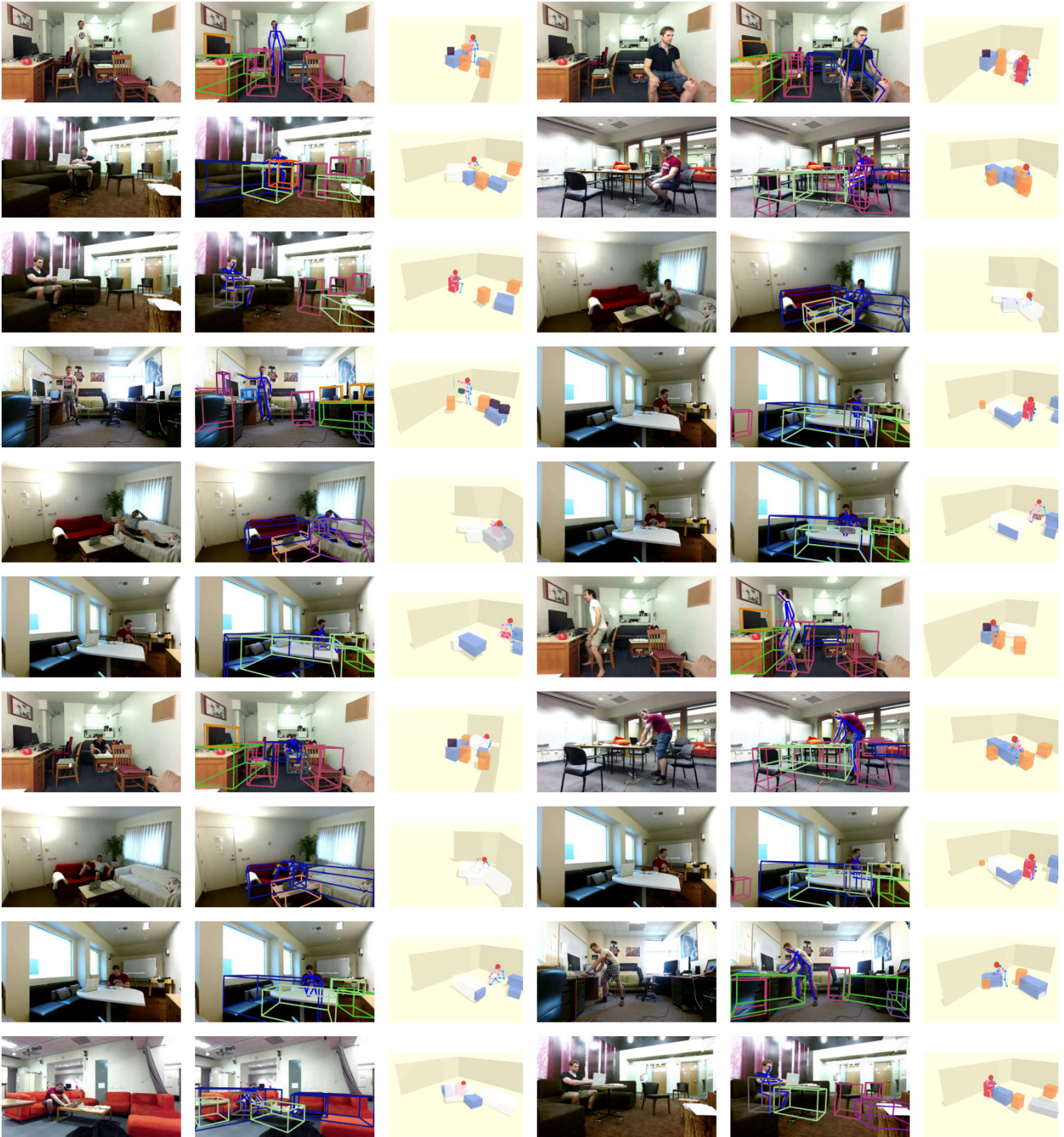


Figure 20. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

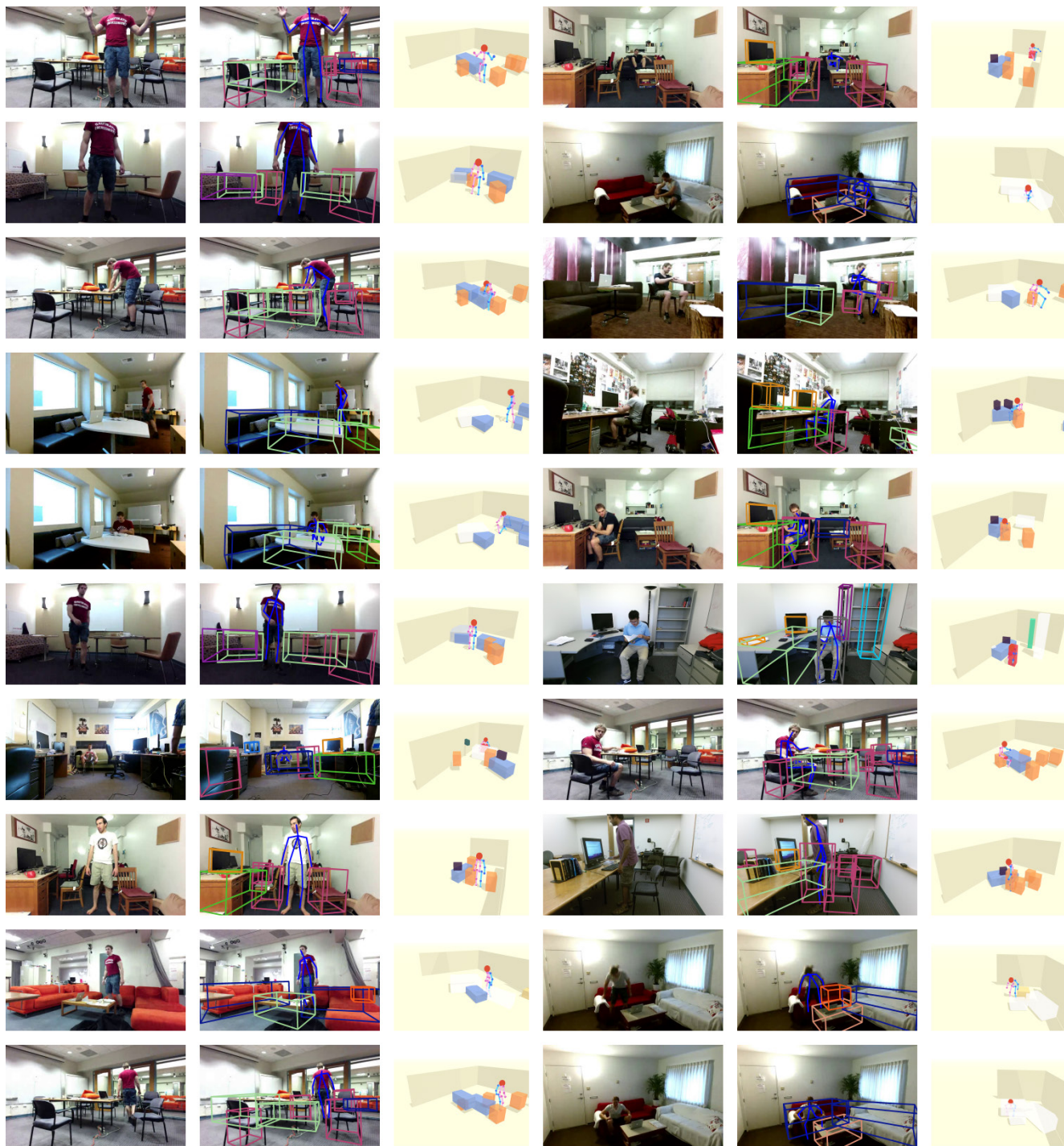


Figure 21. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

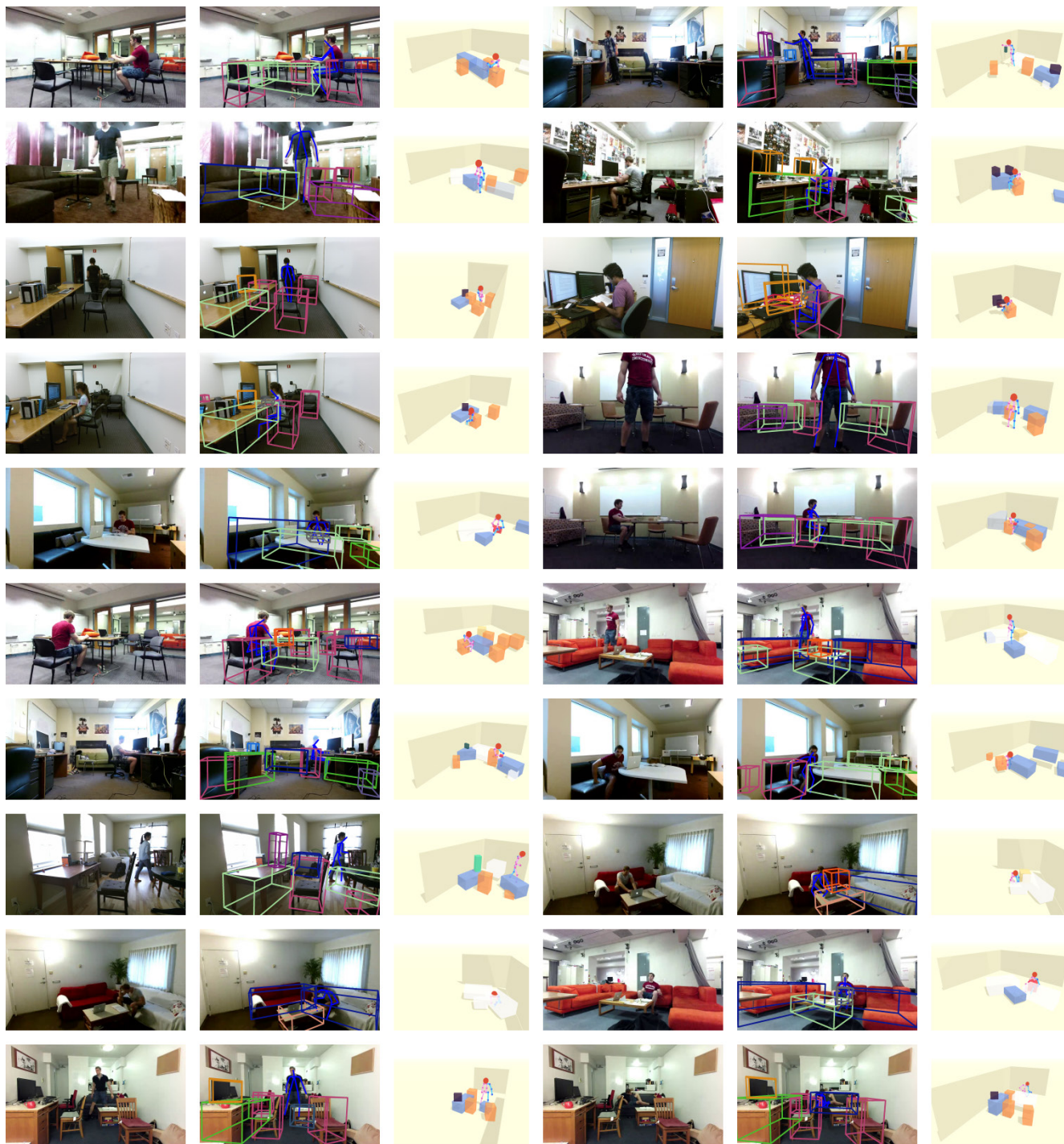


Figure 22. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

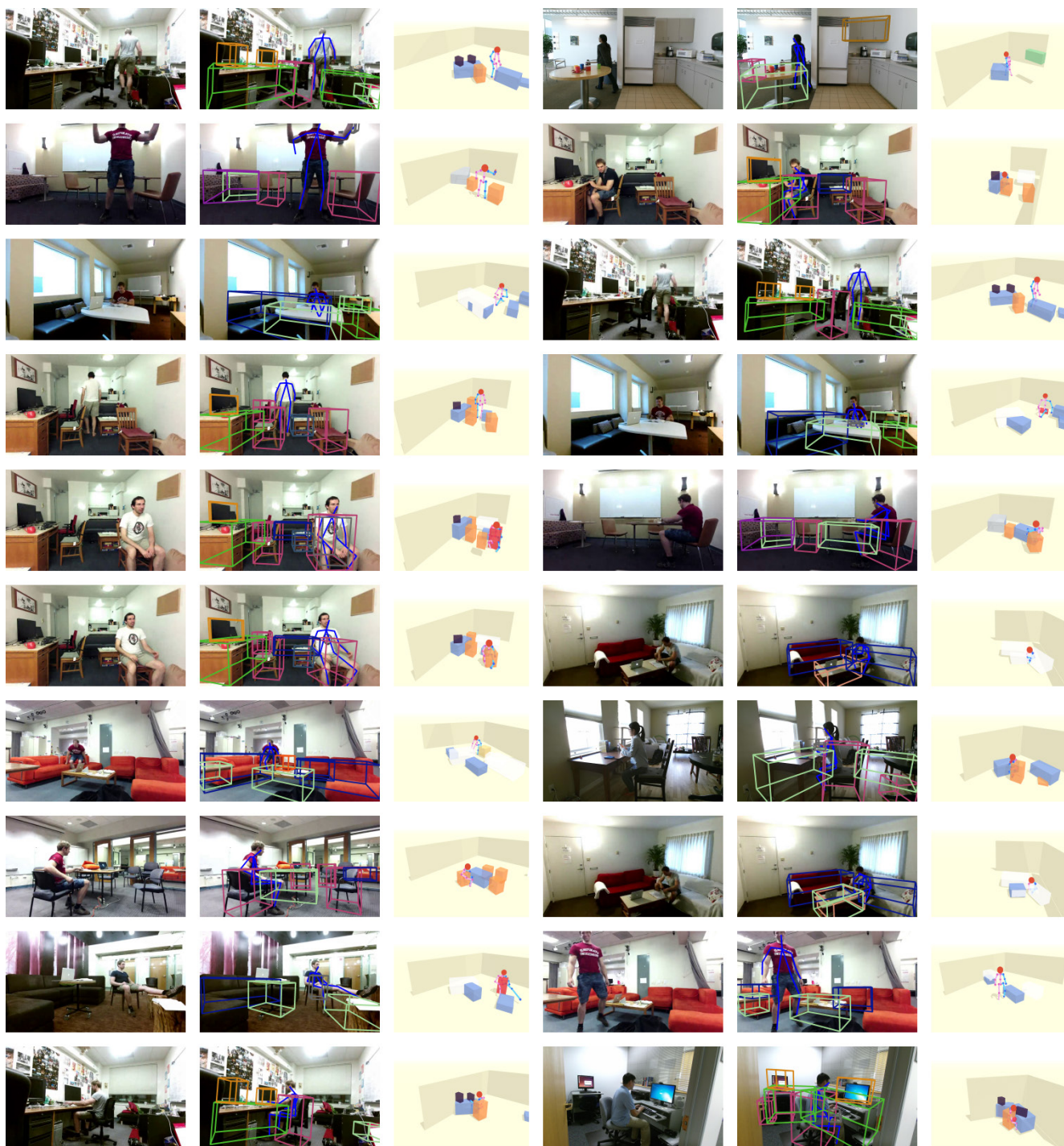


Figure 23. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

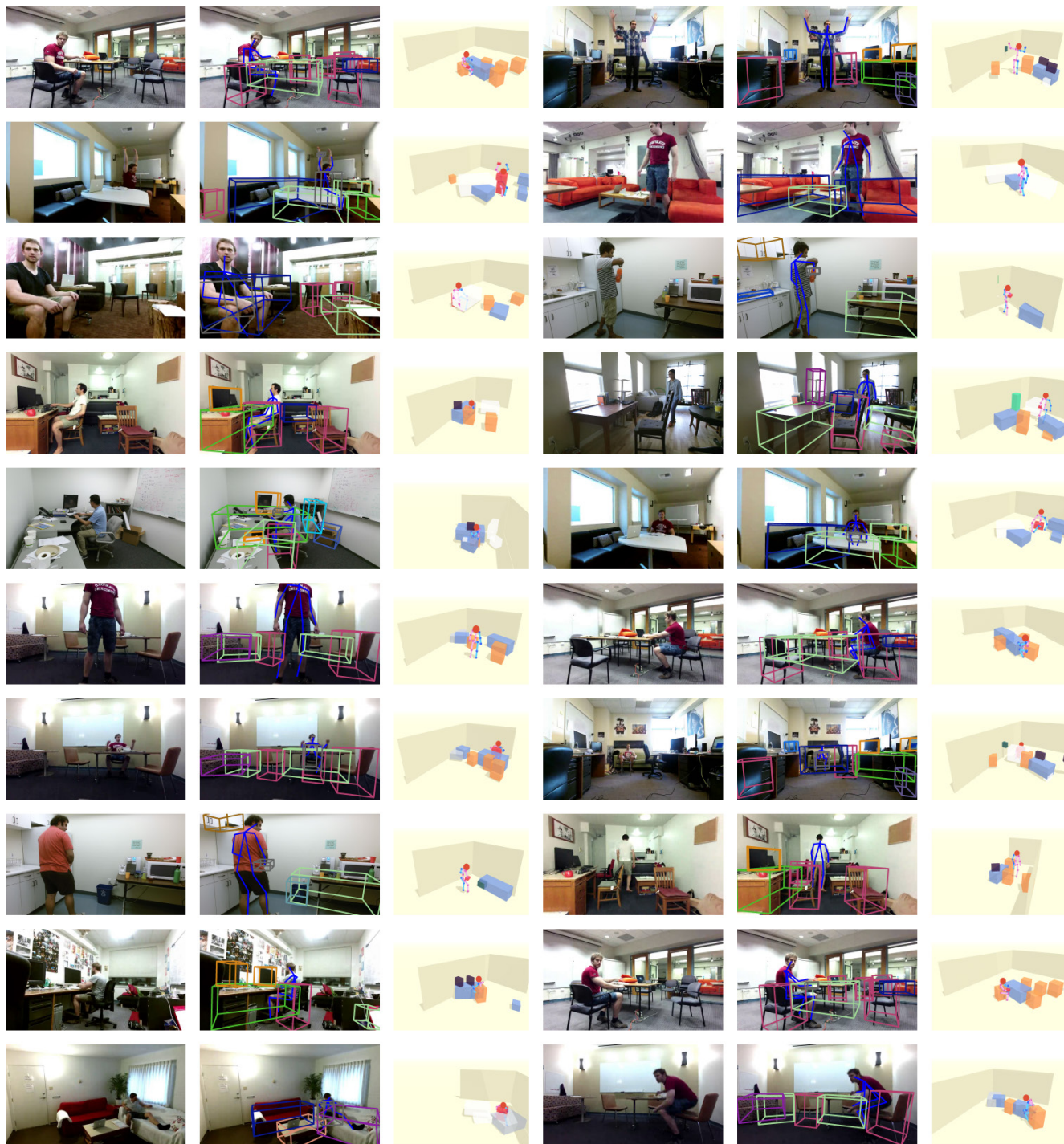


Figure 24. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

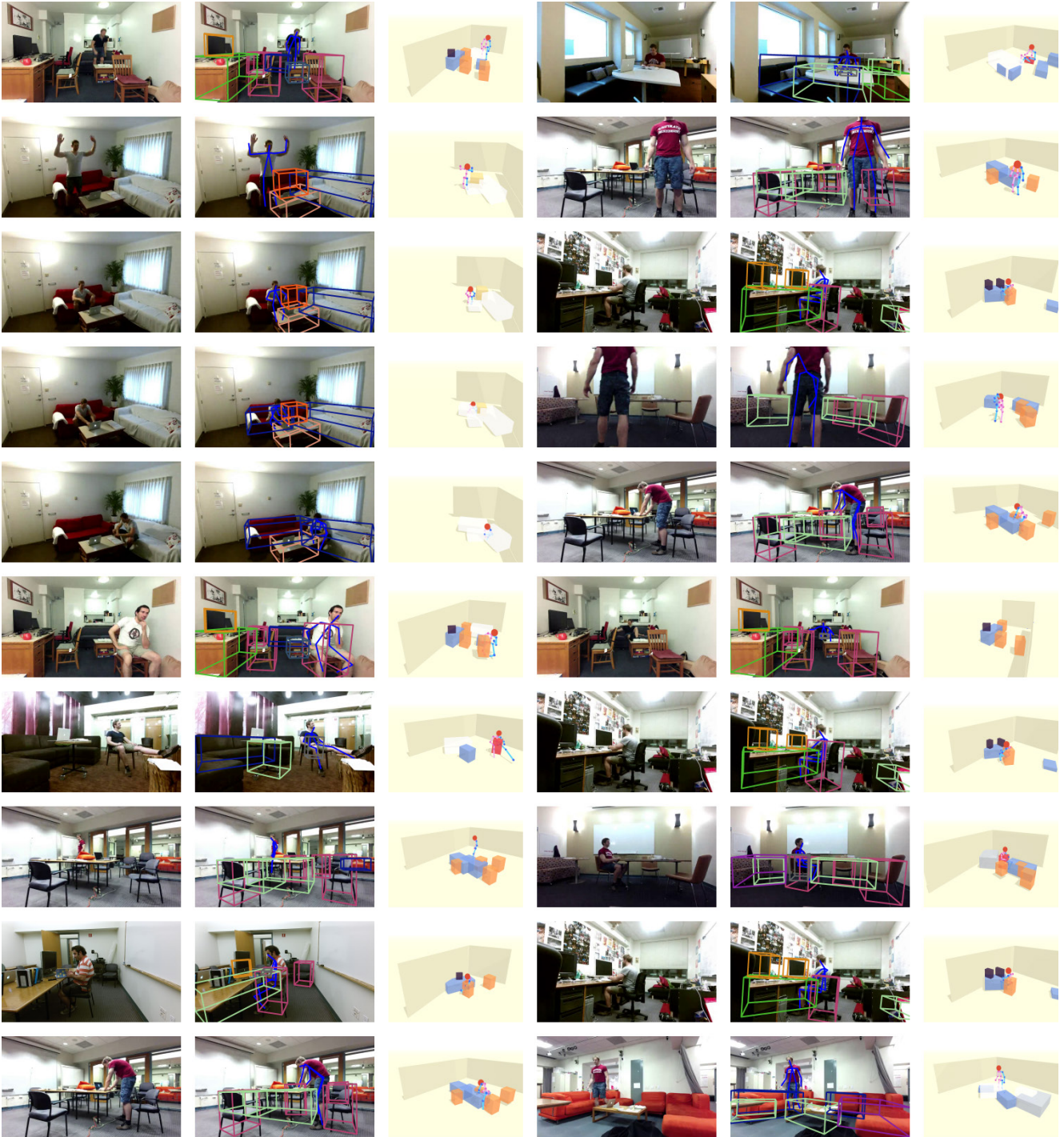


Figure 25. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

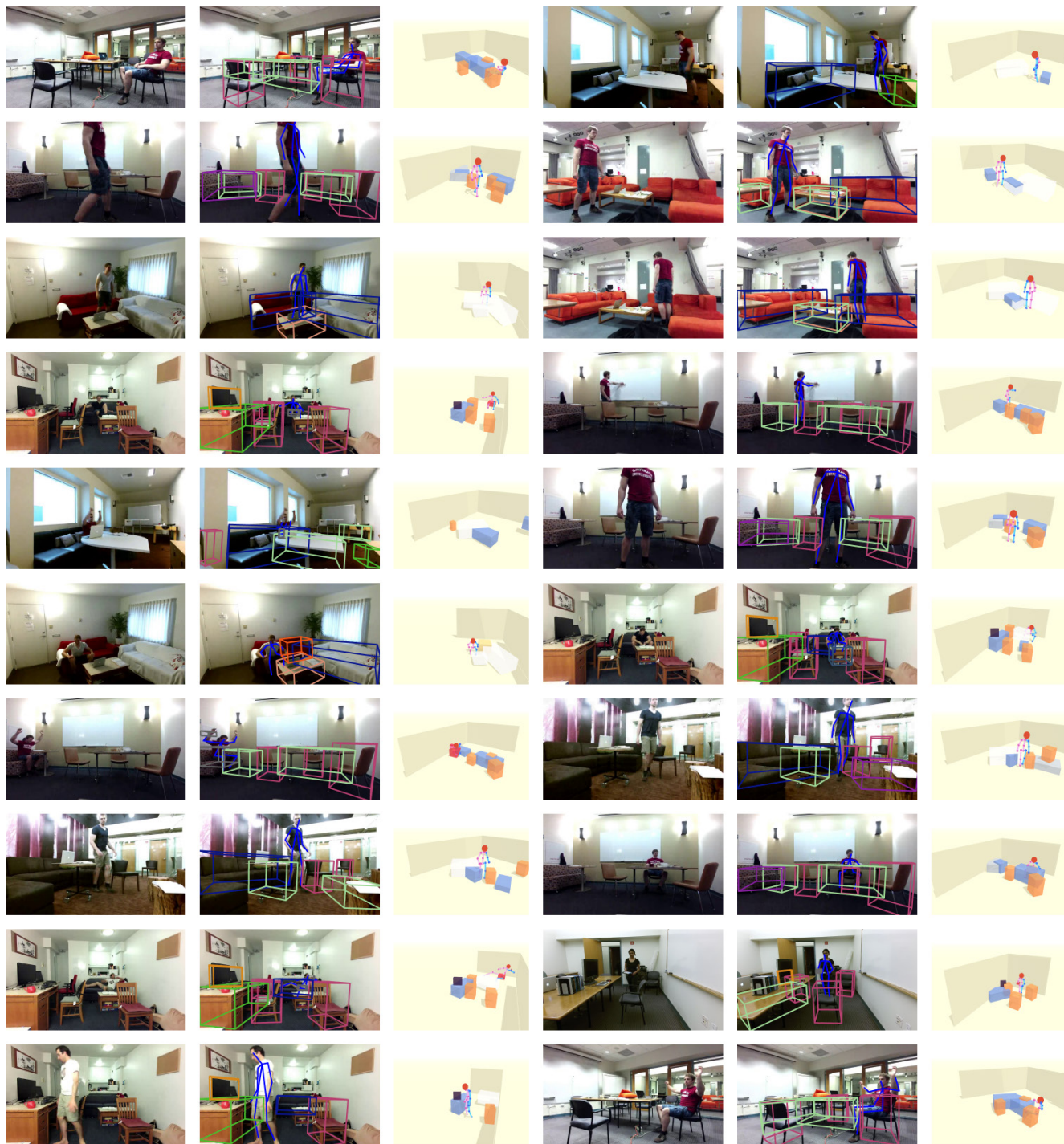


Figure 26. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

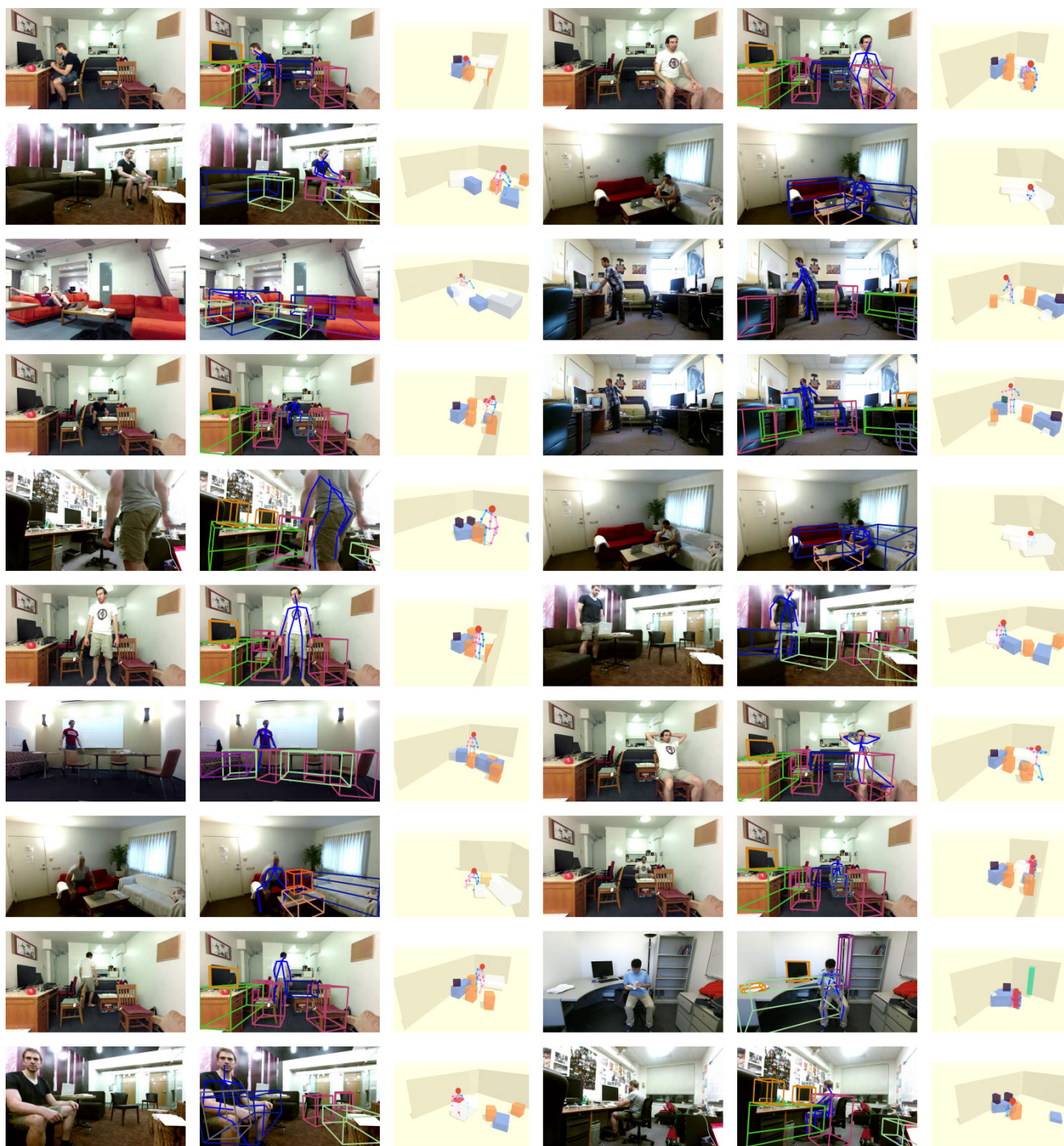


Figure 27. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

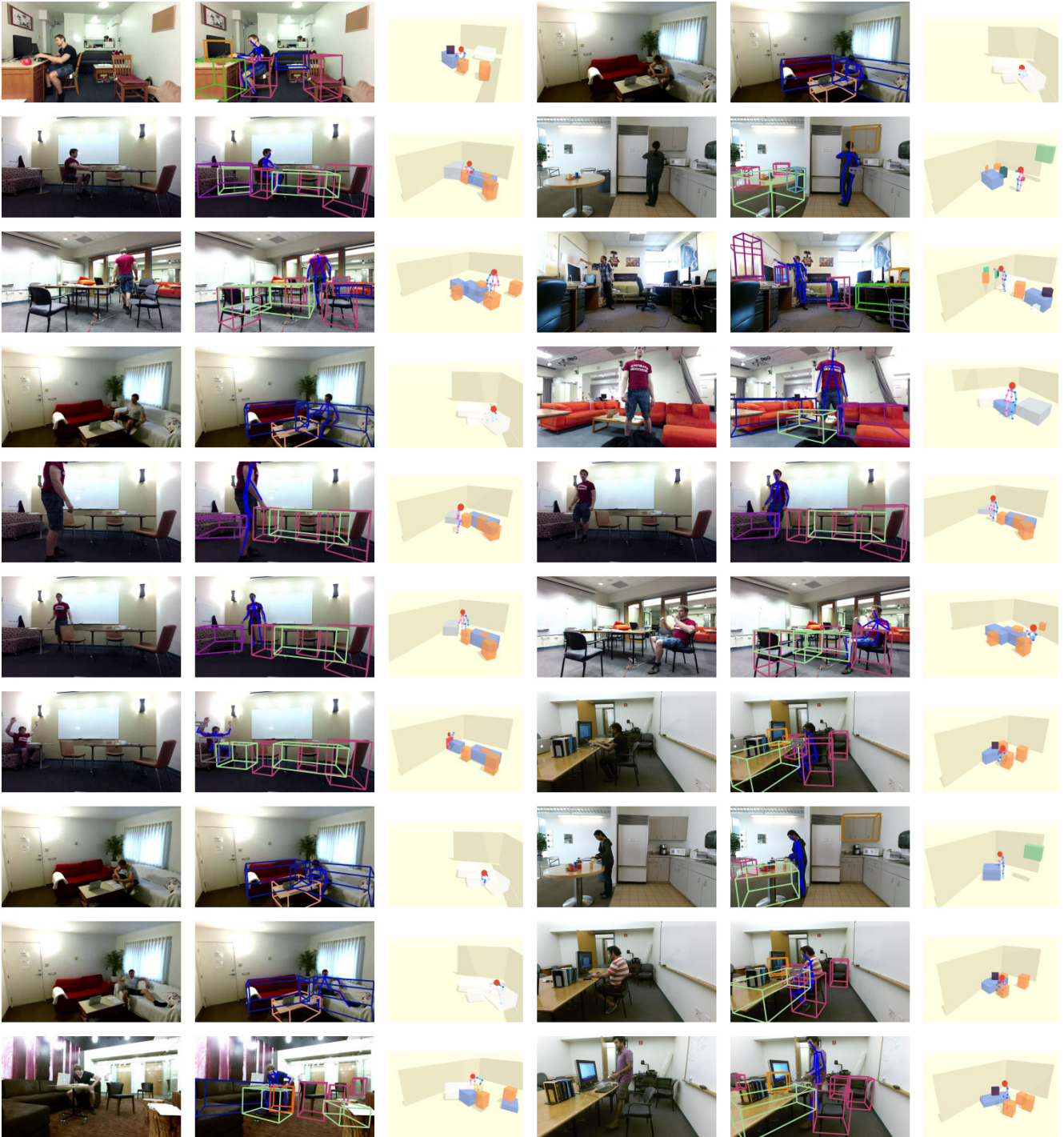


Figure 28. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

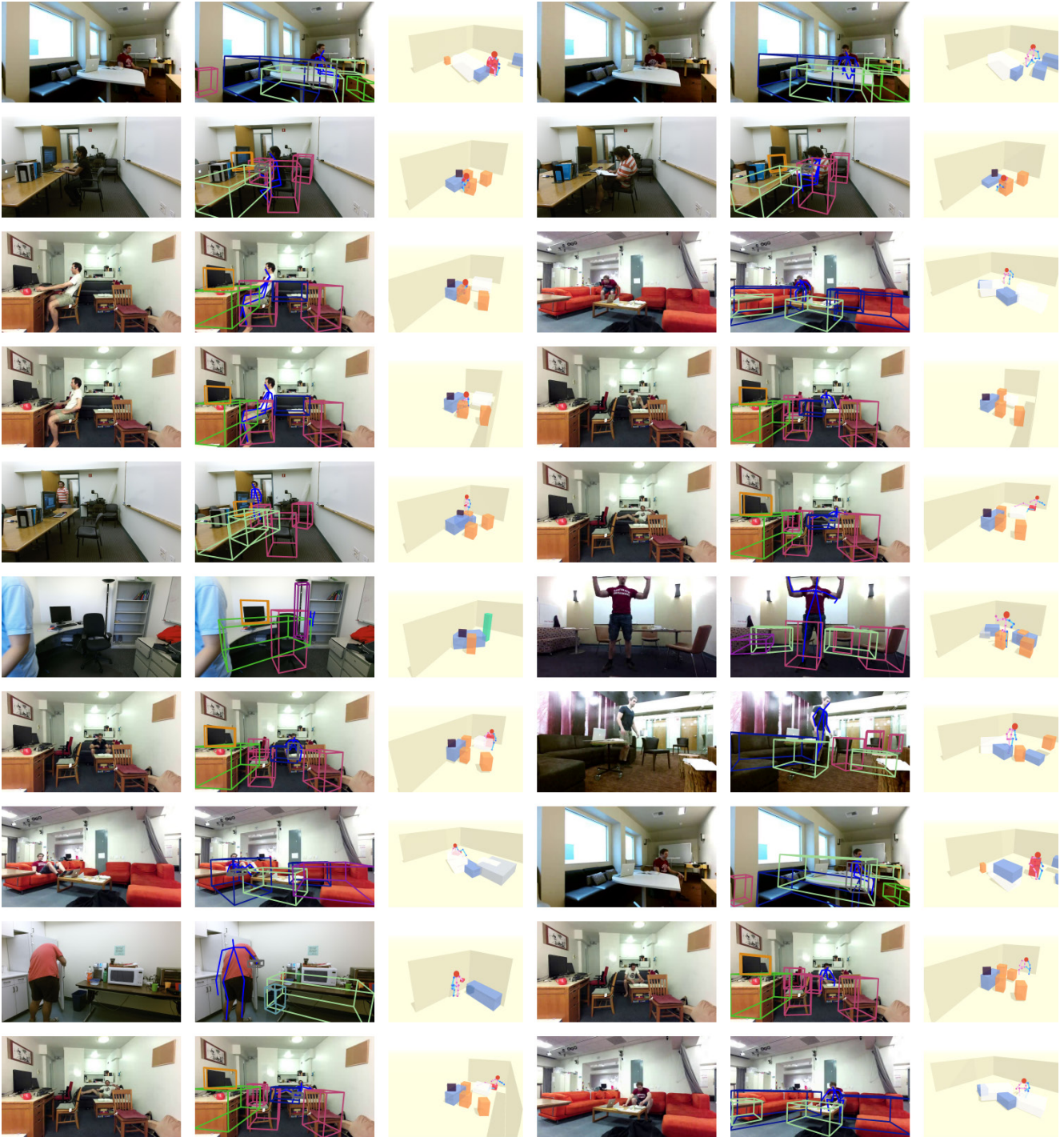


Figure 29. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

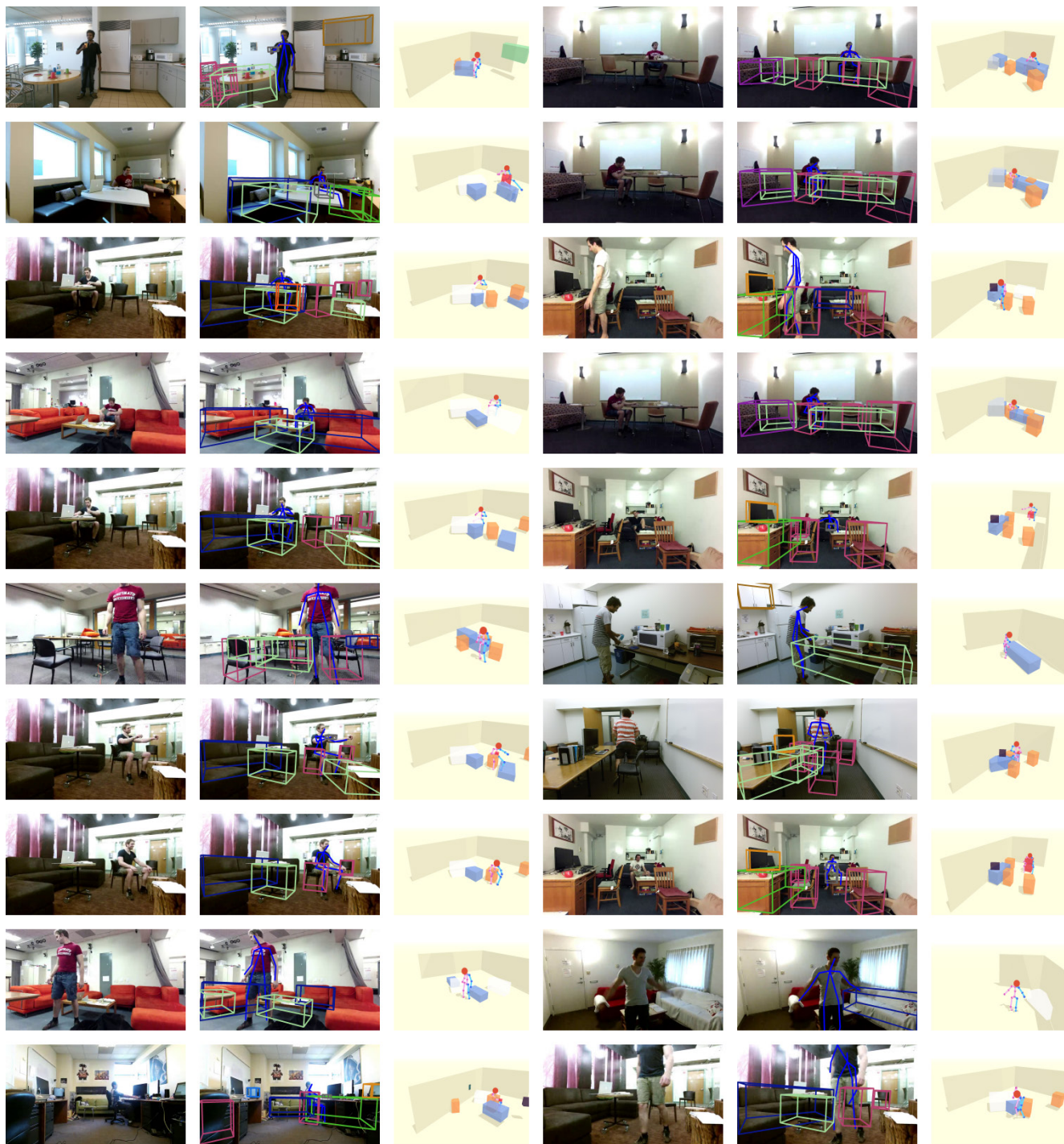


Figure 30. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

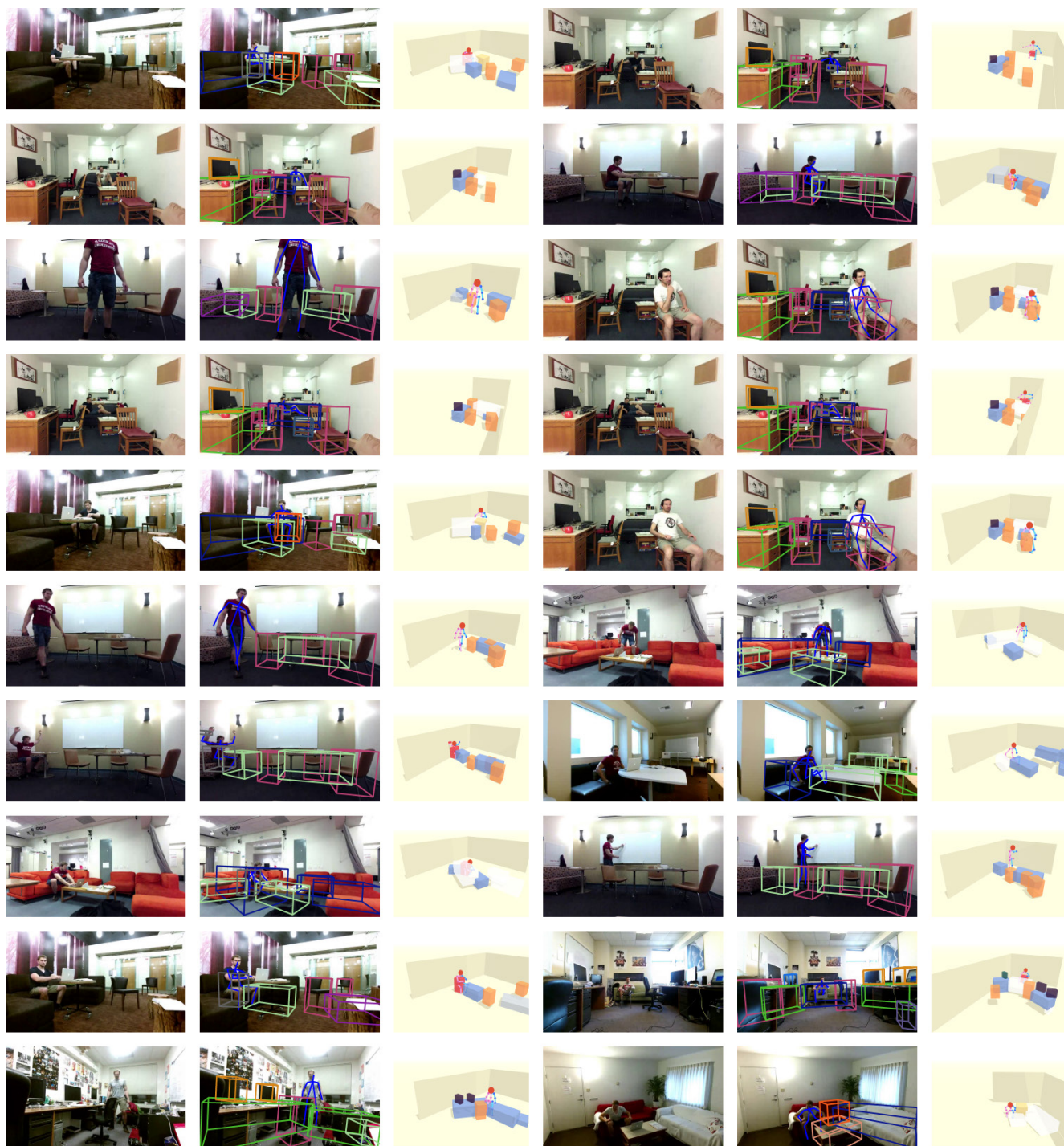


Figure 31. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

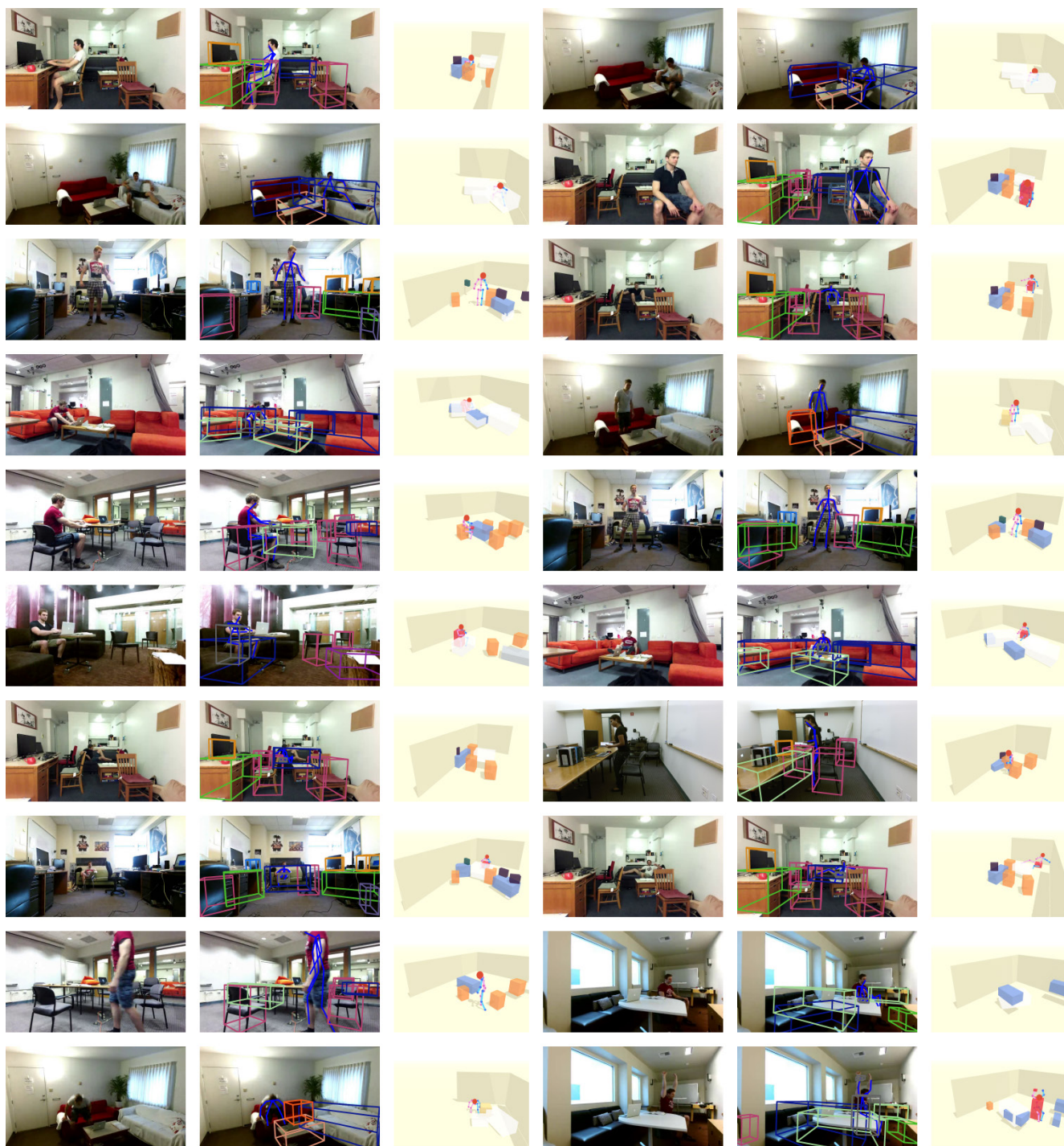


Figure 32. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

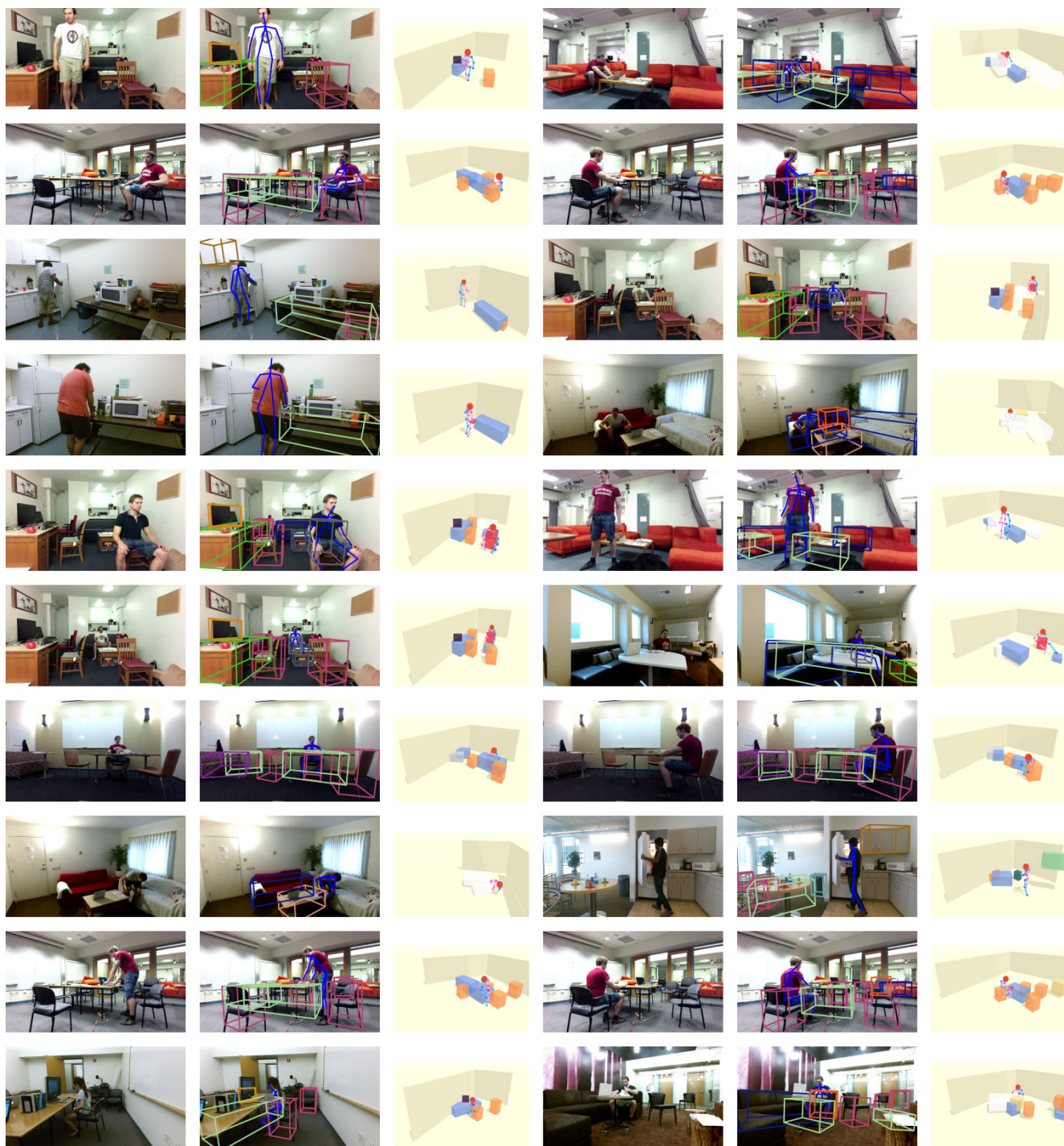


Figure 33. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

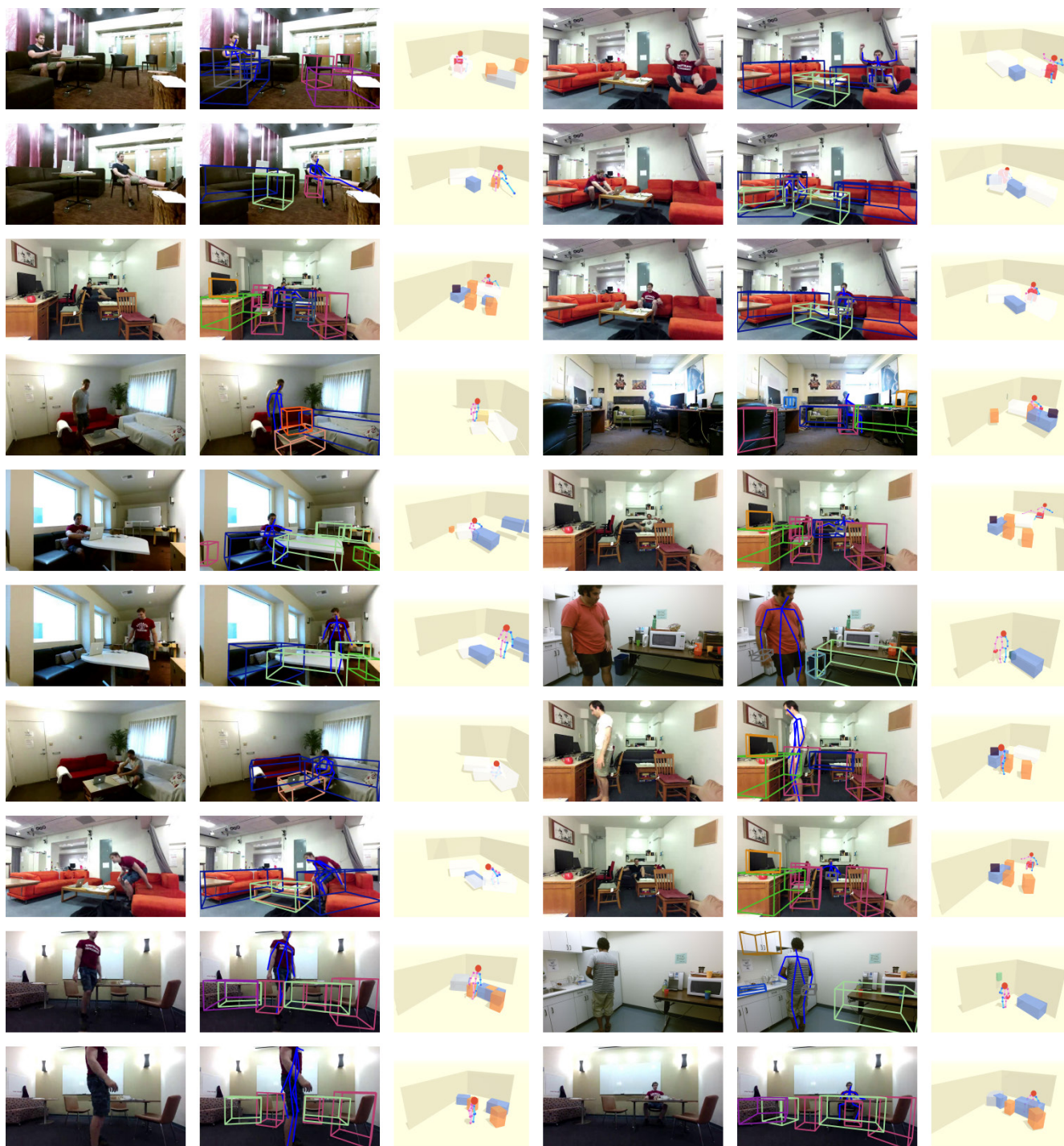


Figure 34. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.



Figure 35. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

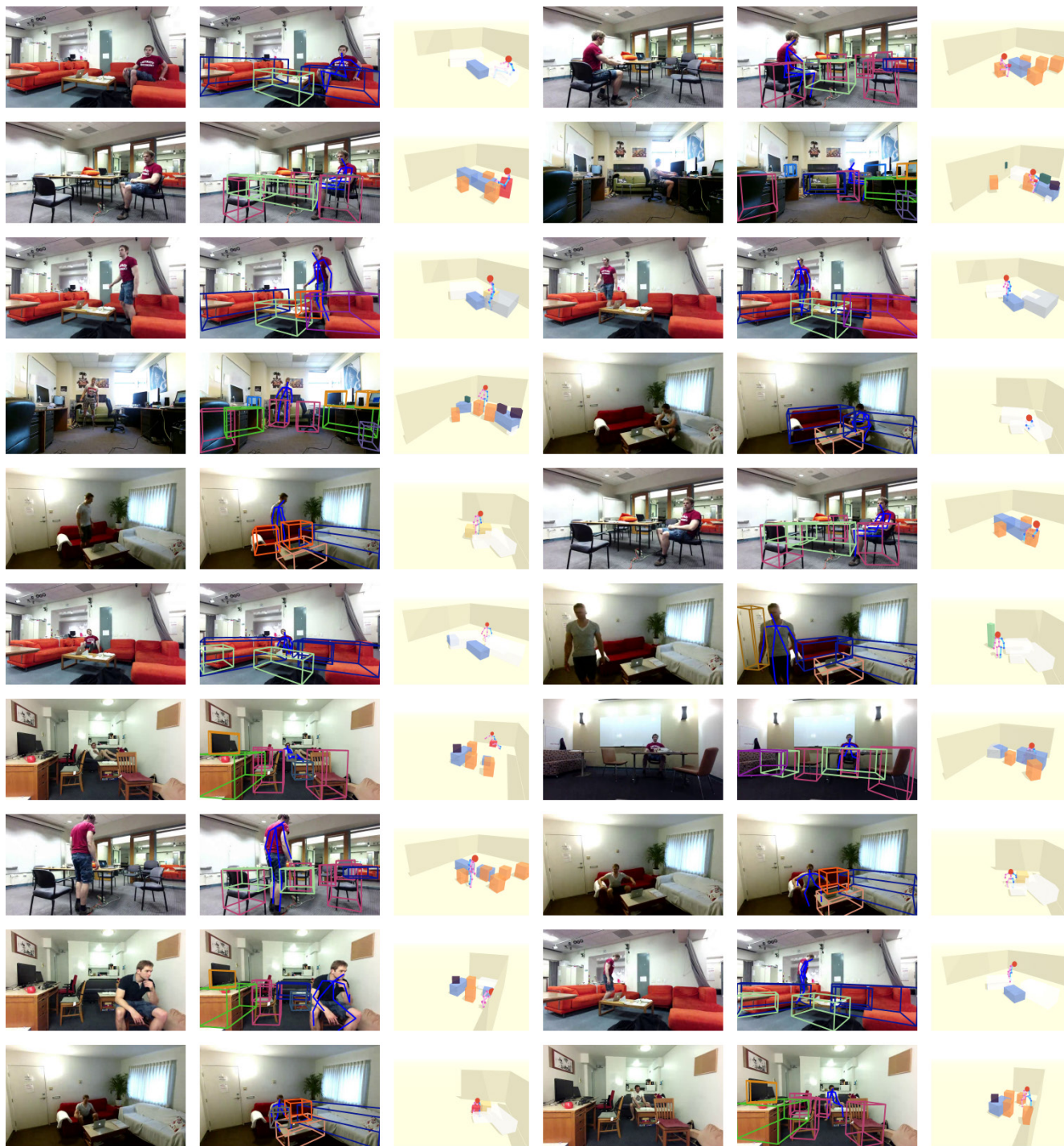


Figure 36. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

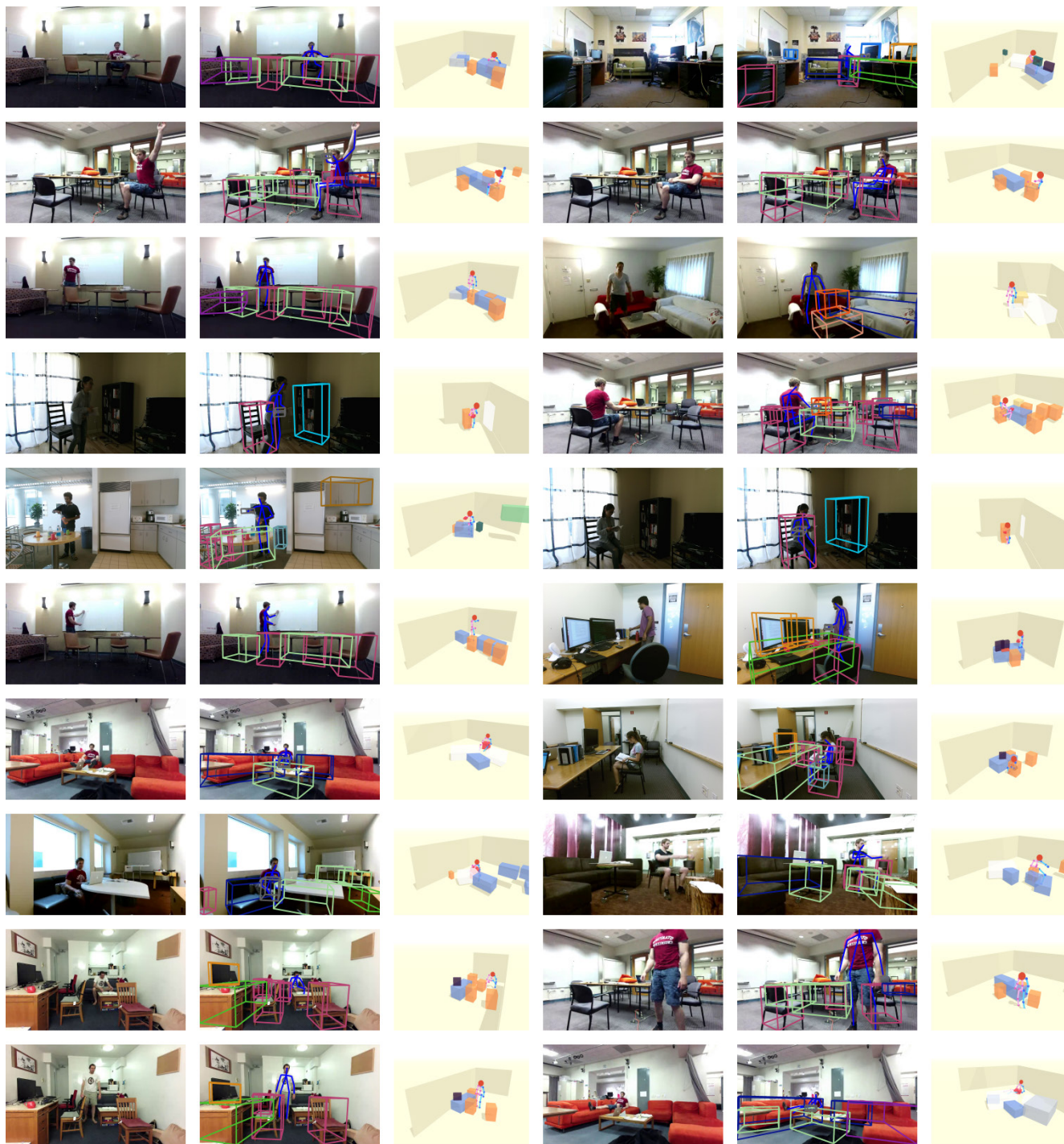


Figure 37. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

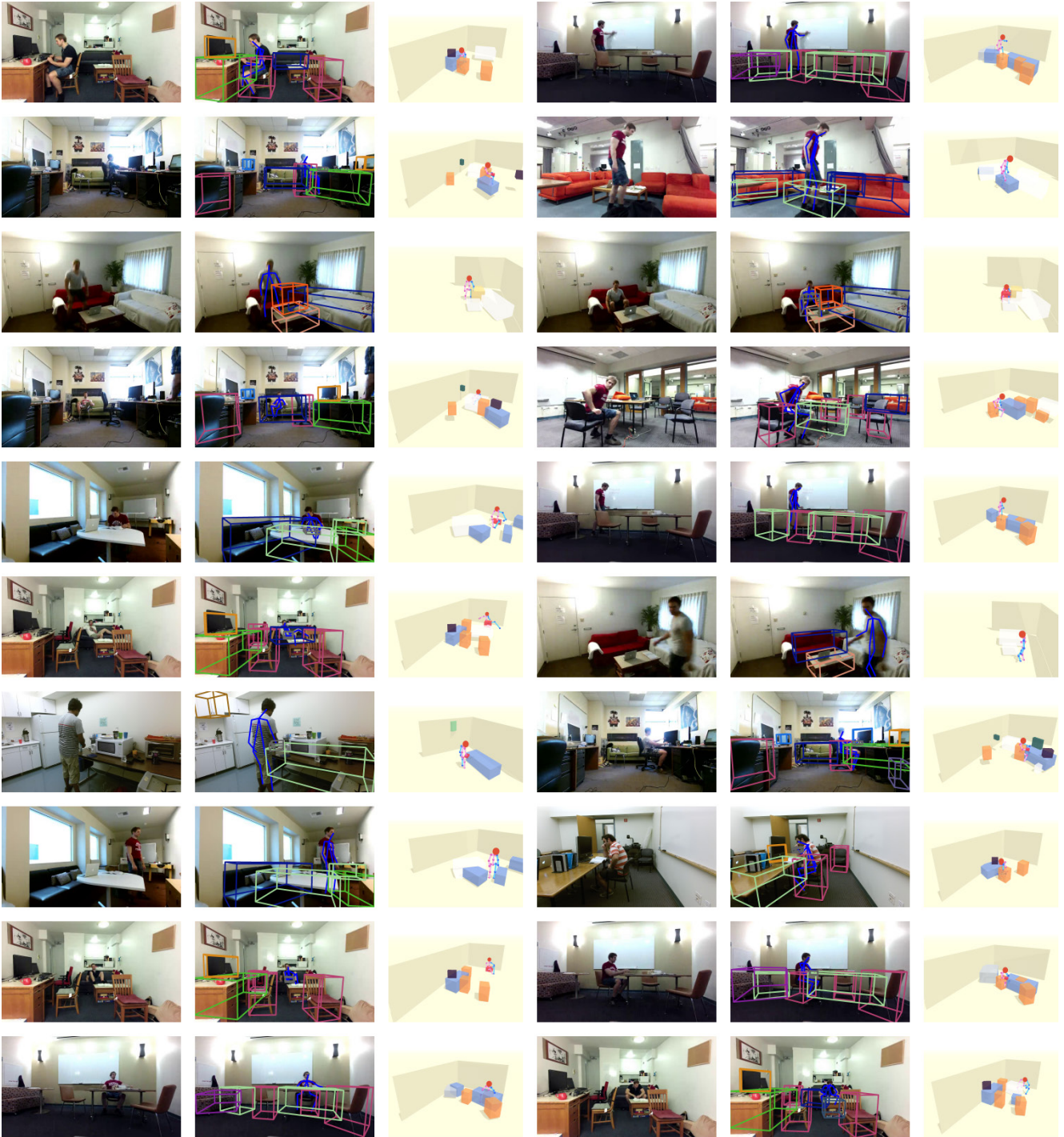


Figure 38. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

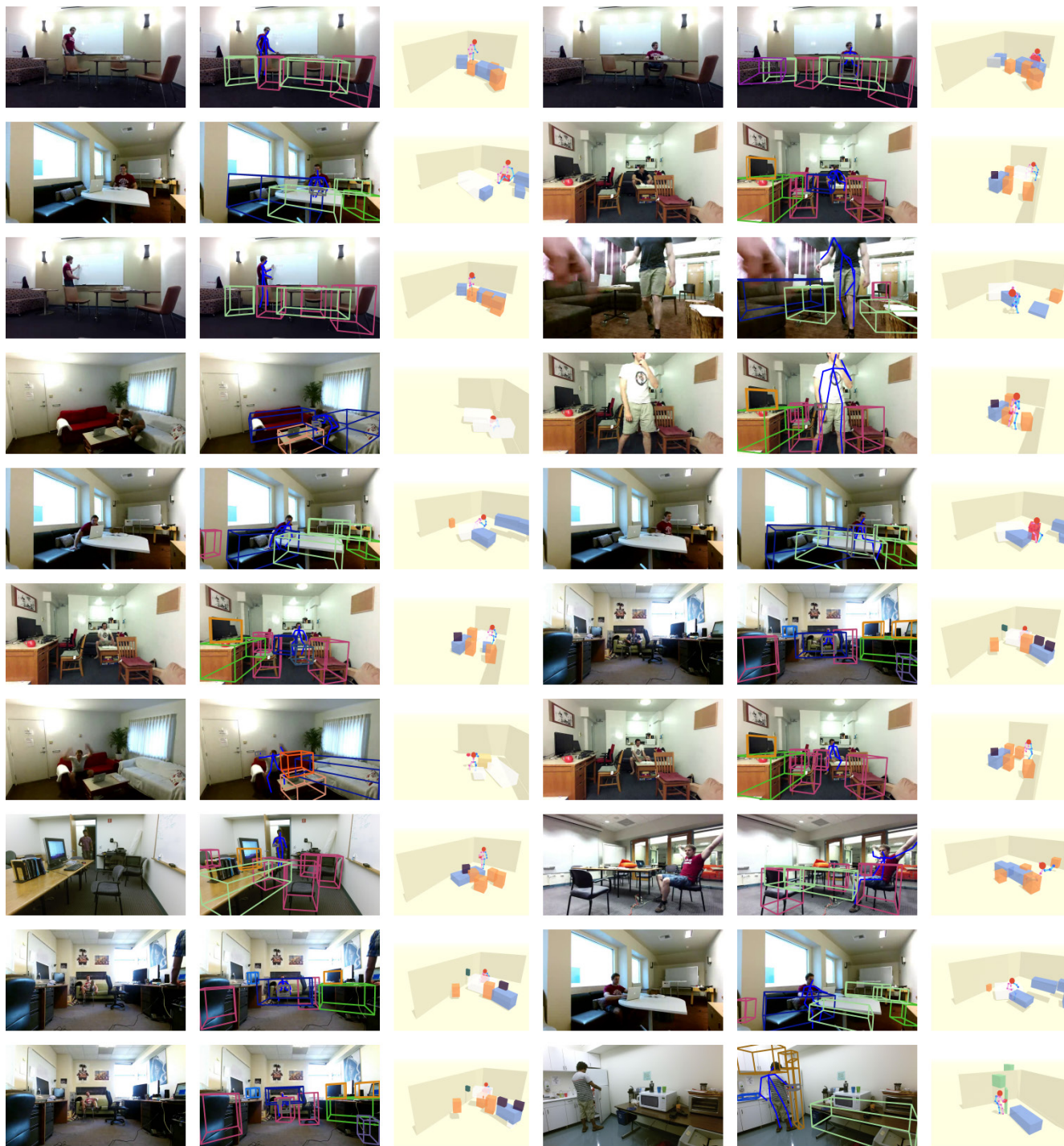


Figure 39. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

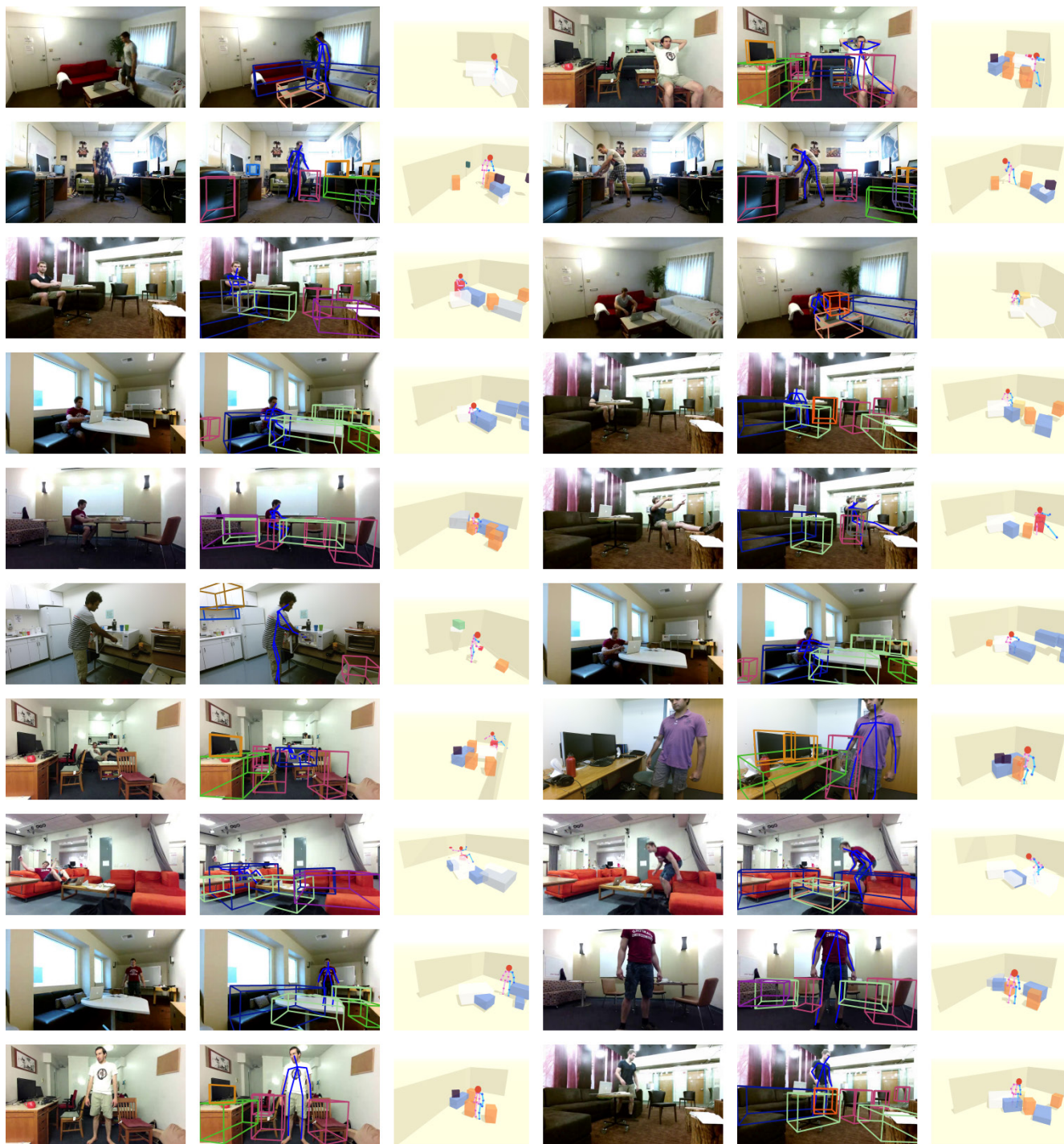


Figure 40. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

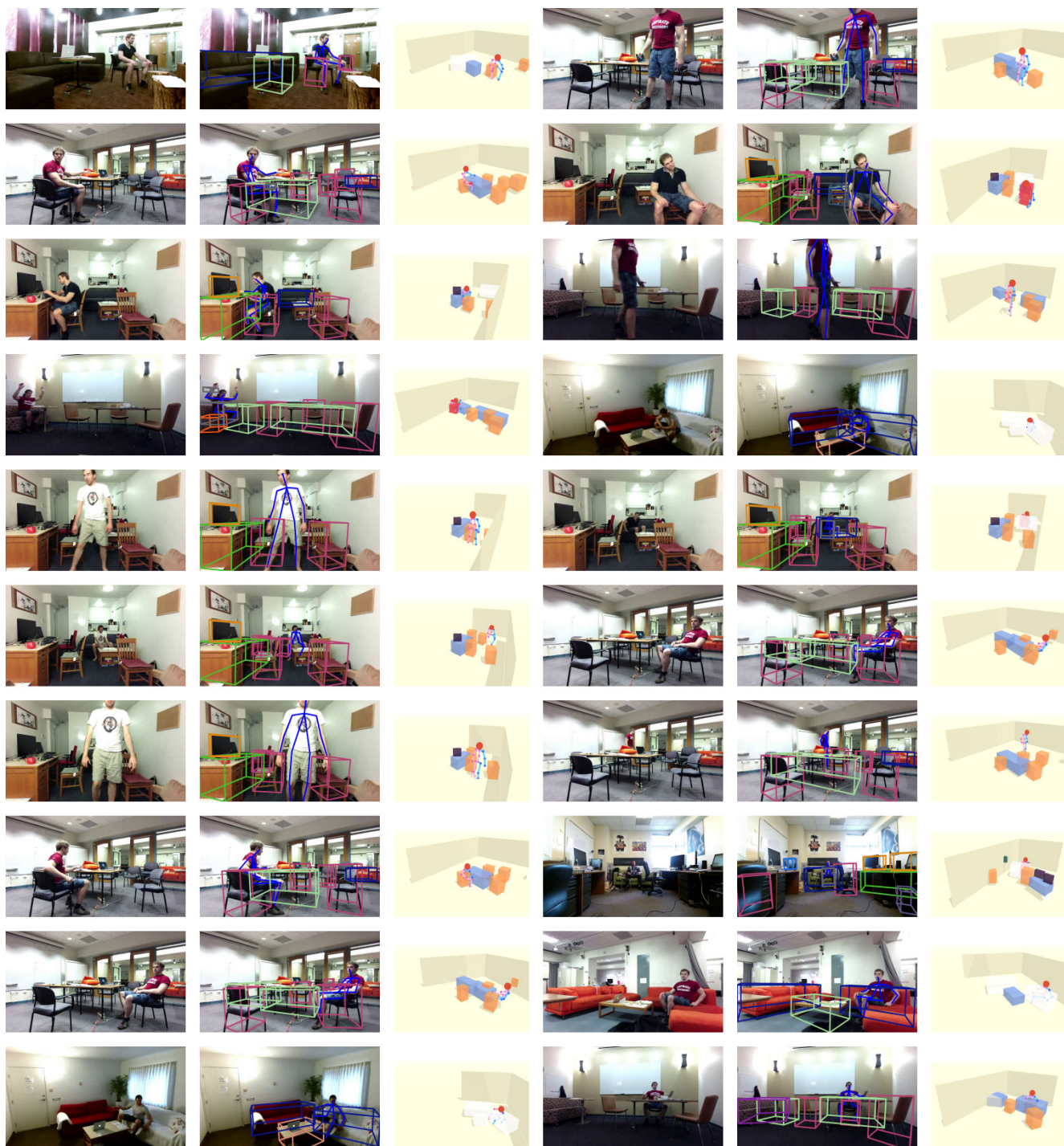


Figure 41. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

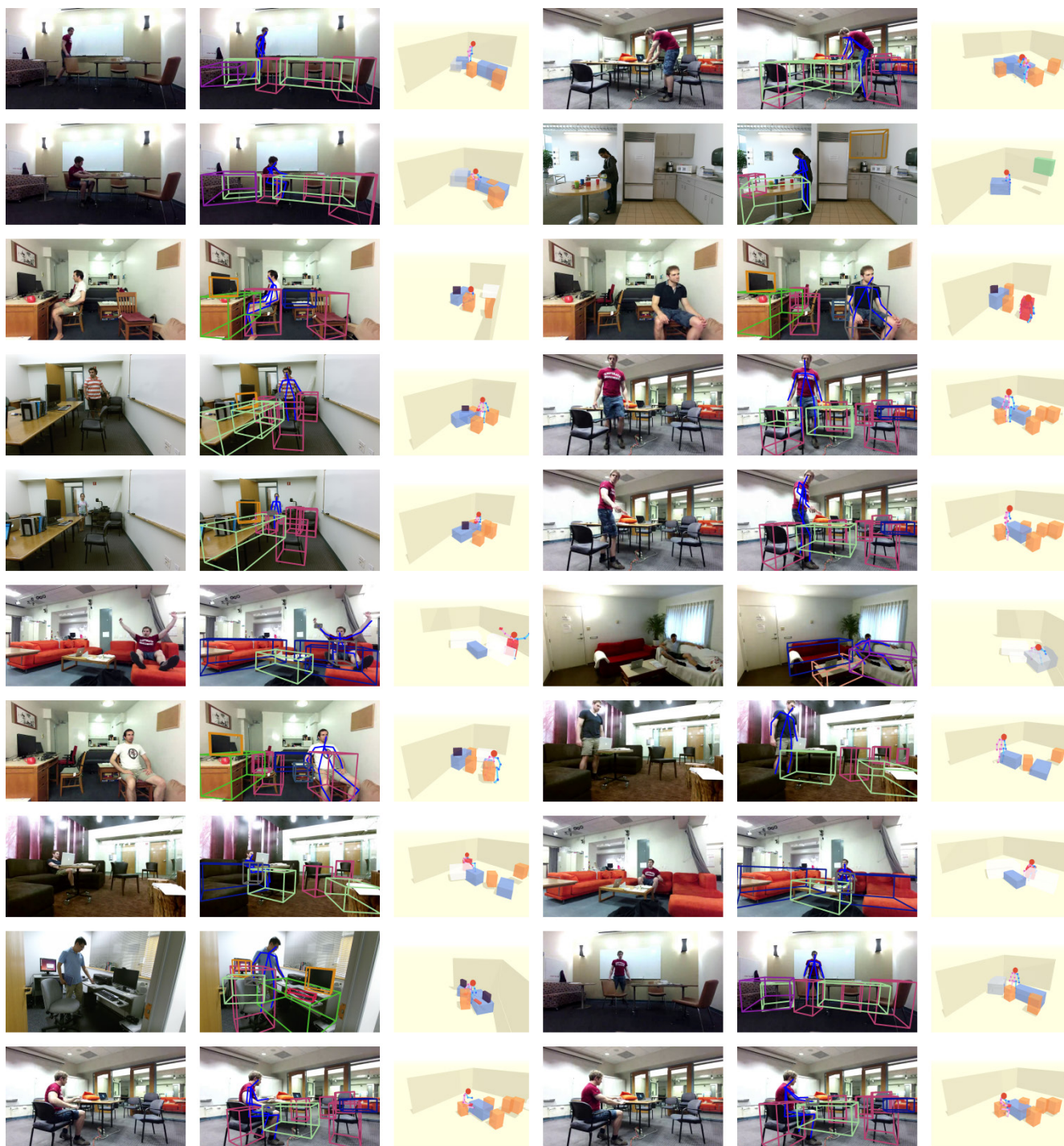


Figure 42. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

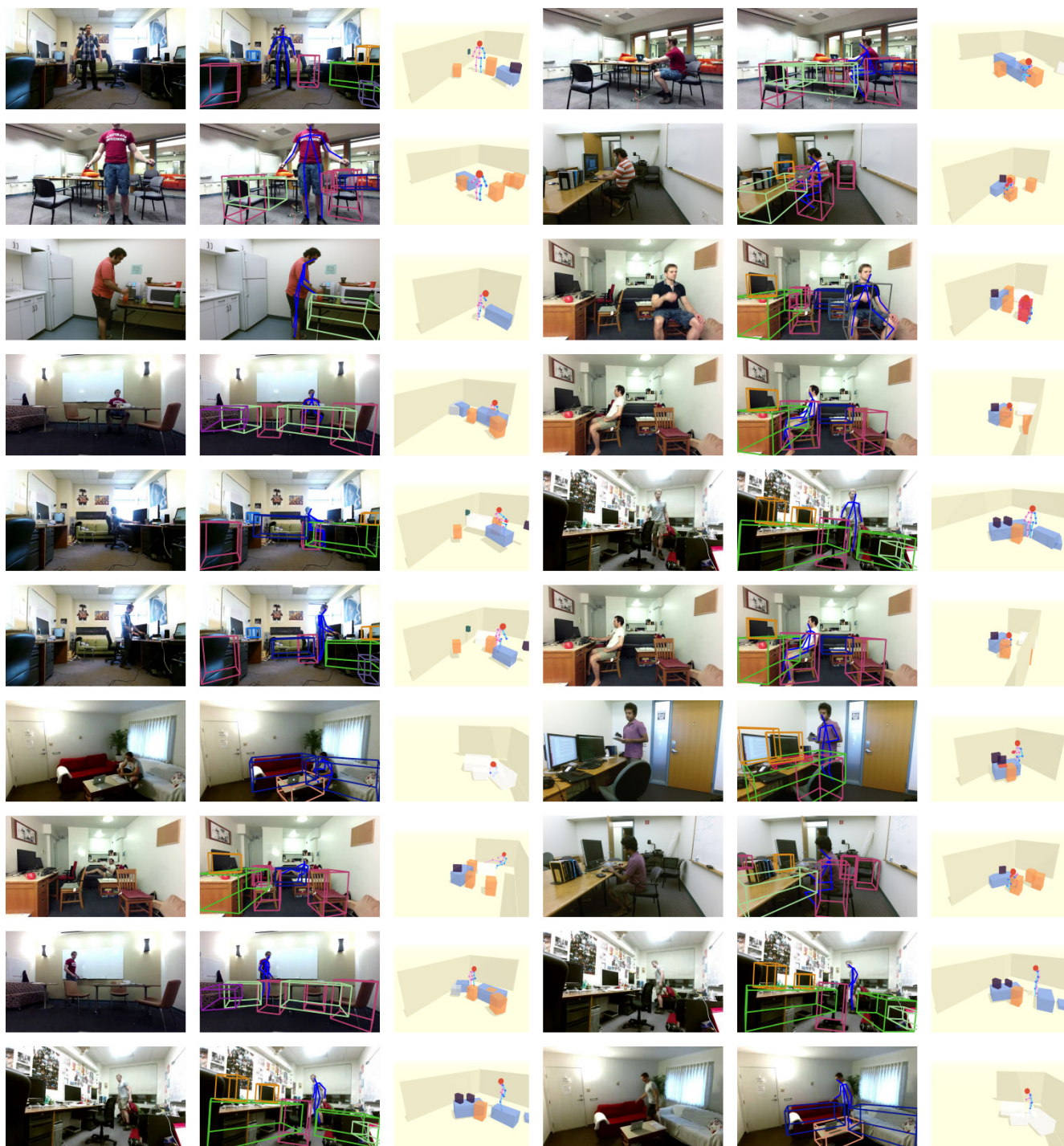


Figure 43. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

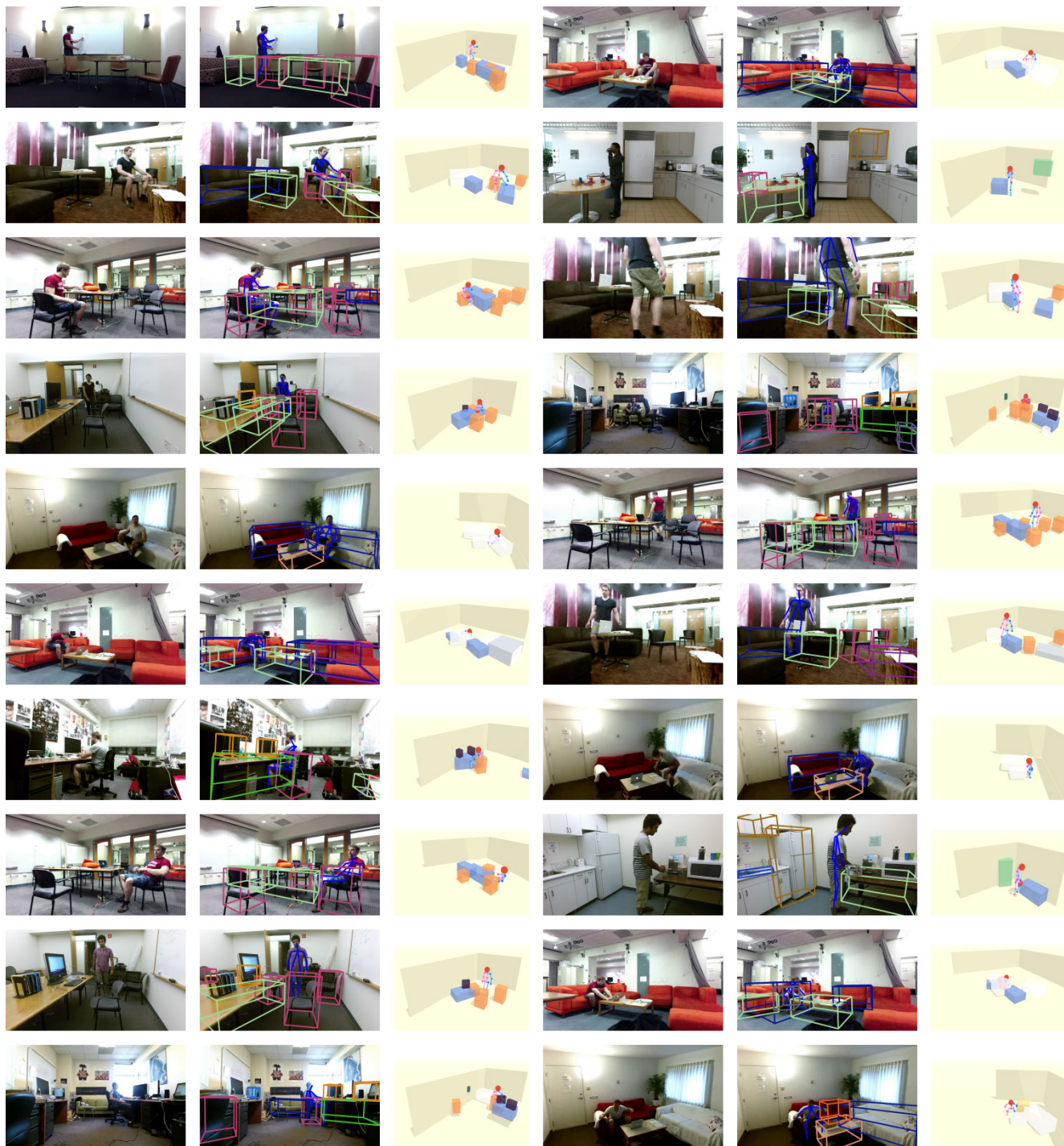


Figure 44. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

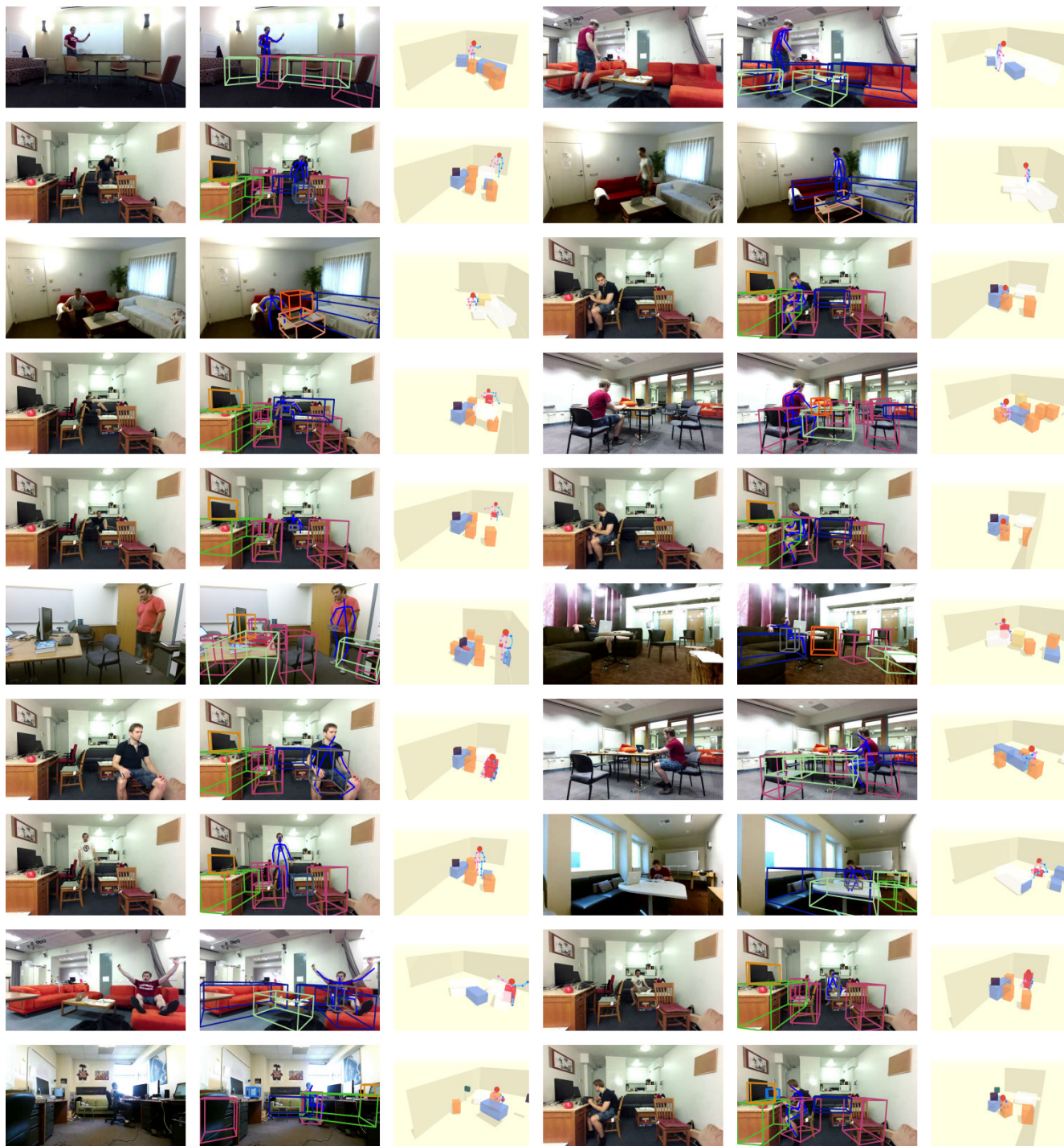


Figure 45. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

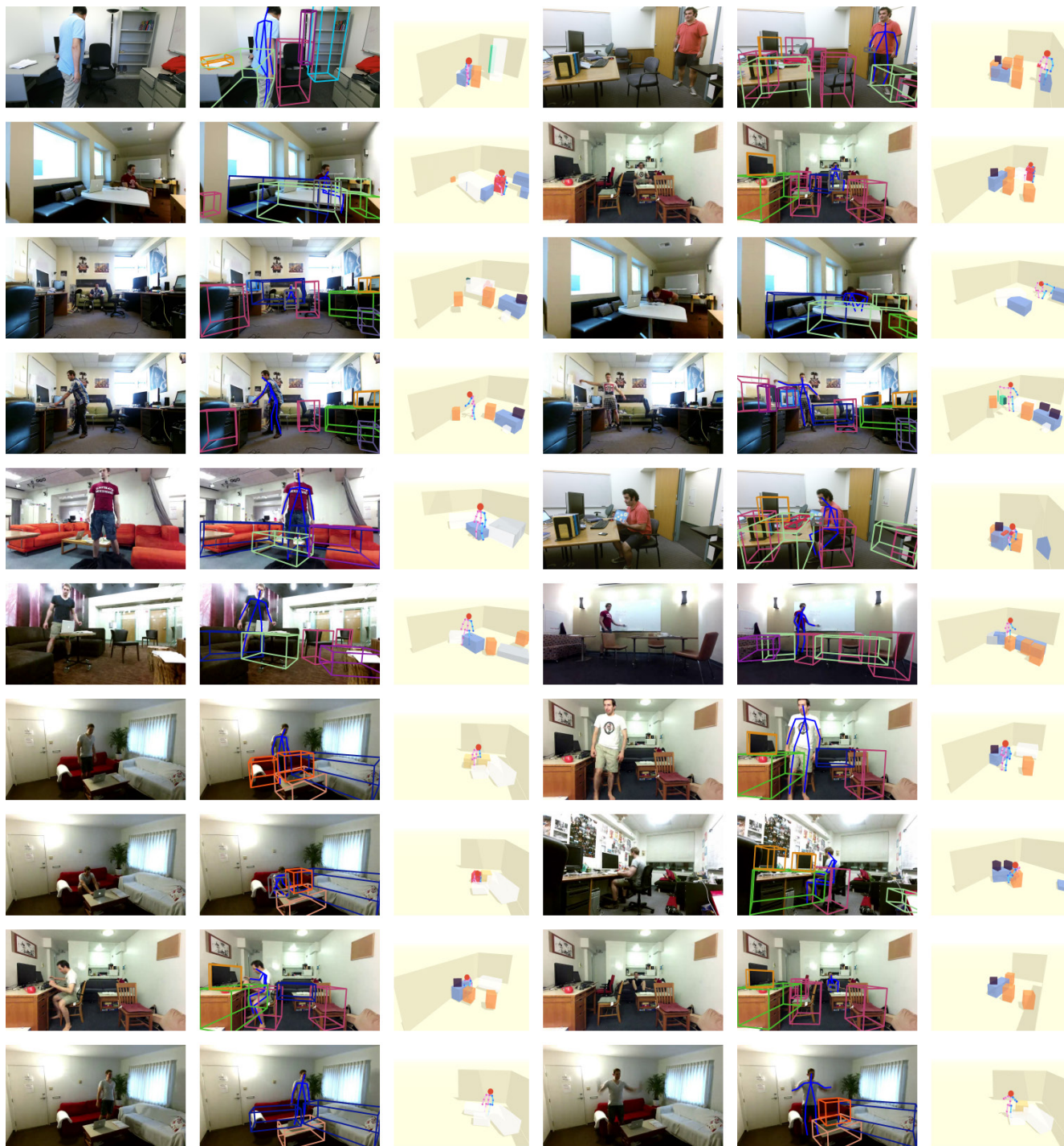


Figure 46. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

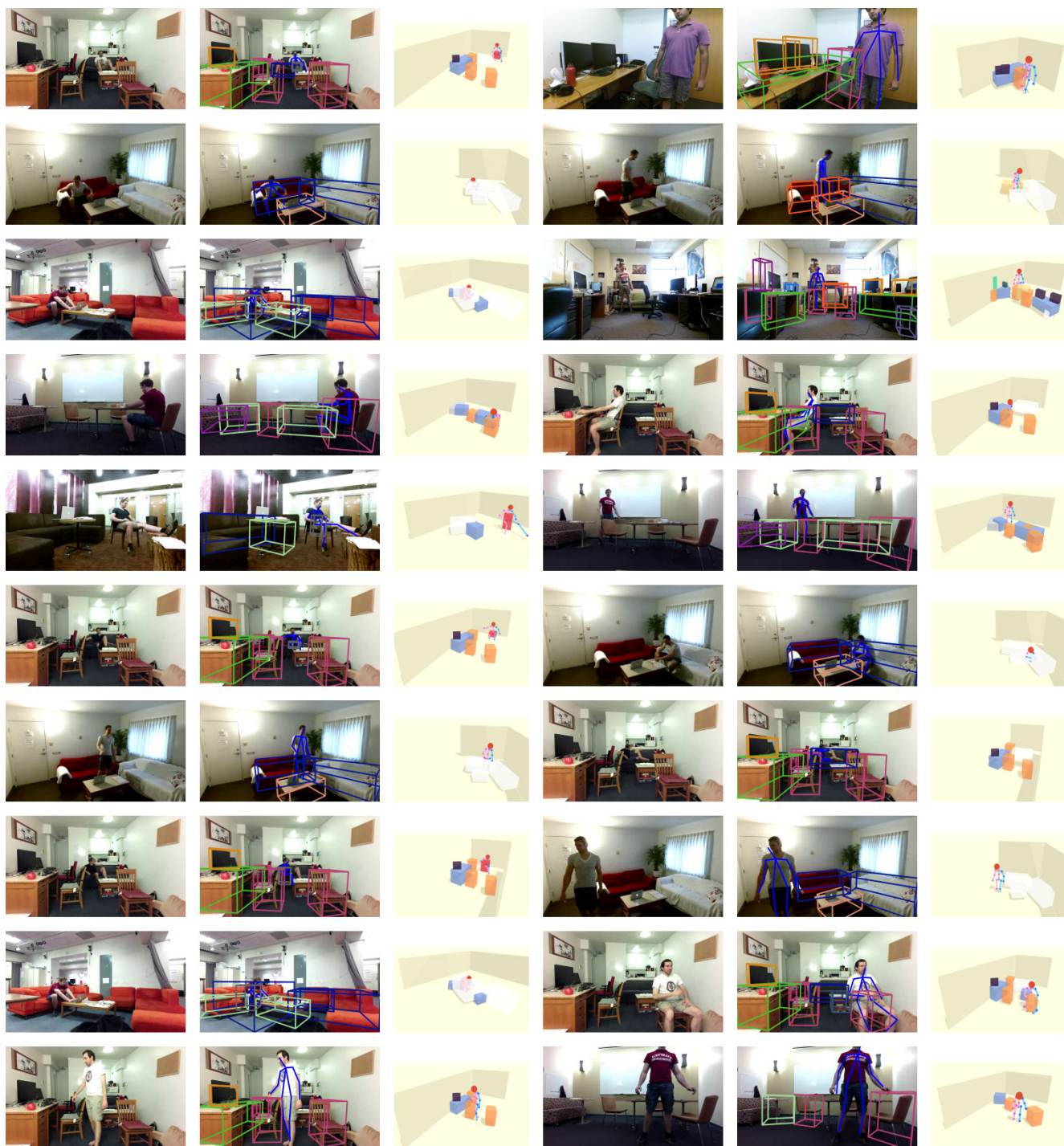


Figure 47. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

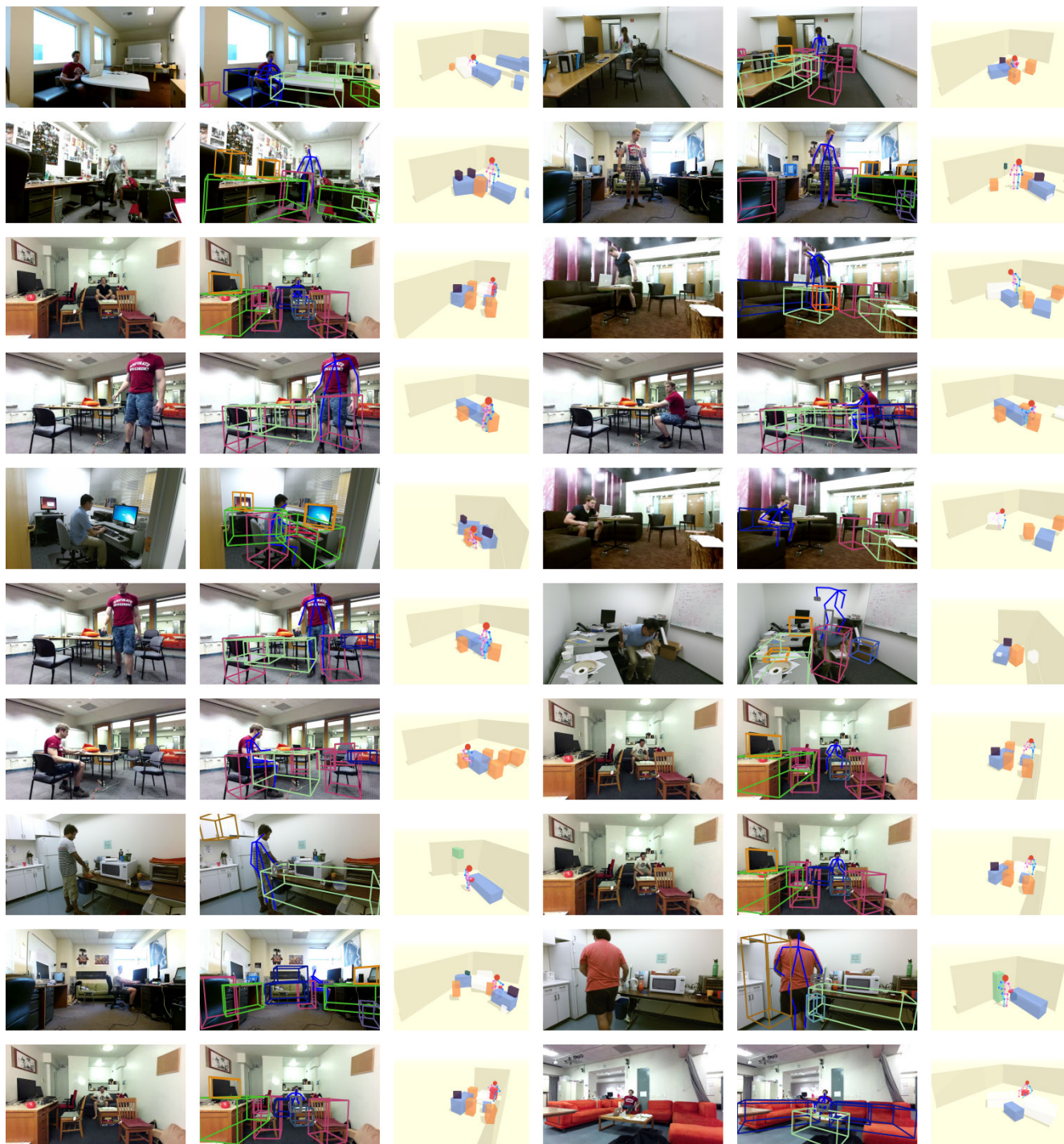


Figure 48. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

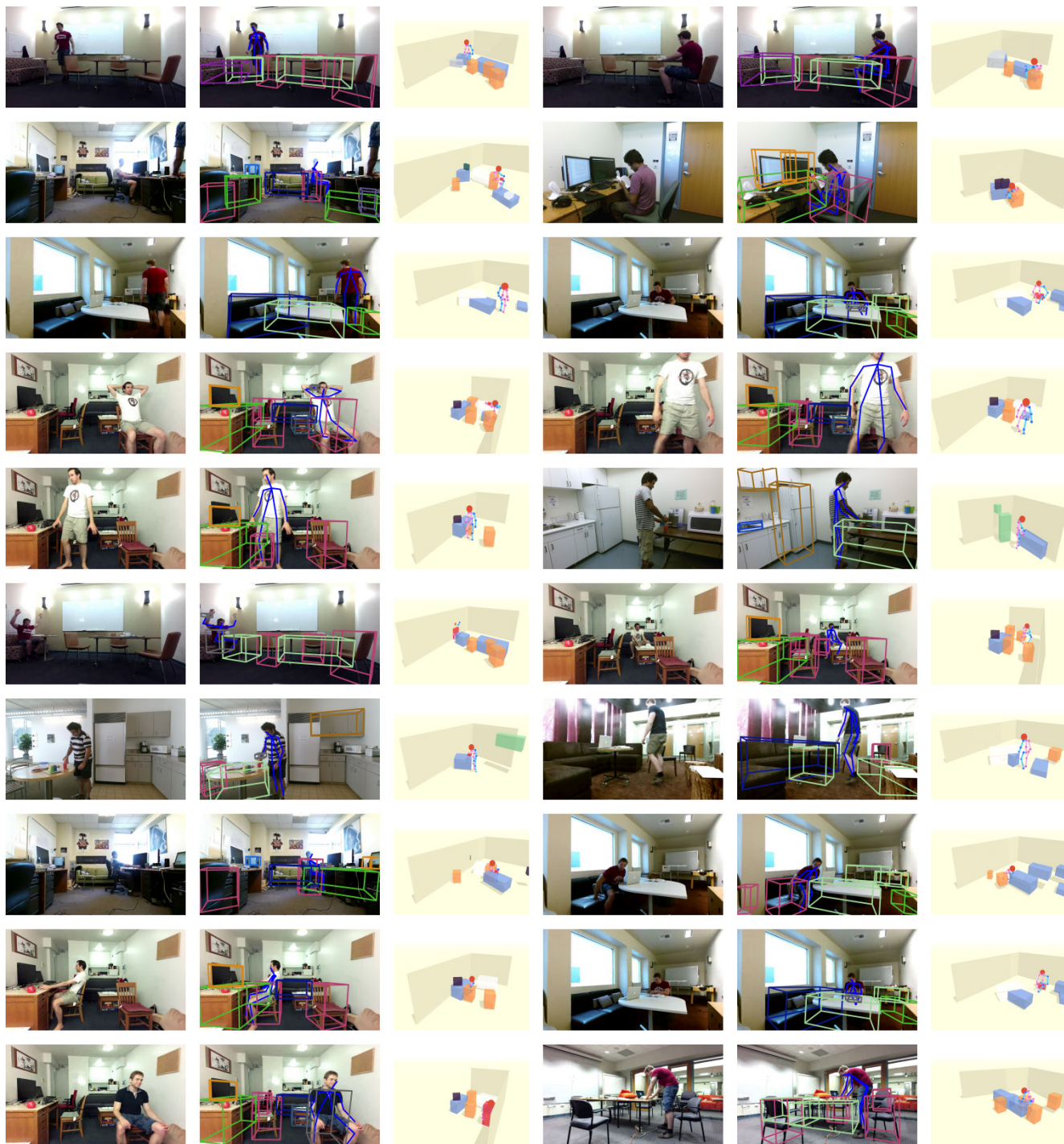


Figure 49. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

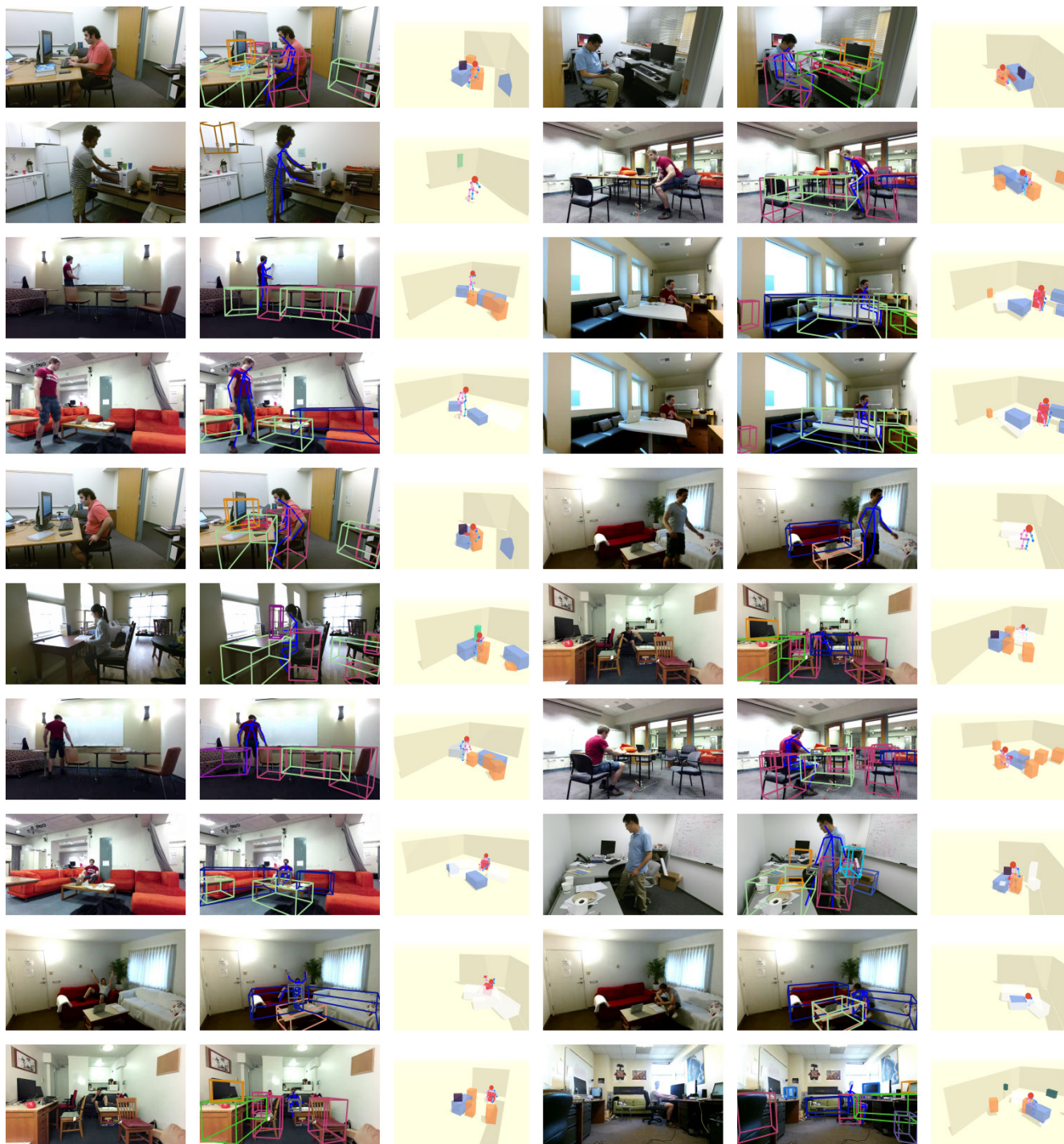


Figure 50. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

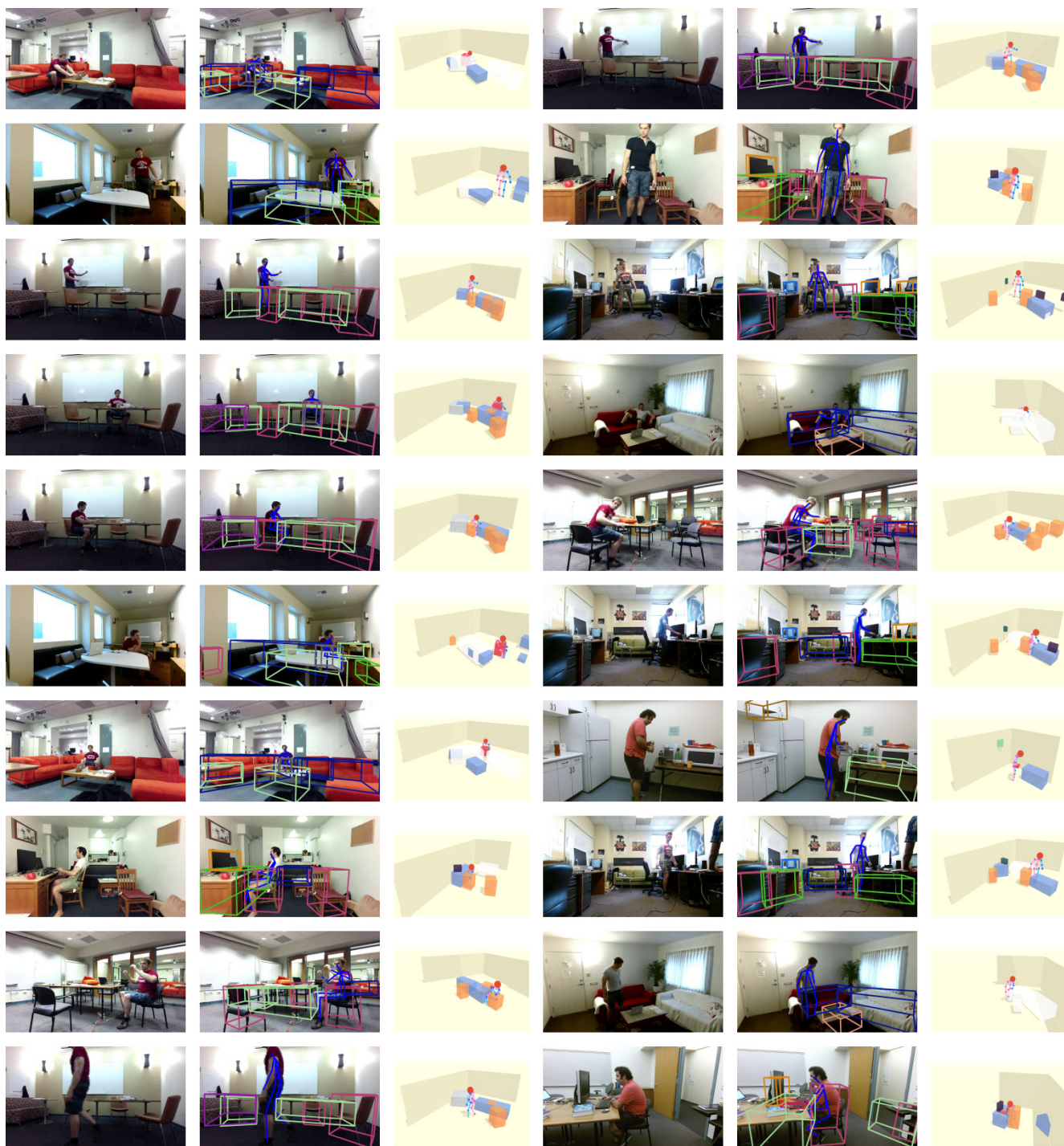


Figure 51. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

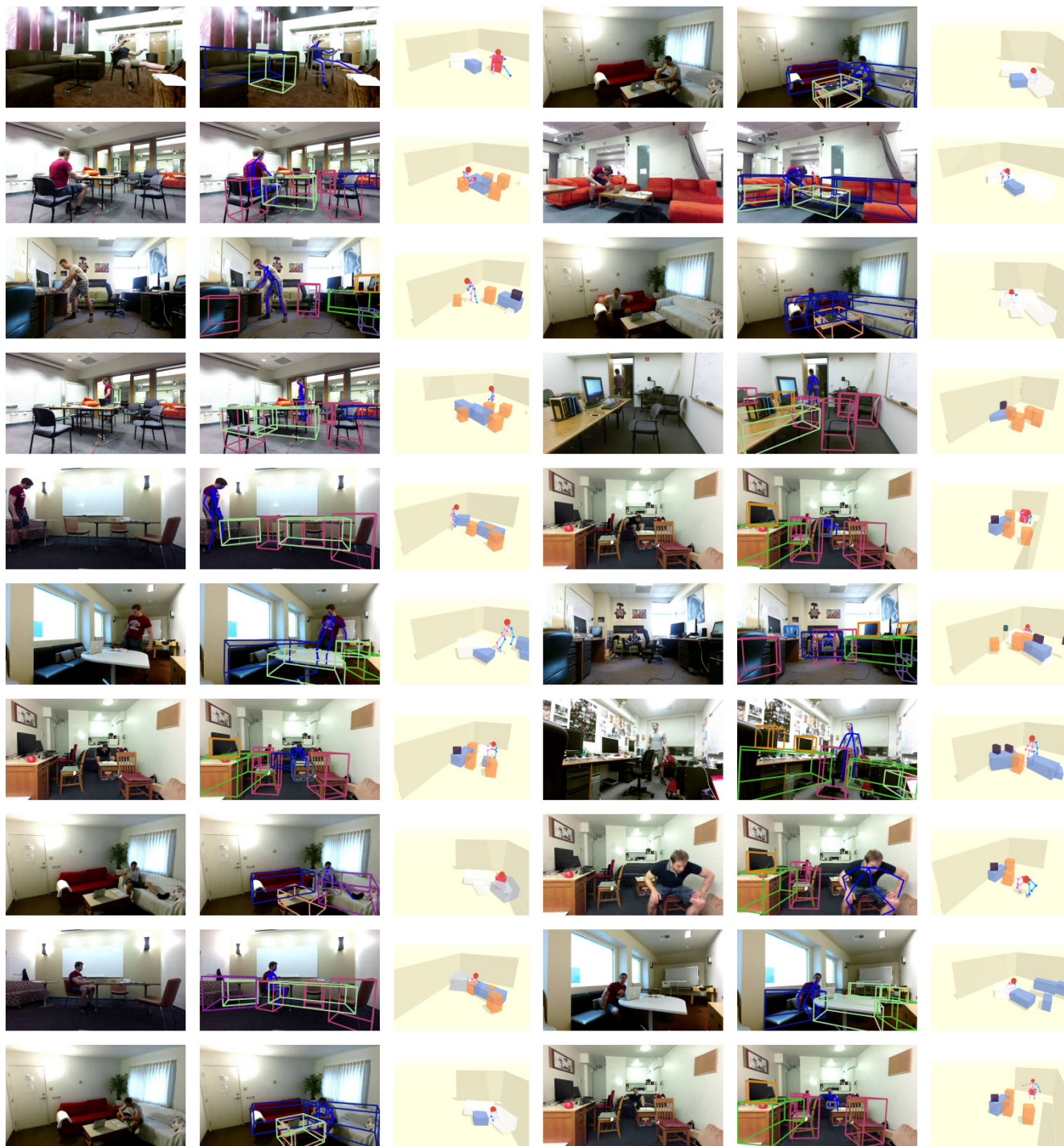


Figure 52. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

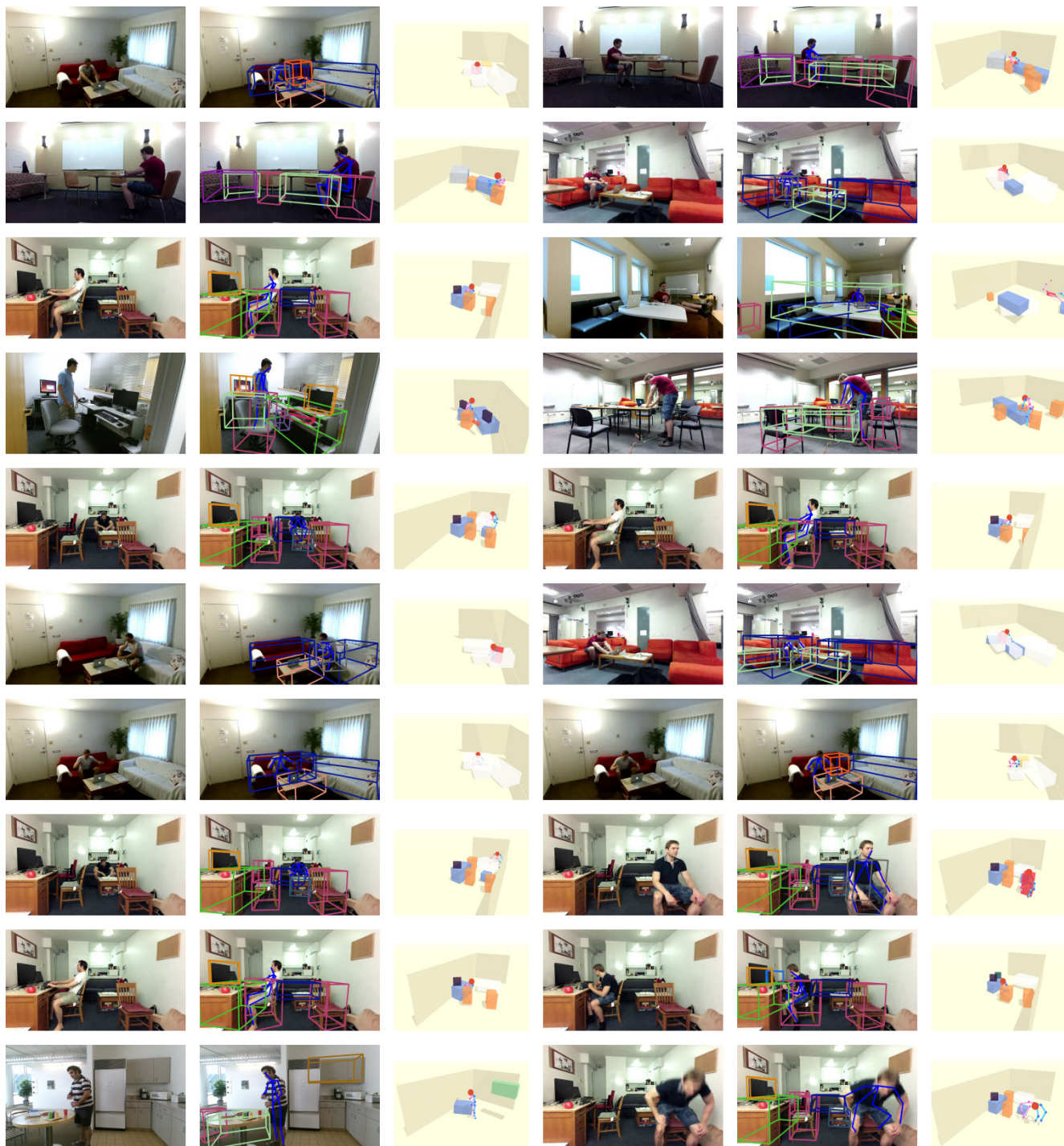


Figure 53. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

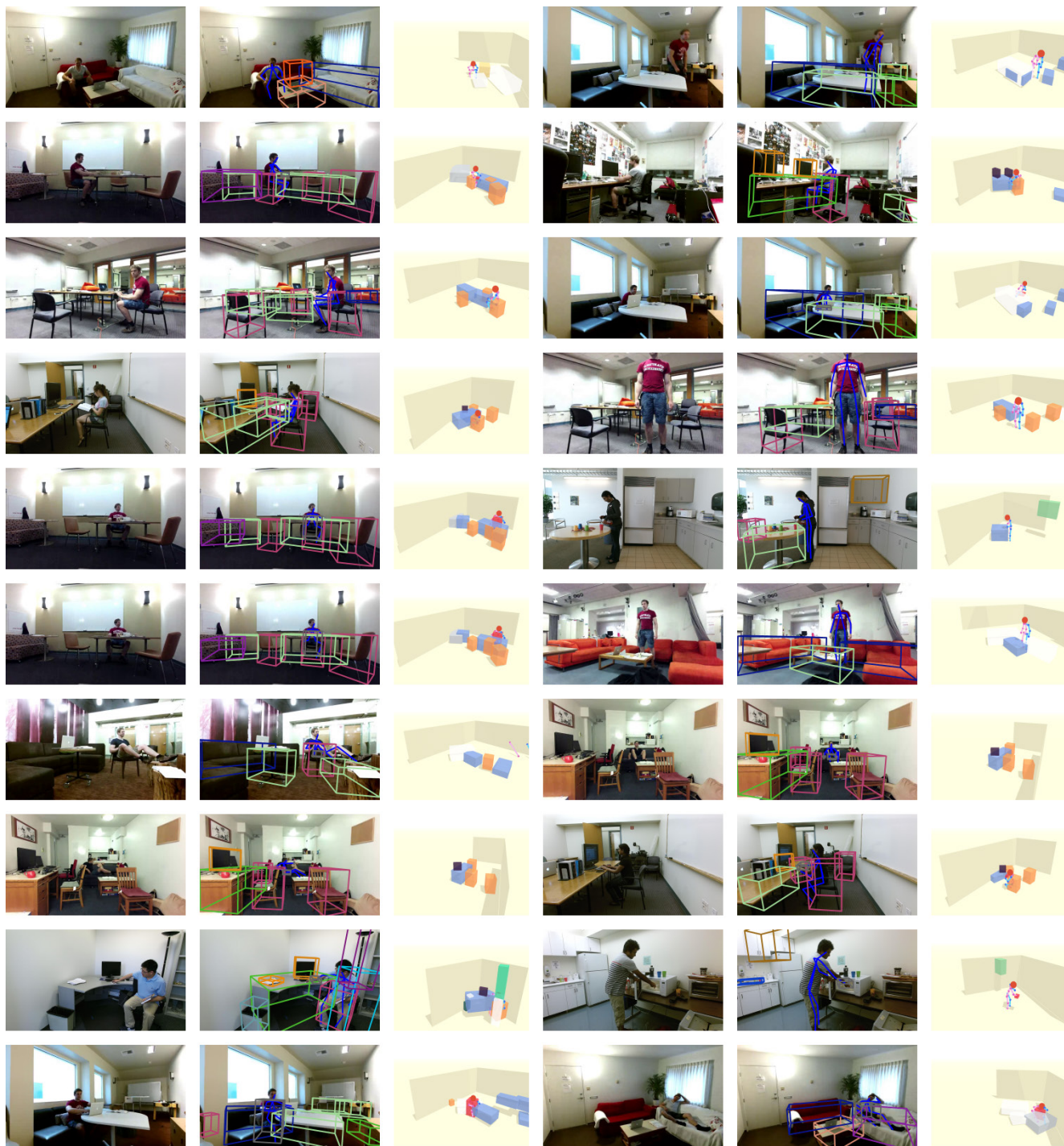


Figure 54. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

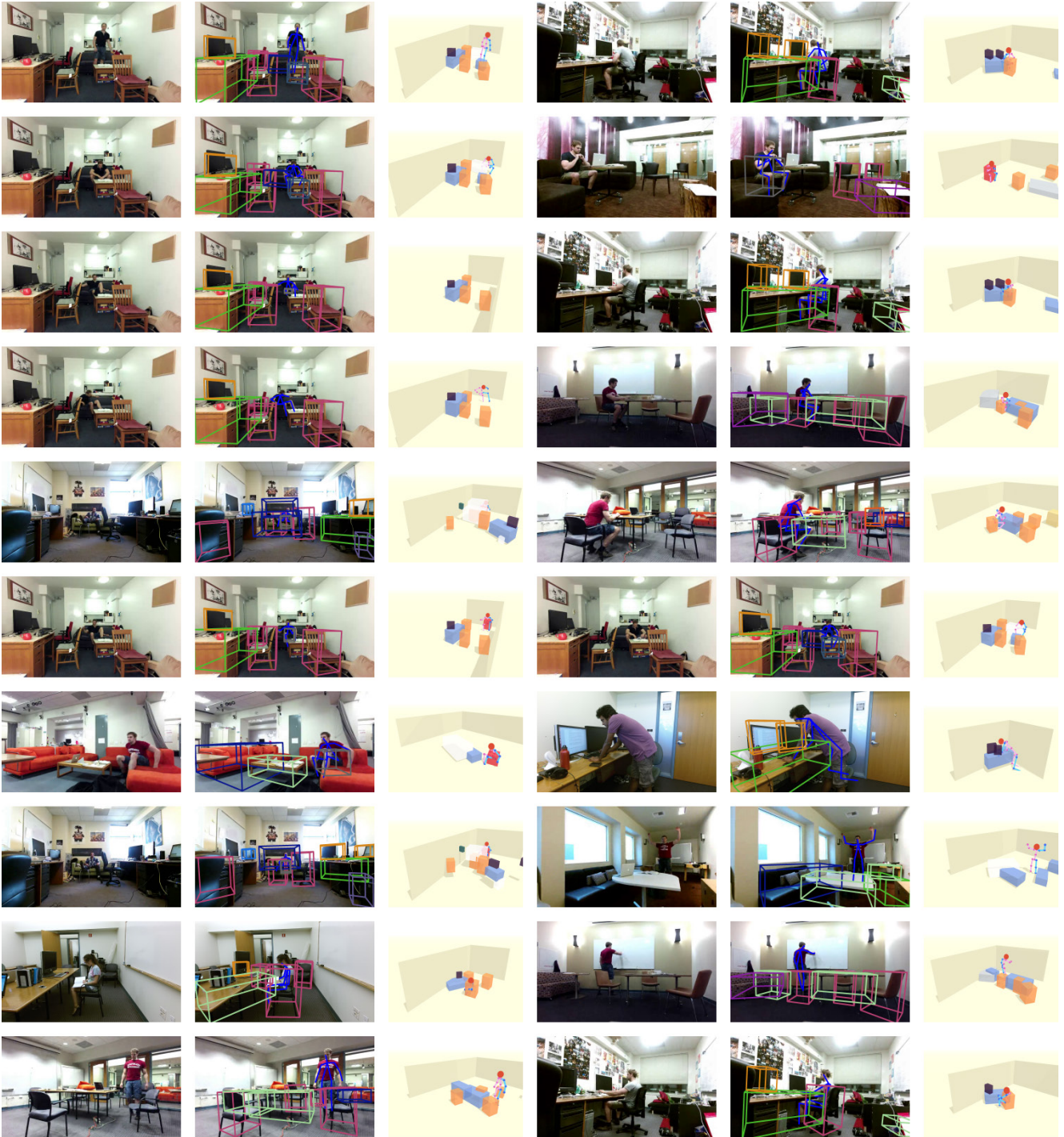


Figure 55. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

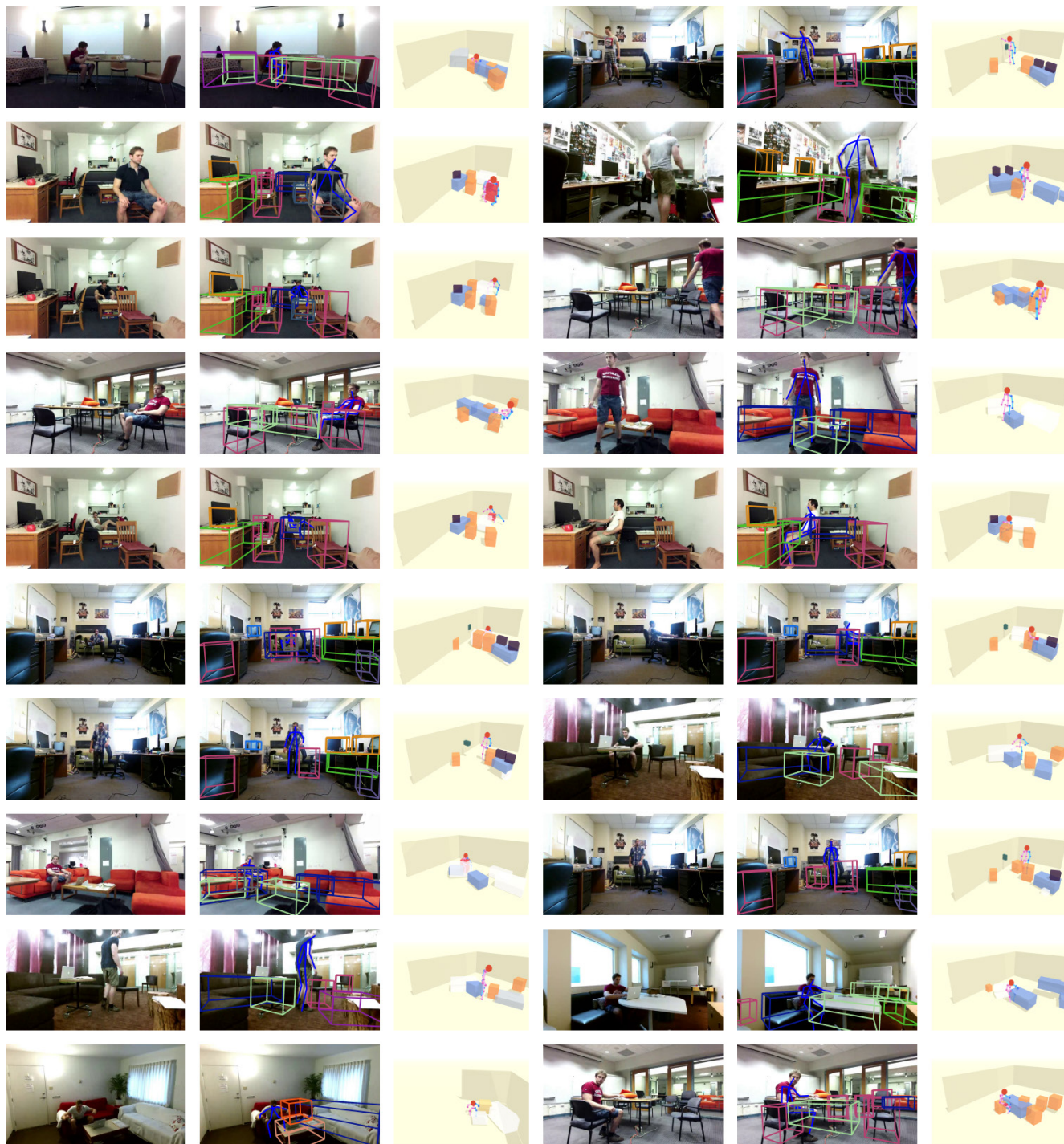


Figure 56. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

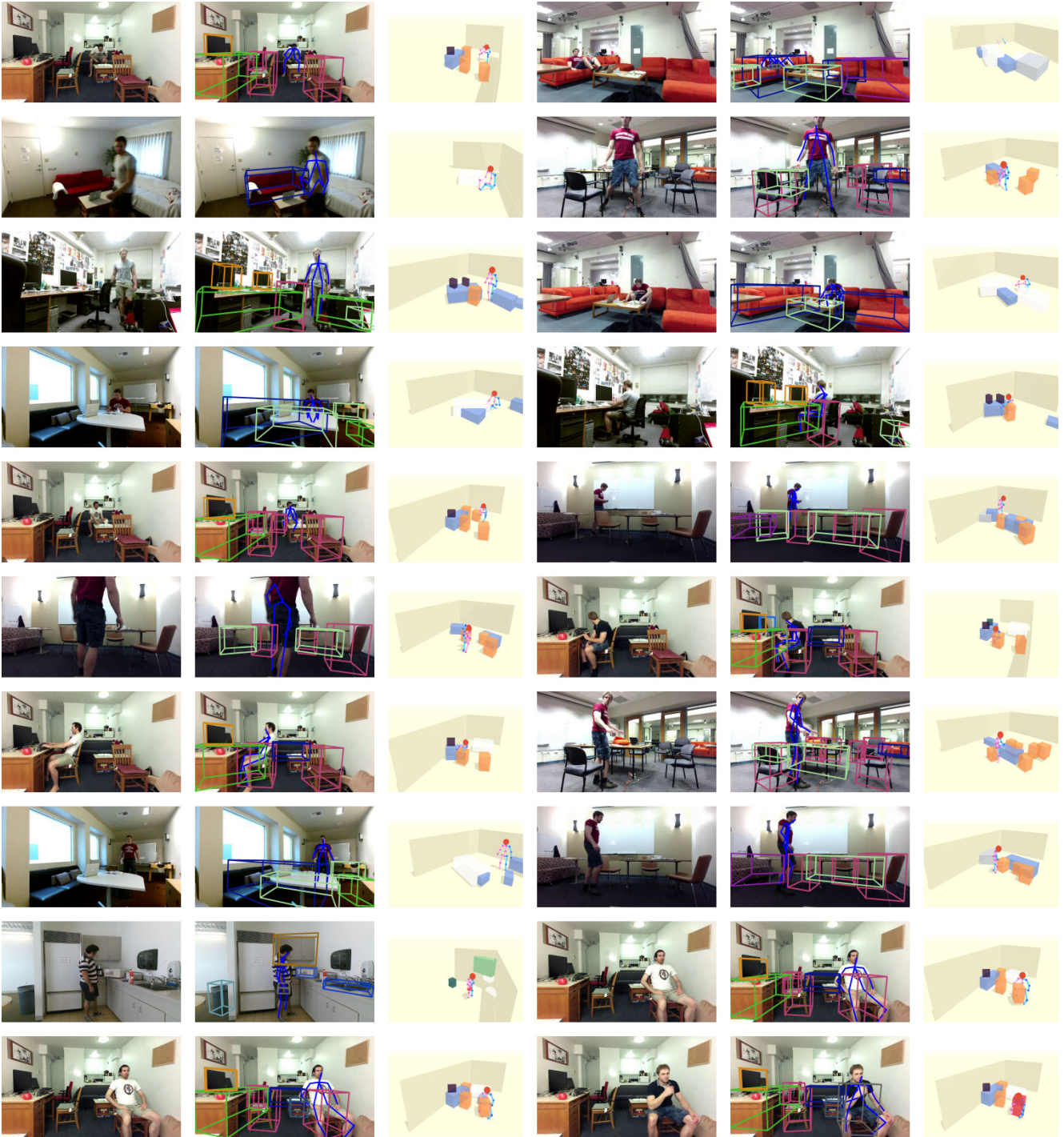


Figure 57. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

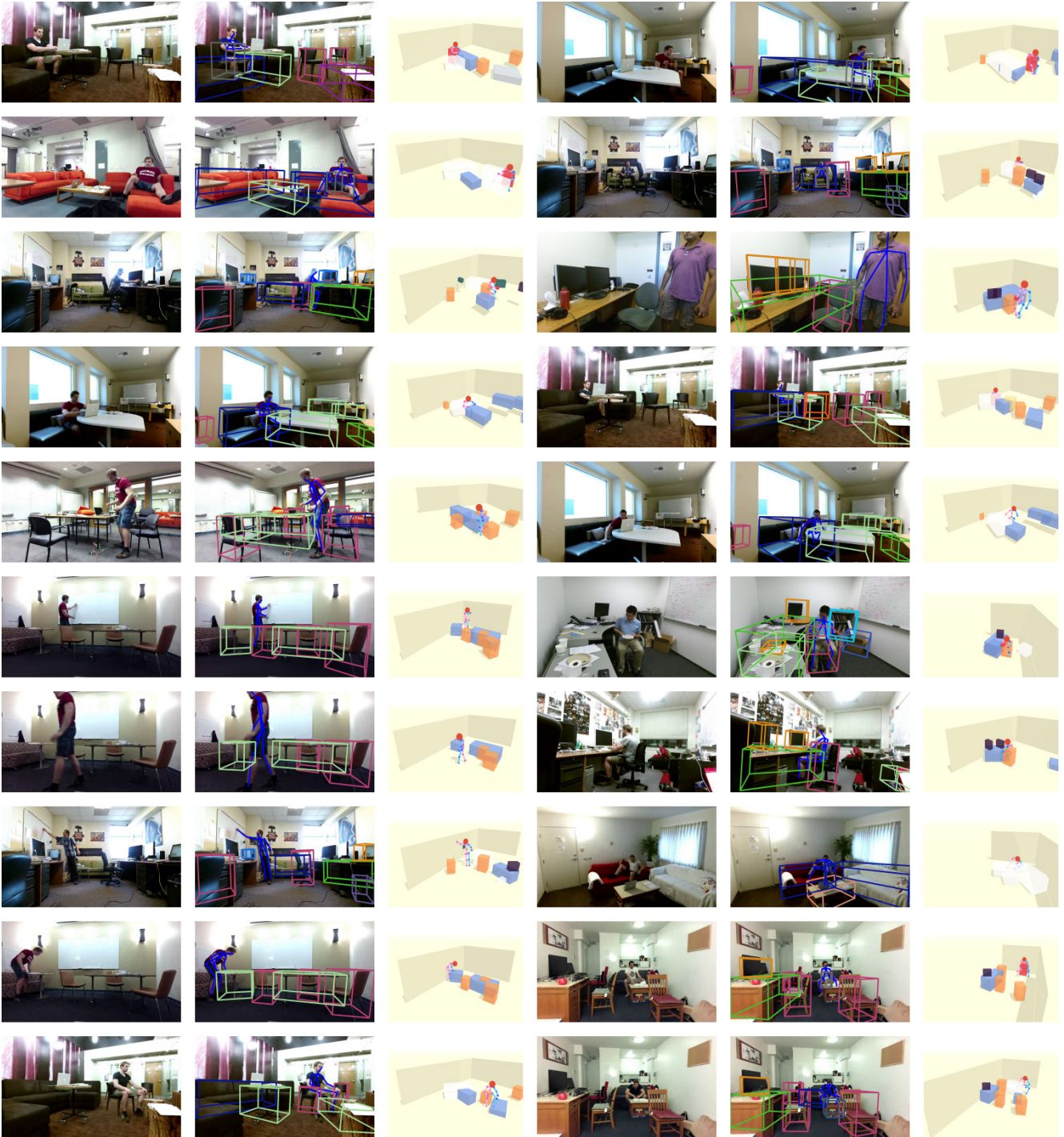


Figure 58. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

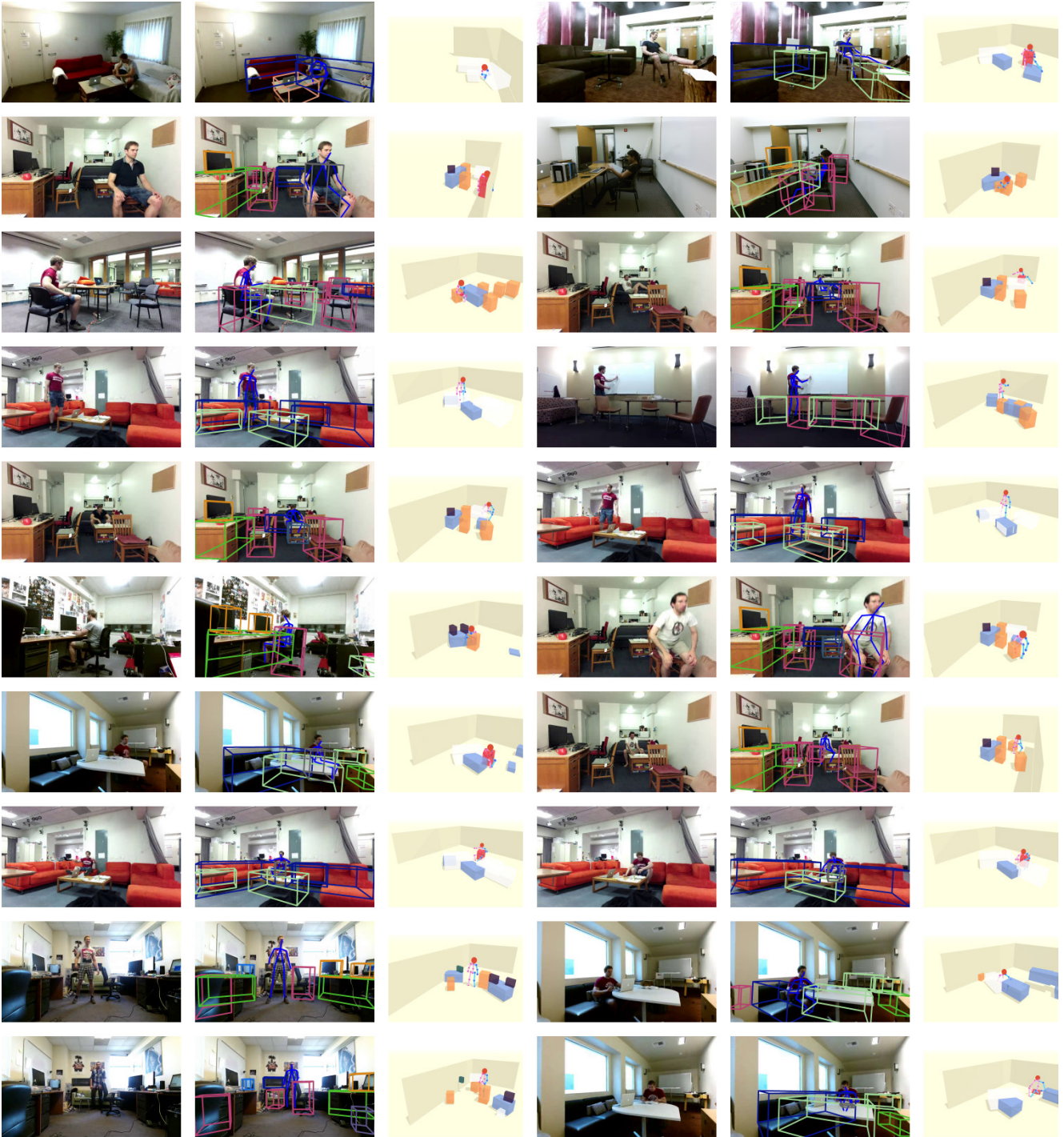


Figure 59. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

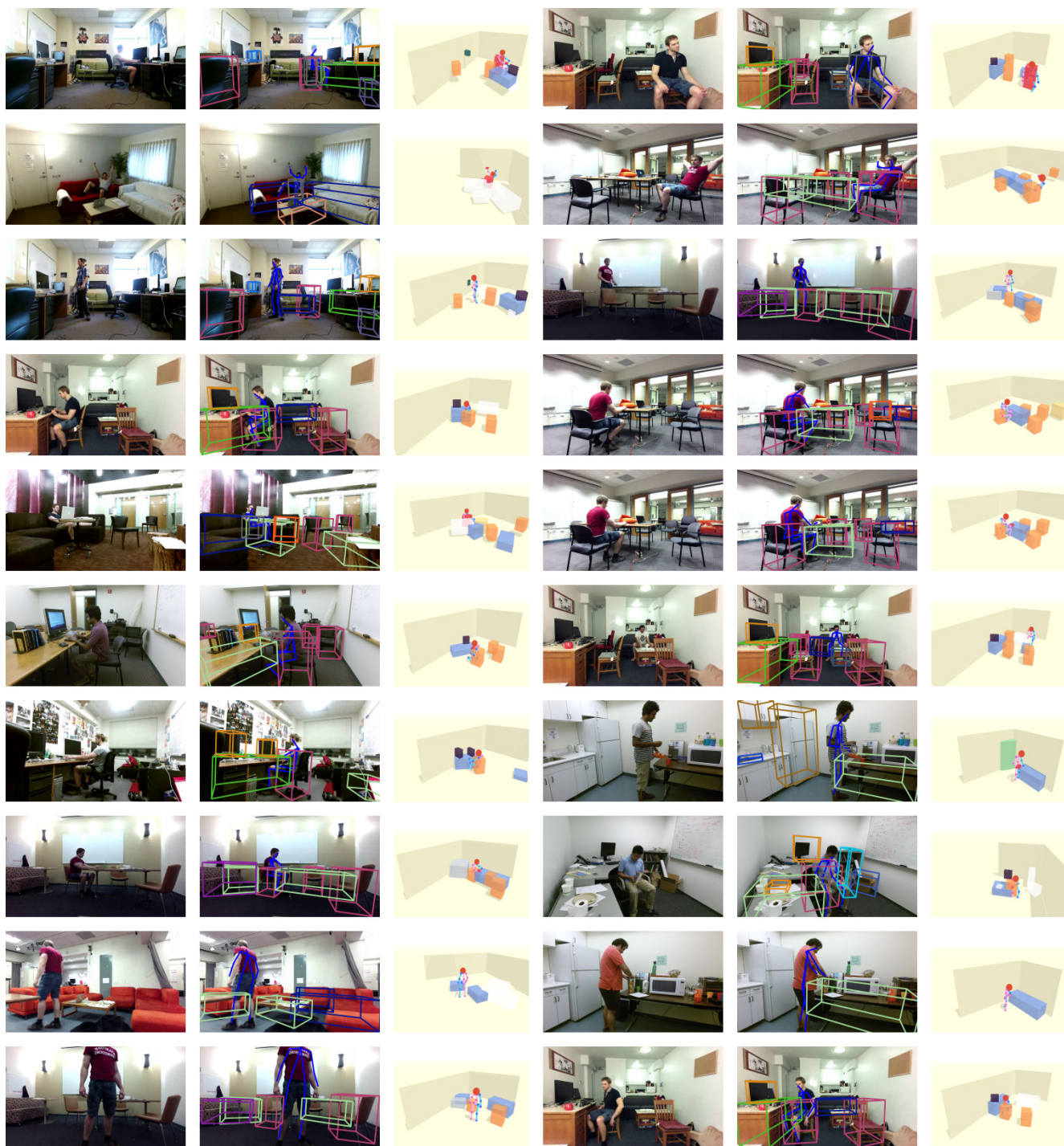


Figure 60. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

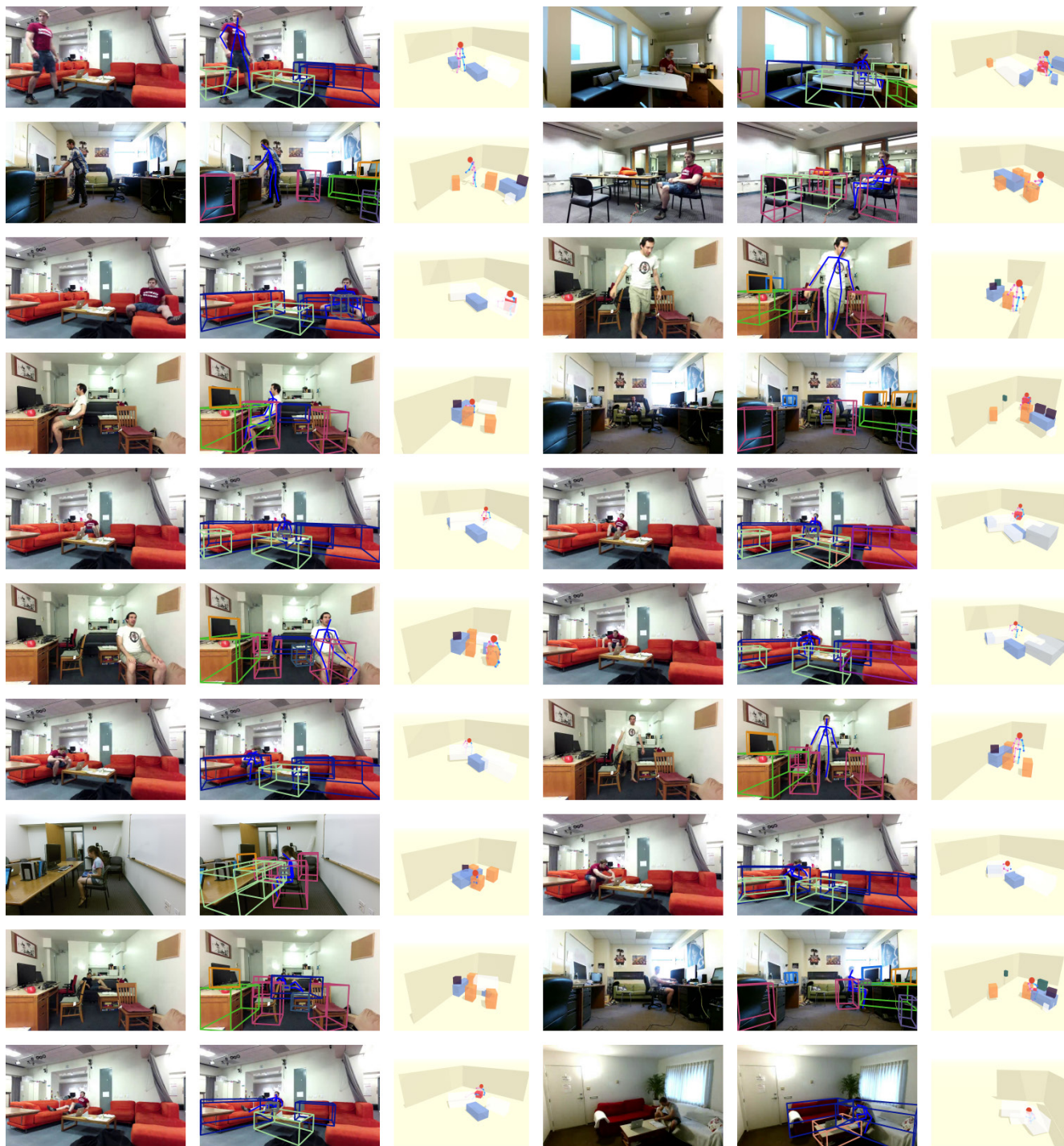


Figure 61. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

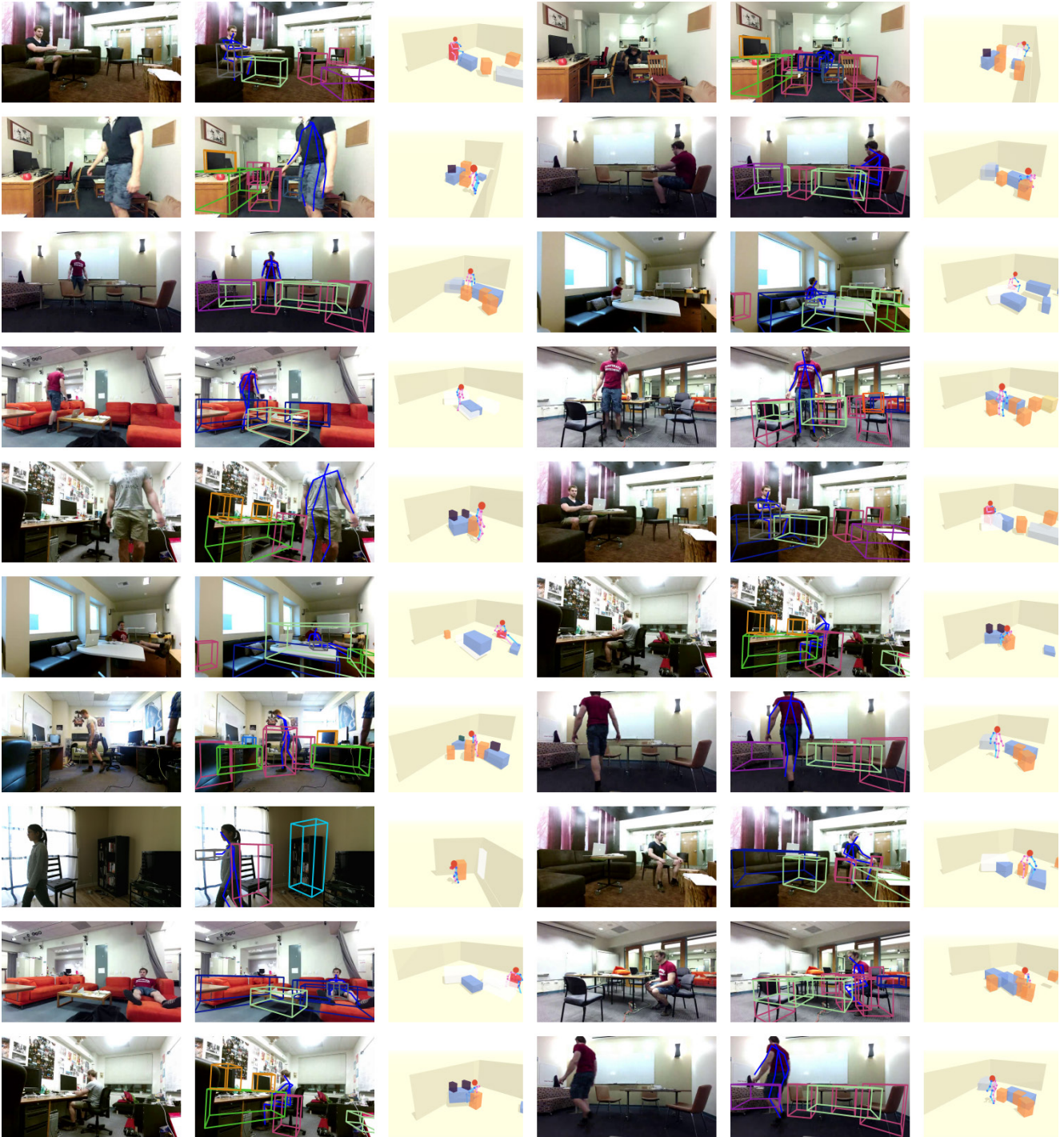


Figure 62. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

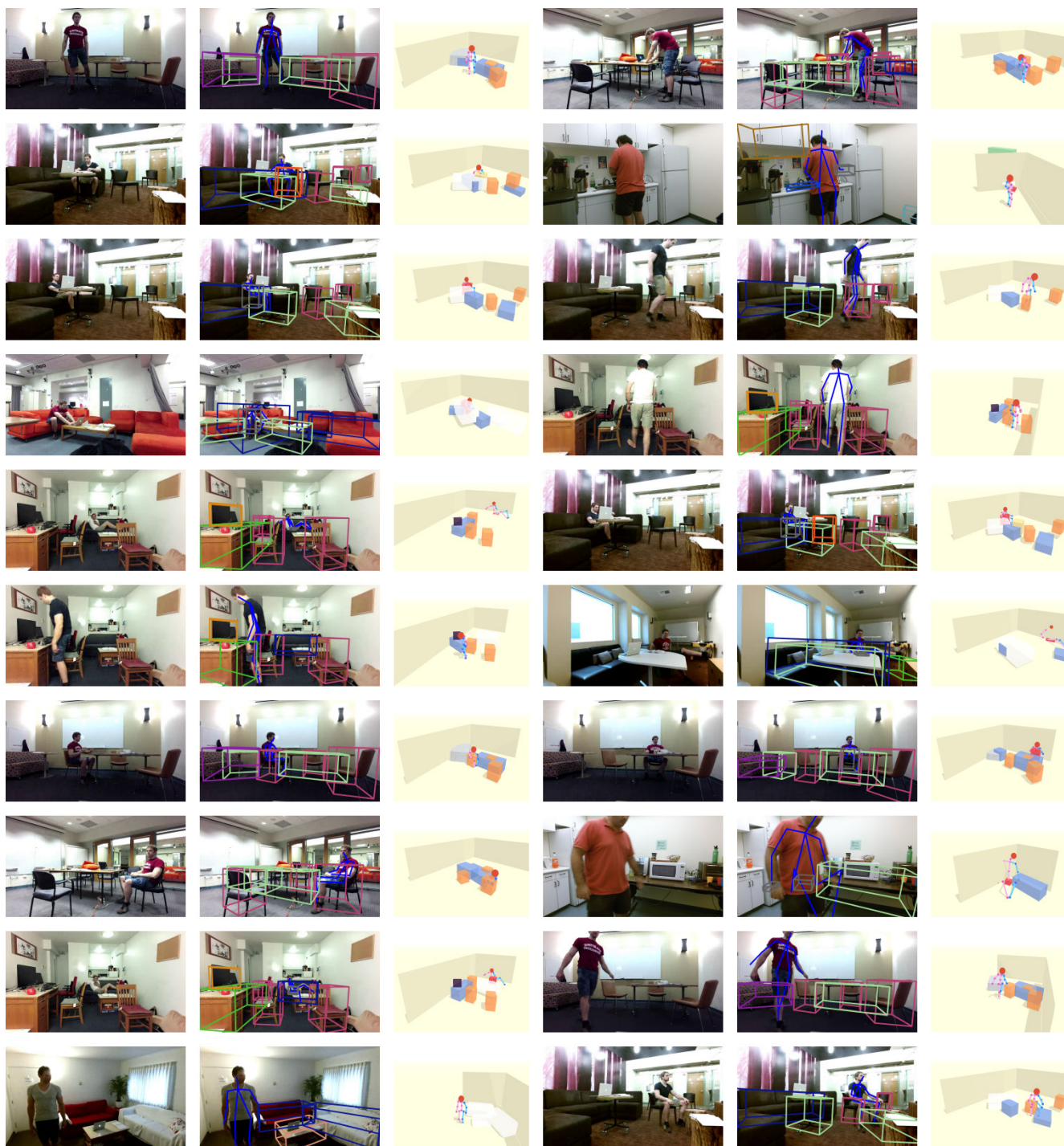


Figure 63. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.



Figure 64. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

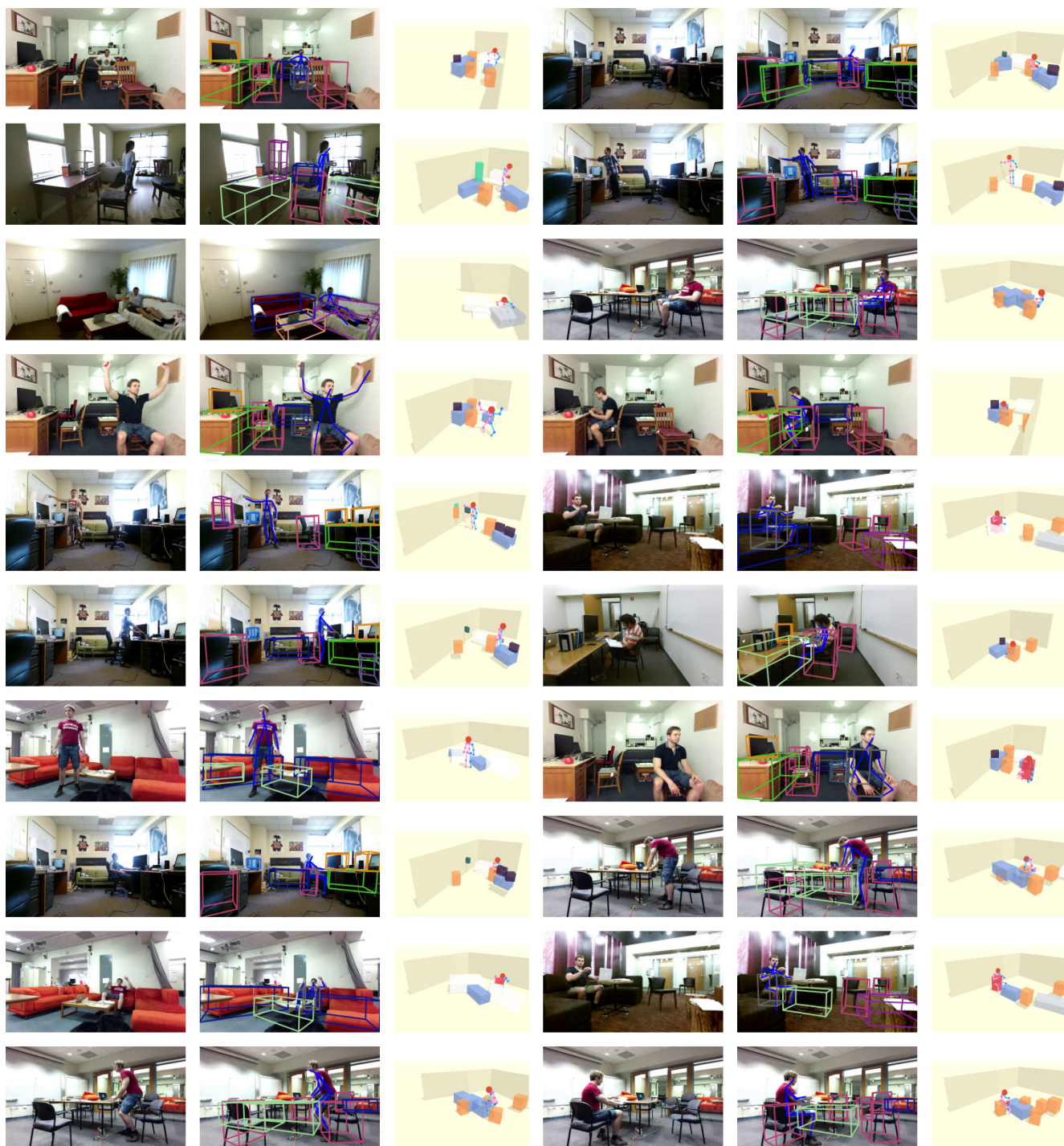


Figure 65. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

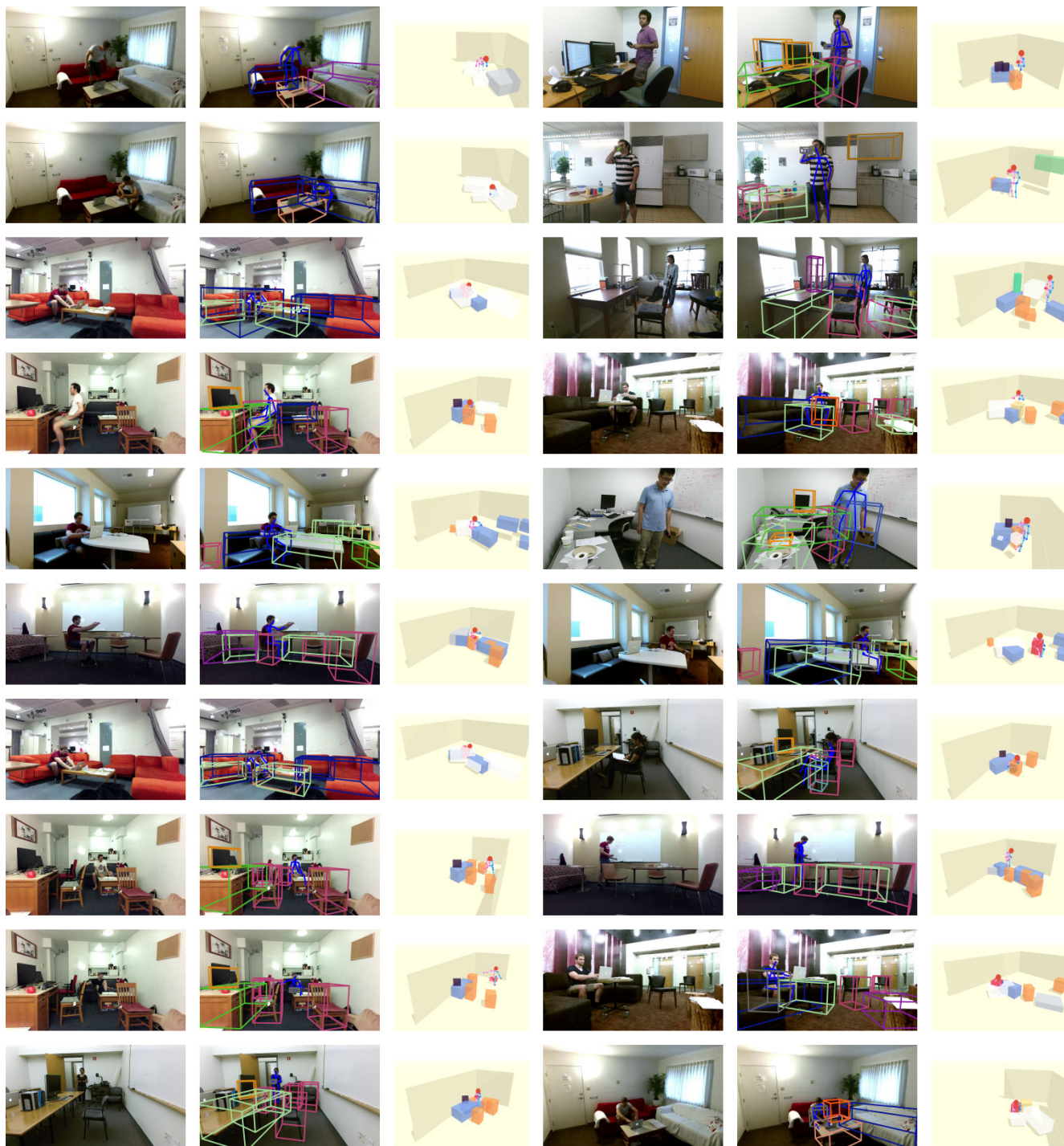


Figure 66. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

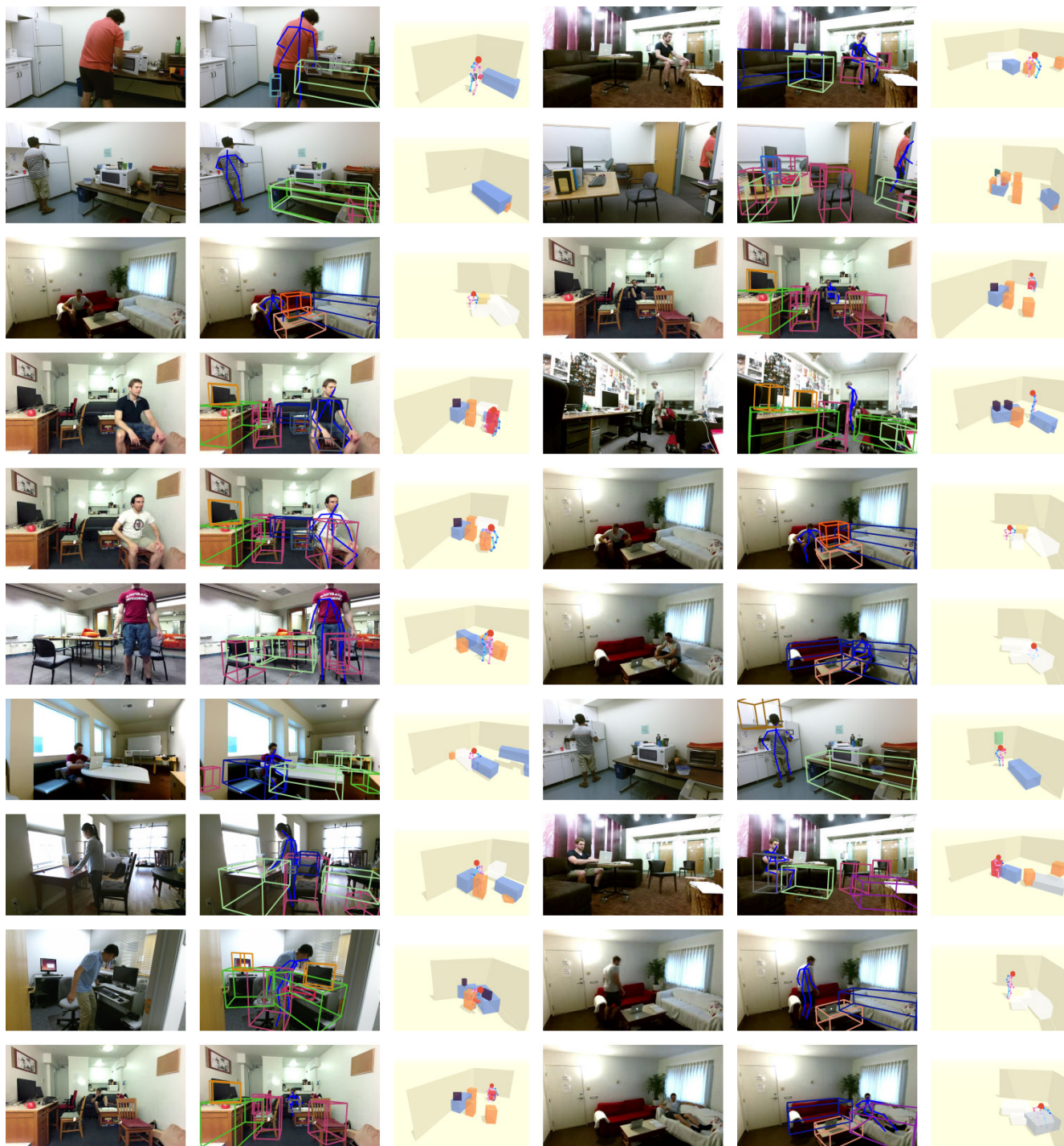


Figure 67. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

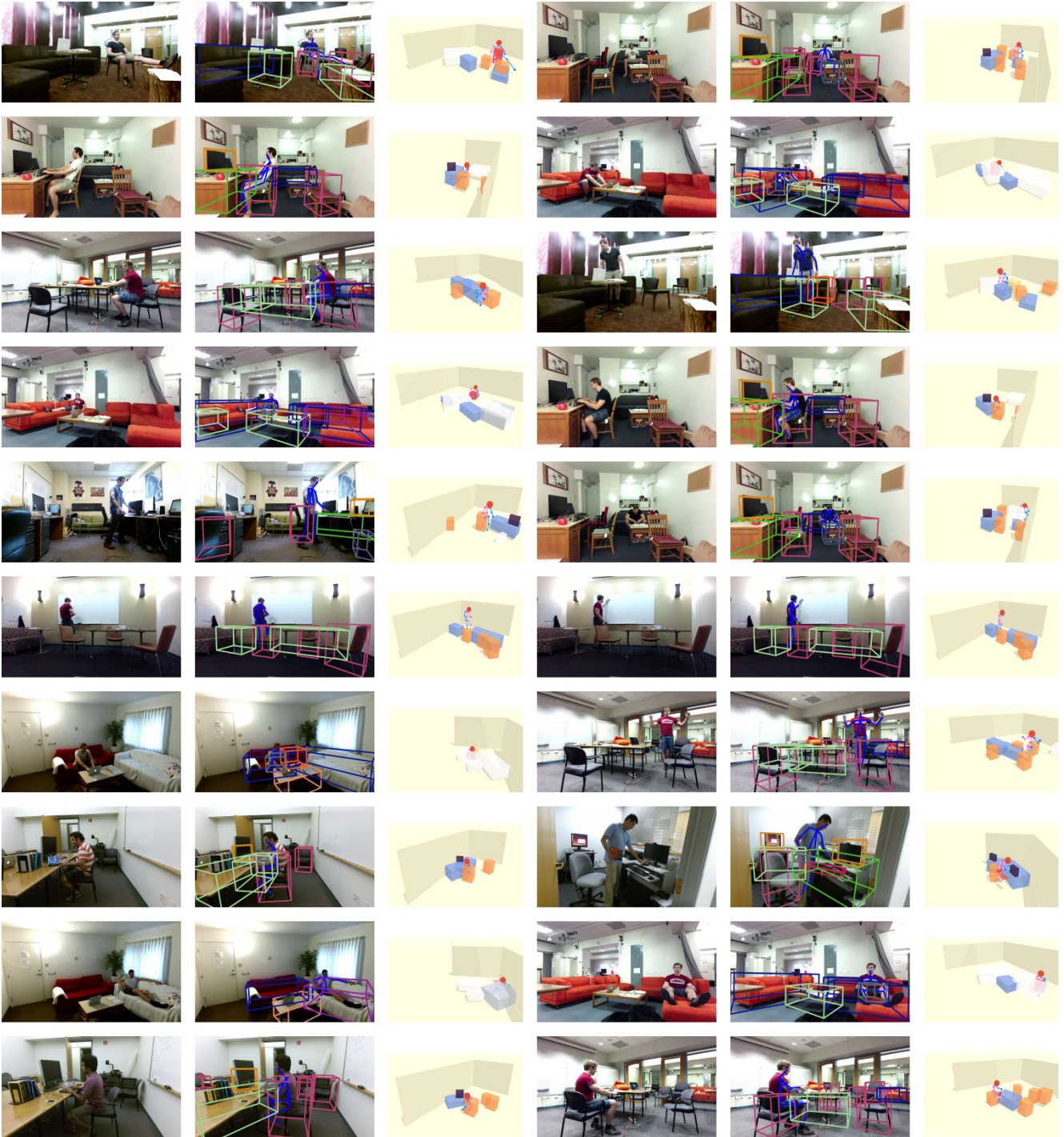


Figure 68. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

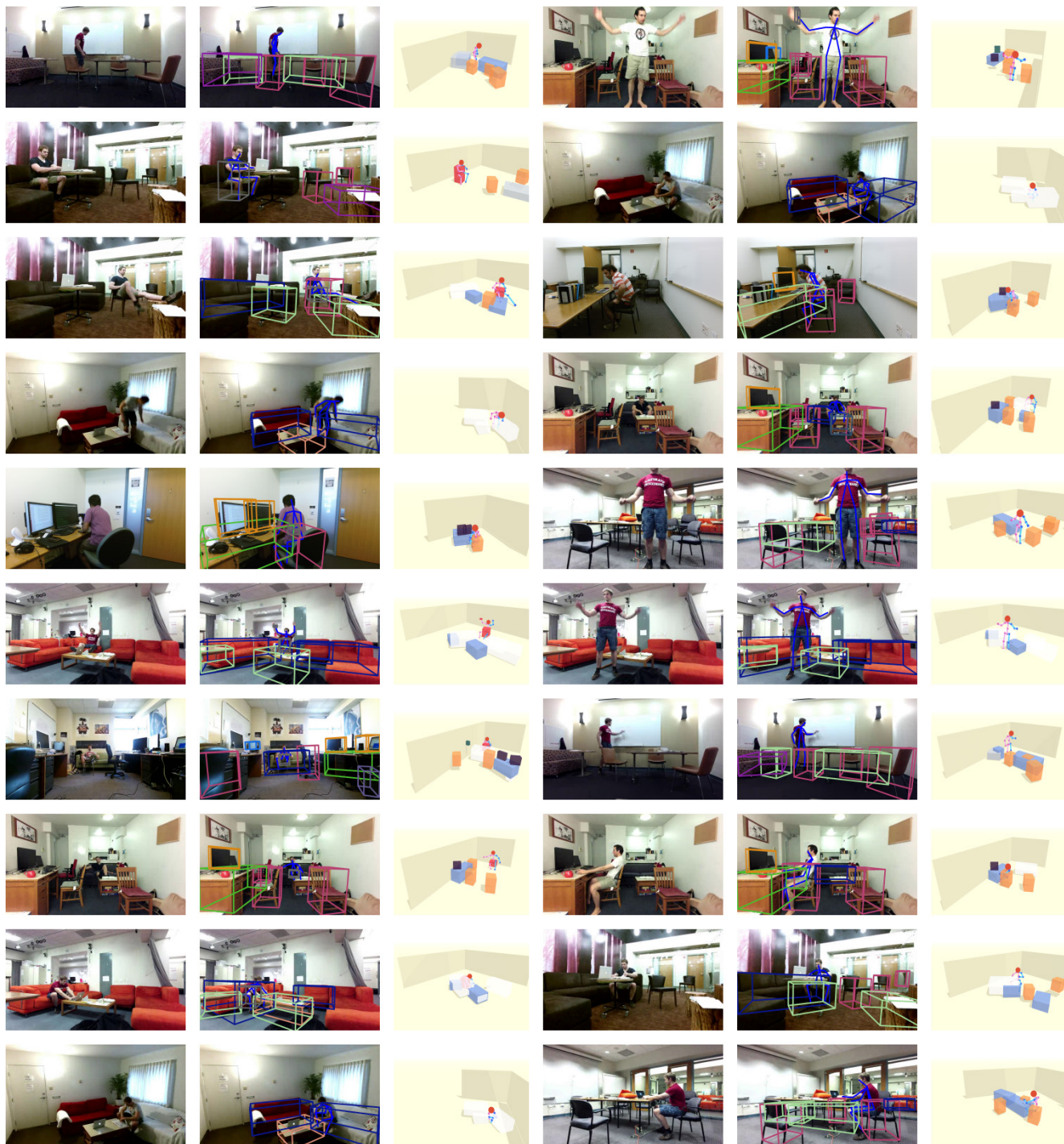


Figure 69. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

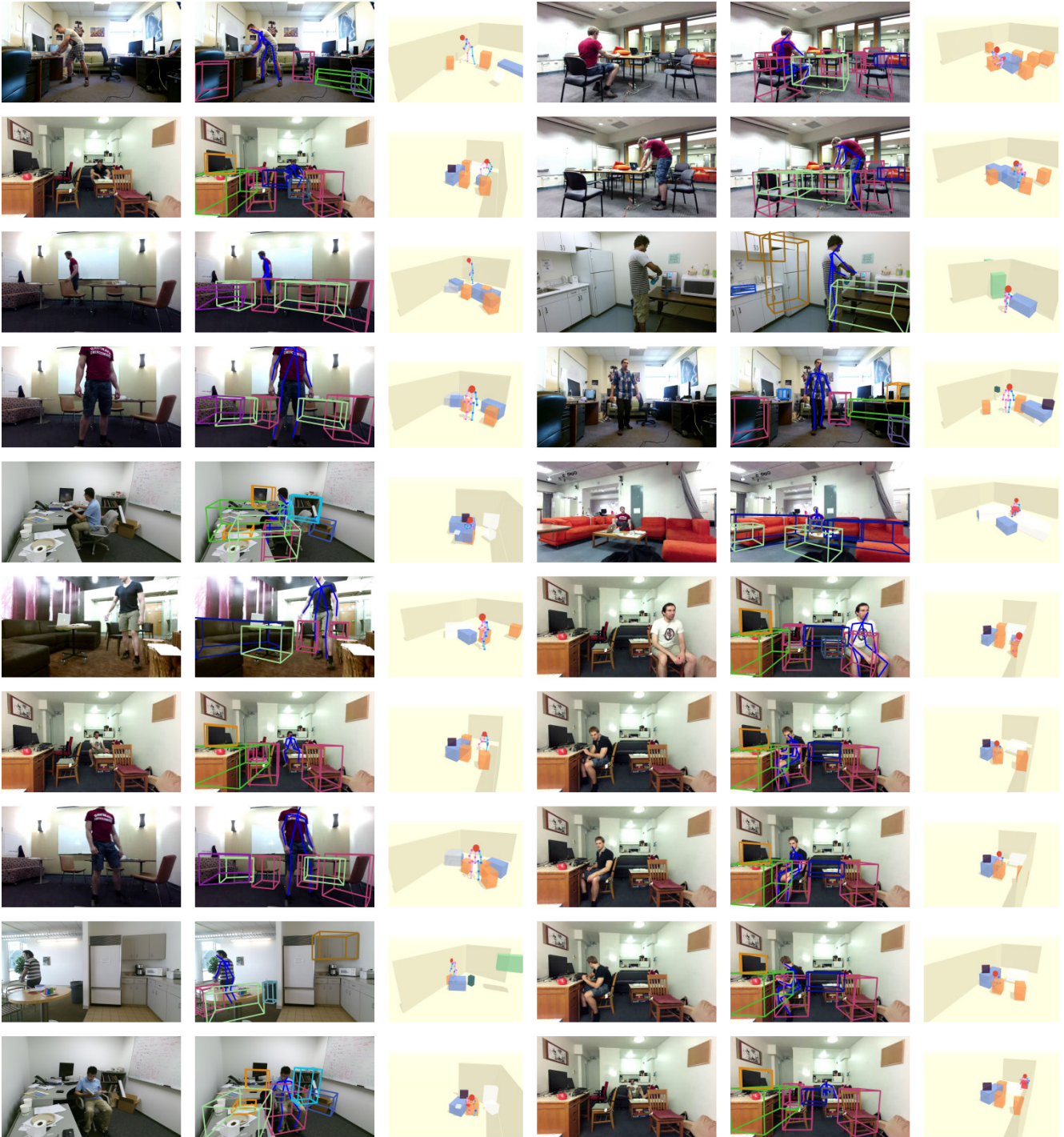


Figure 70. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

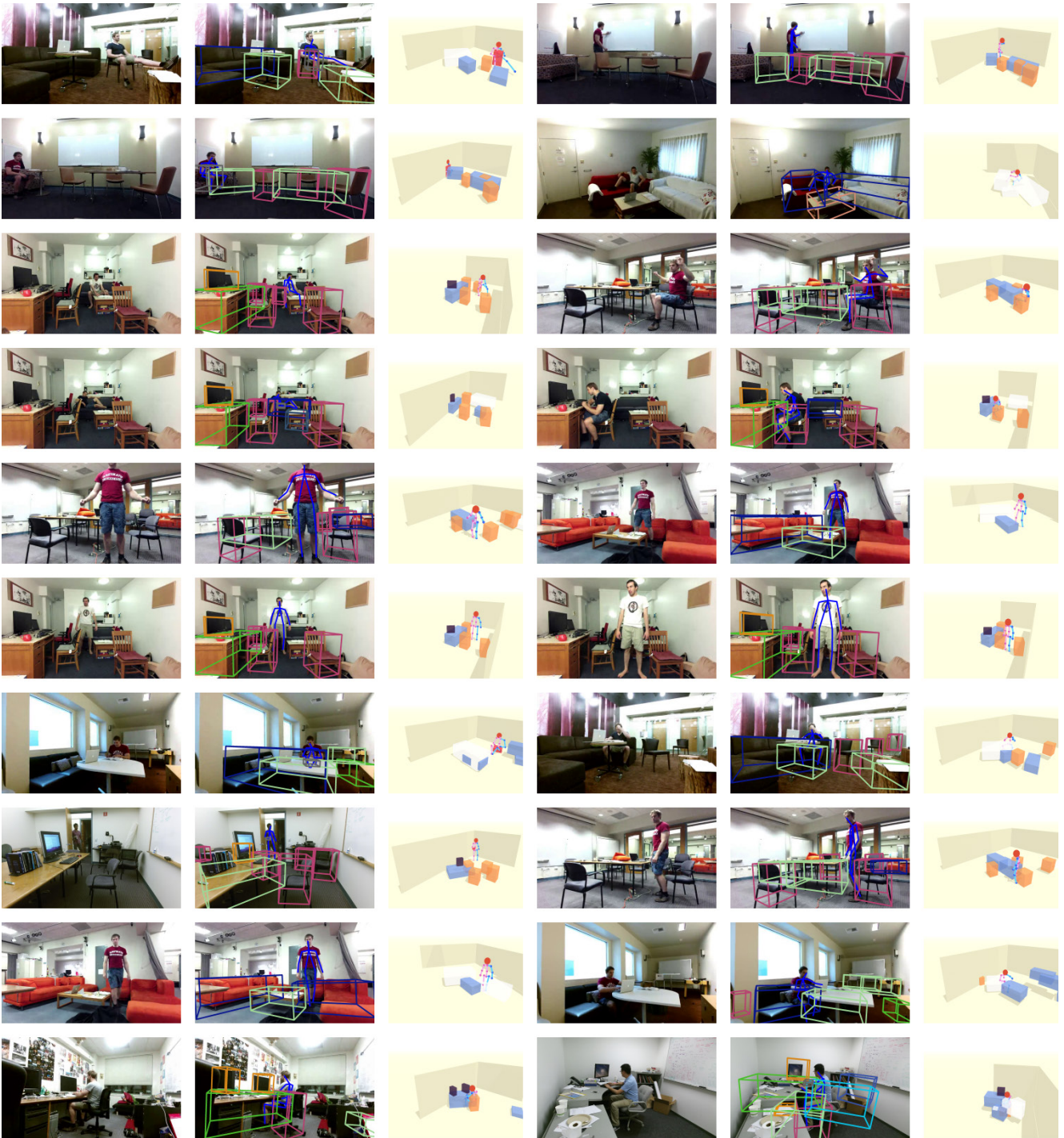


Figure 71. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

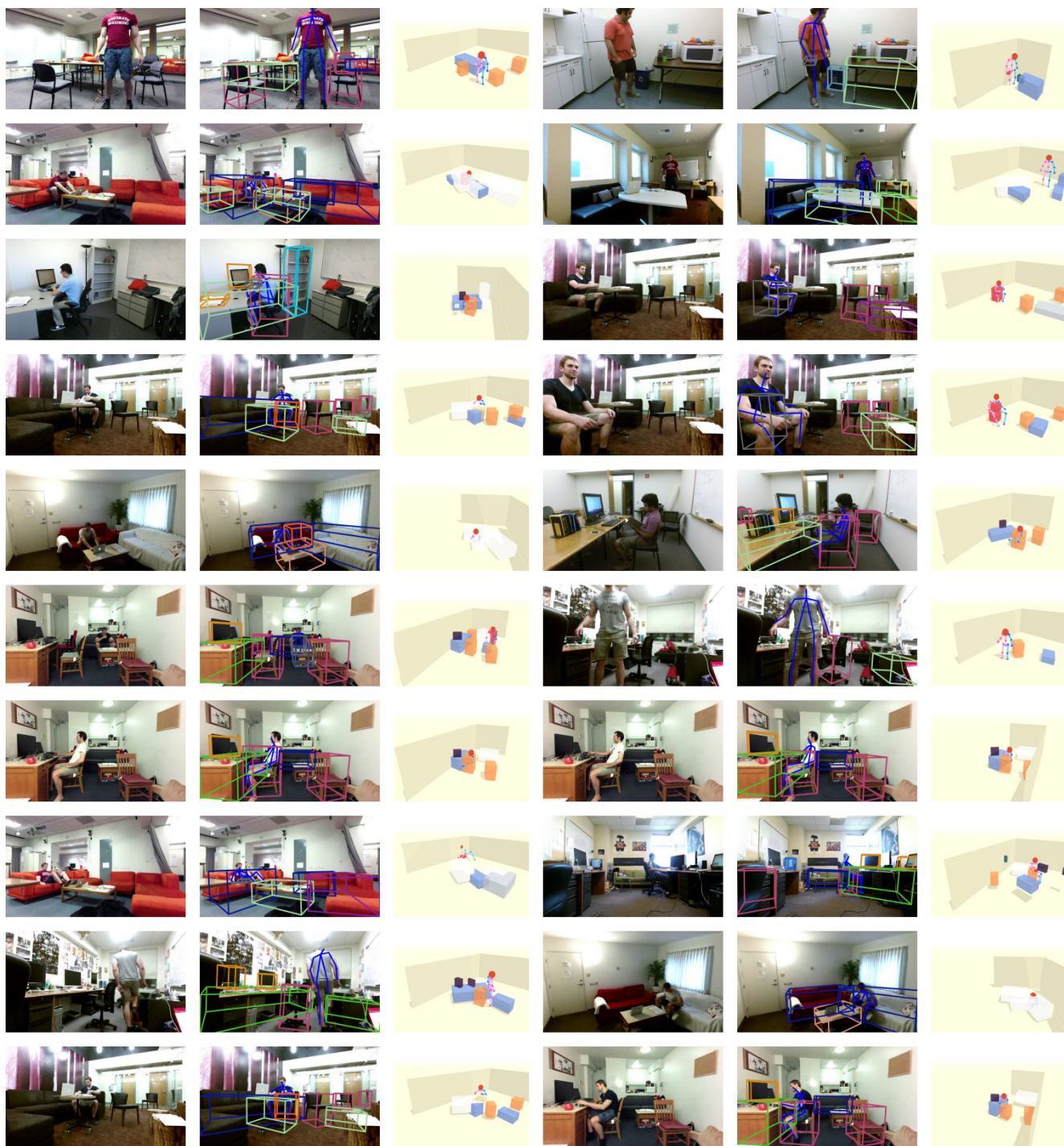


Figure 72. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

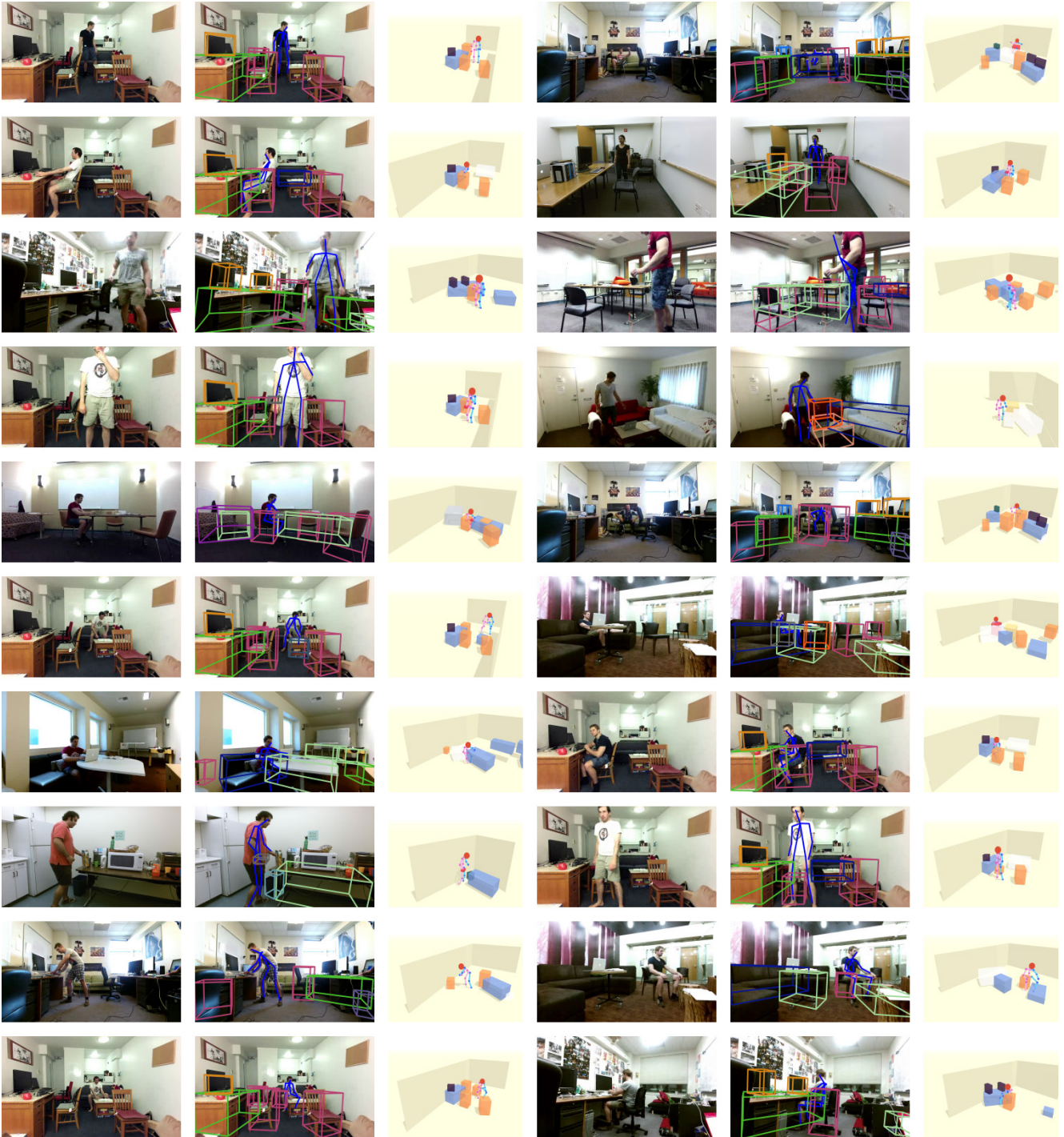


Figure 73. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

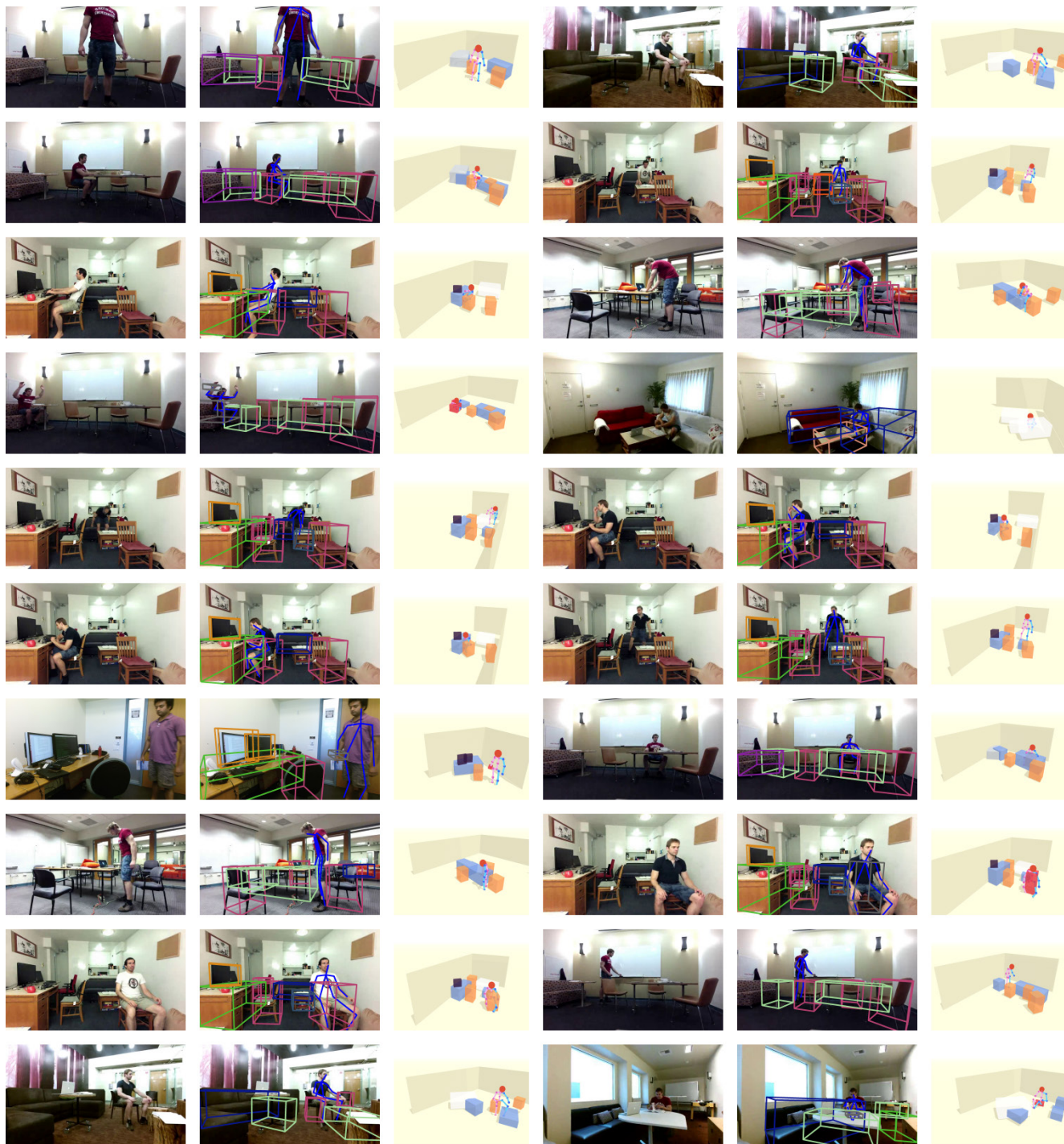


Figure 74. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

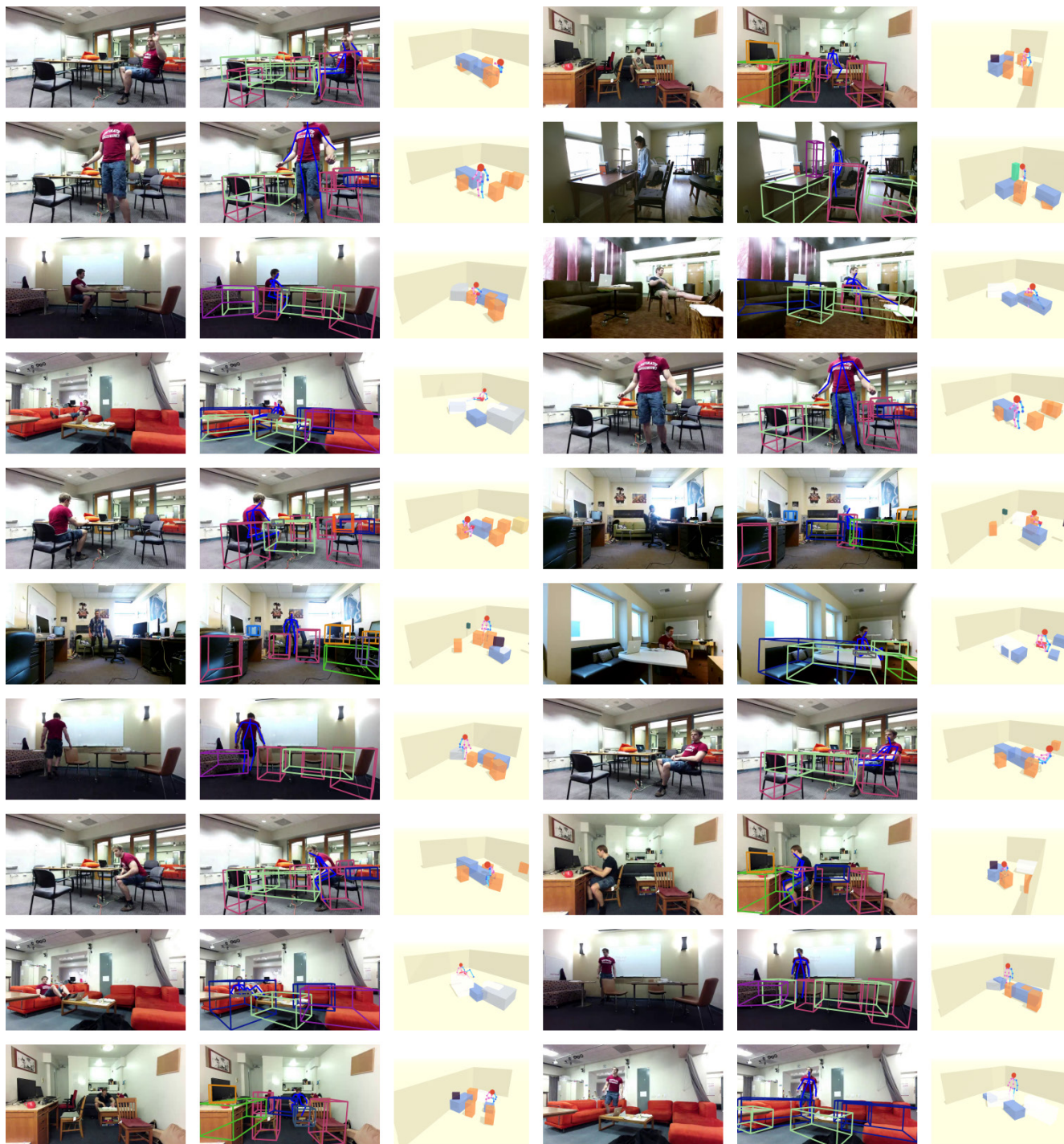


Figure 75. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

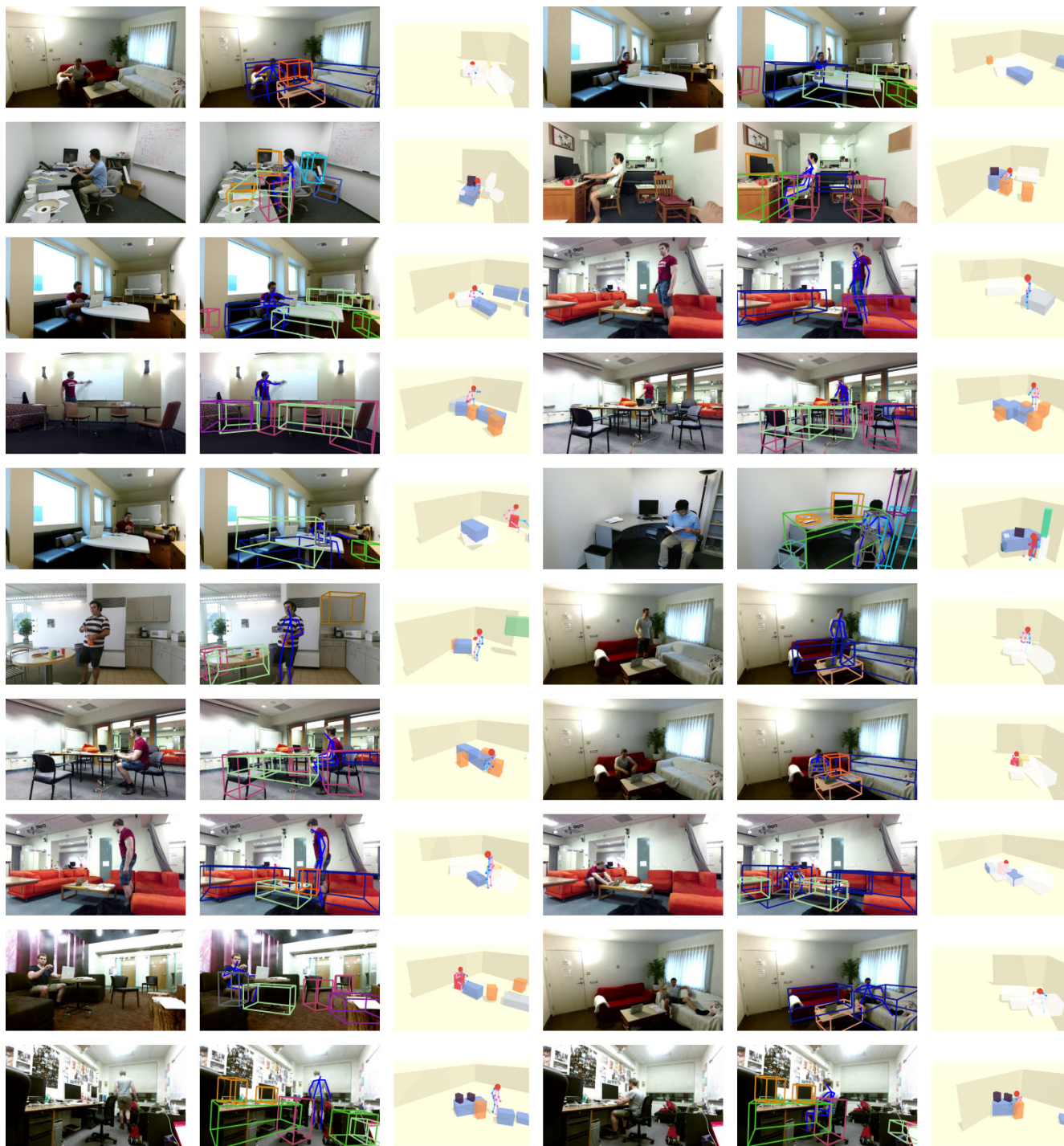


Figure 76. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

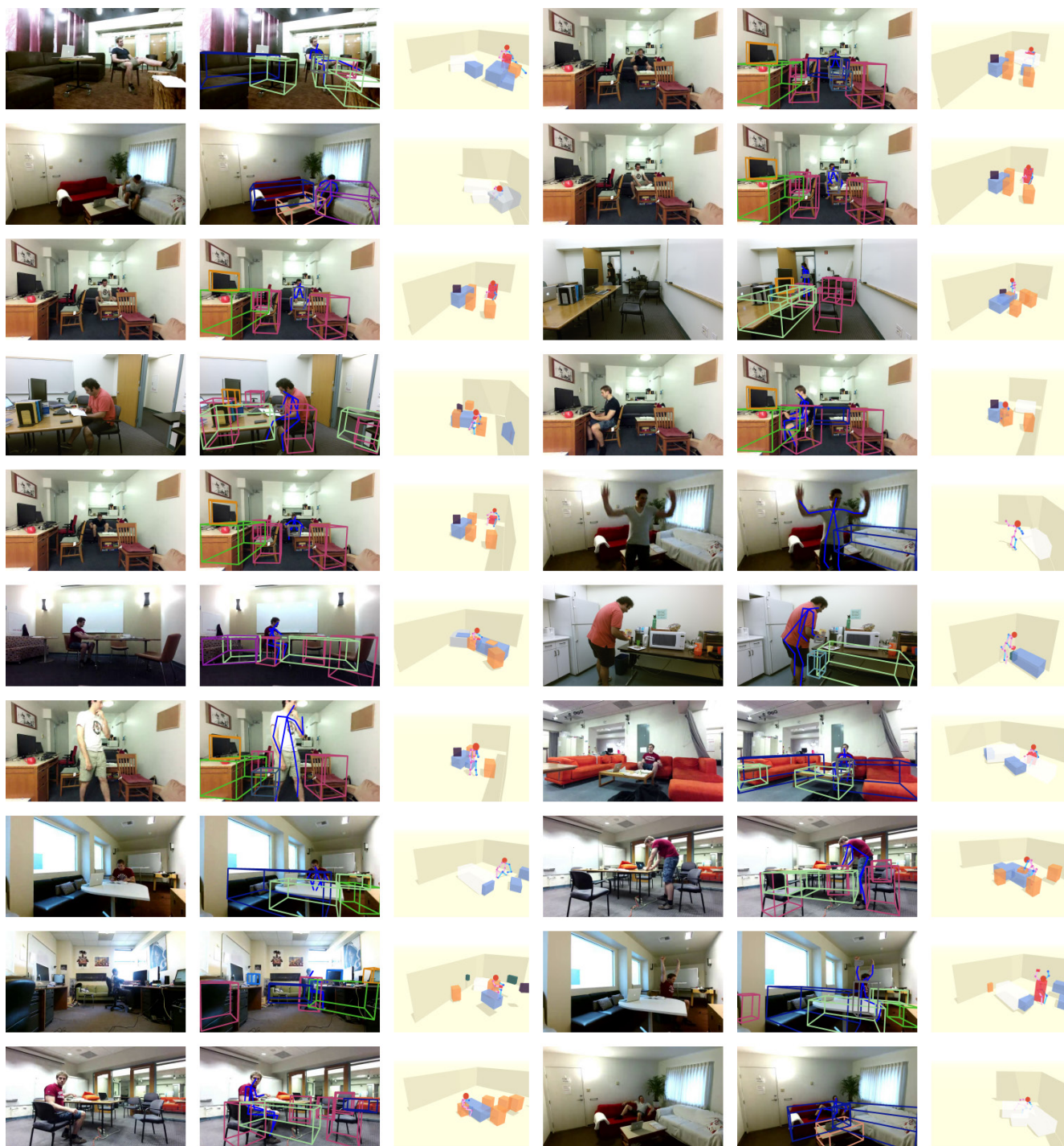


Figure 77. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

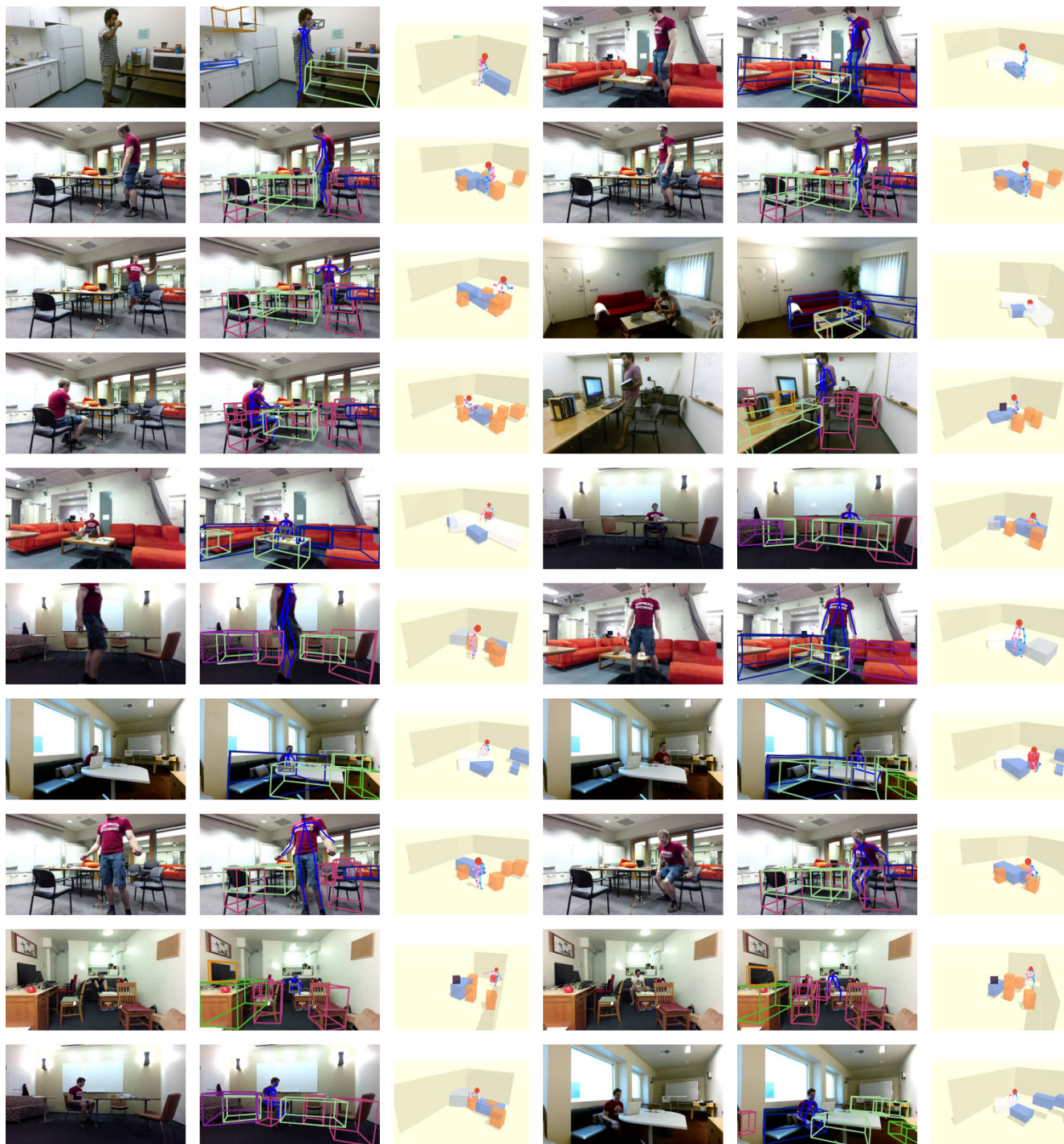


Figure 78. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

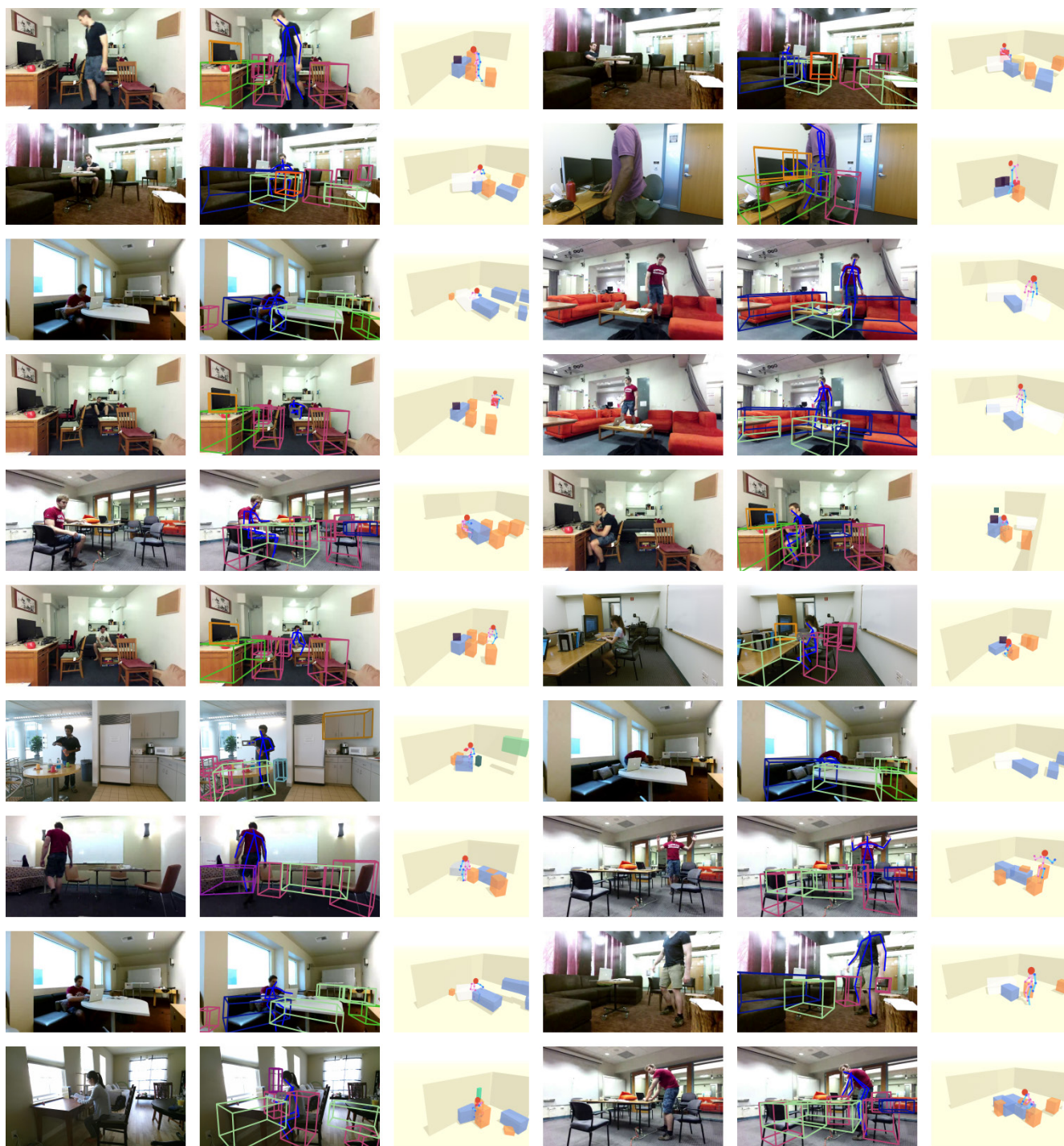


Figure 79. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

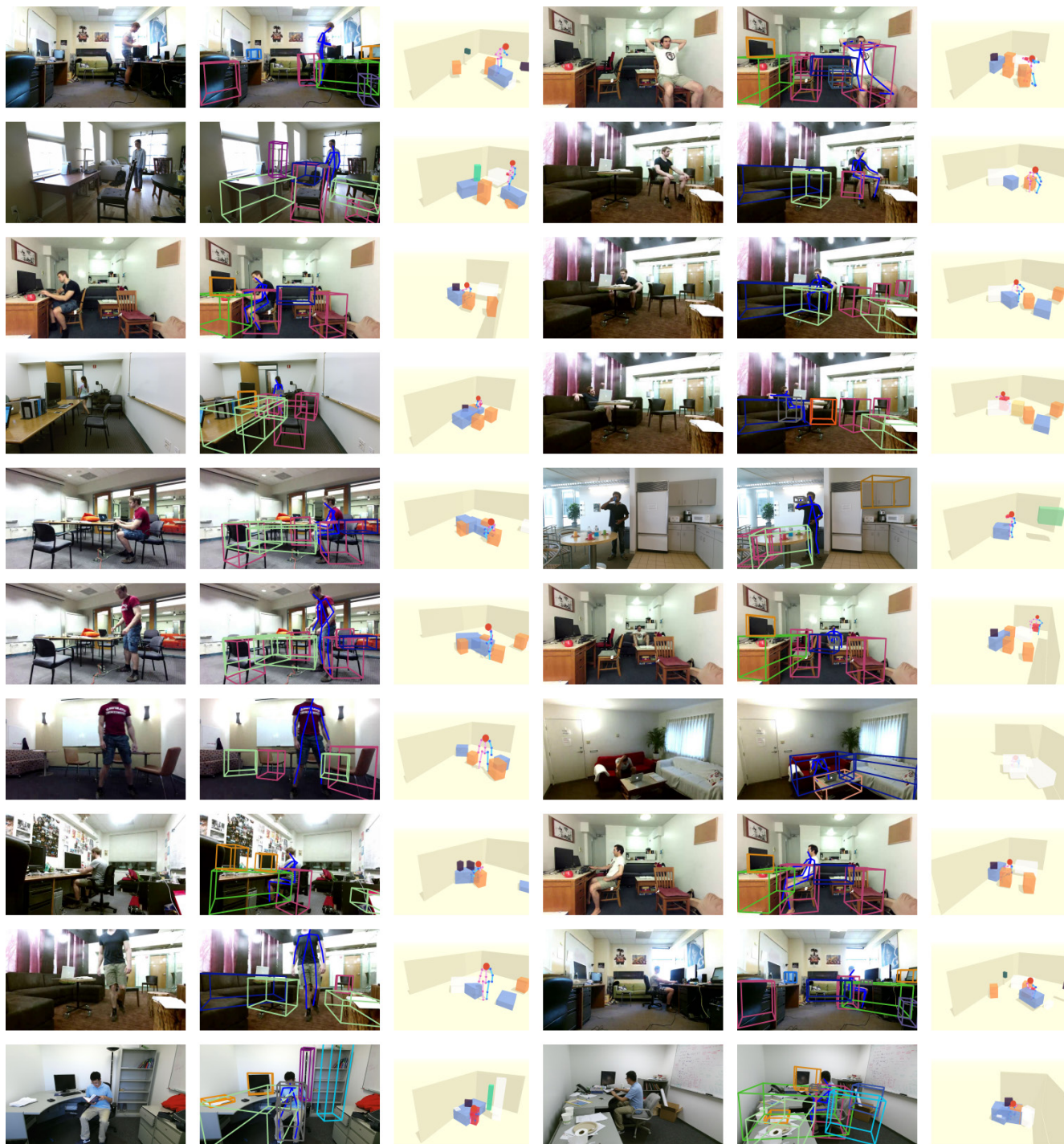


Figure 80. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

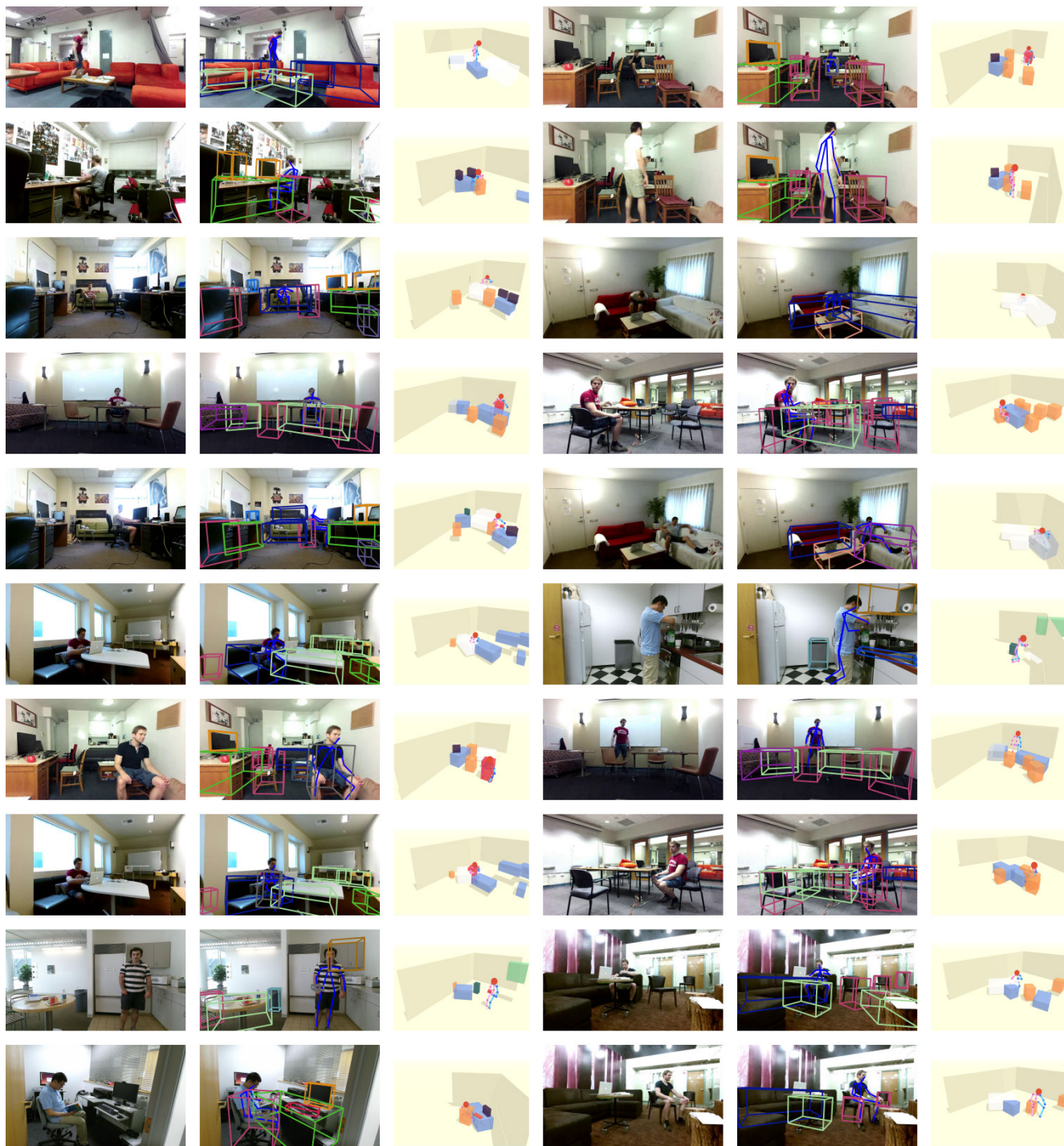


Figure 81. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

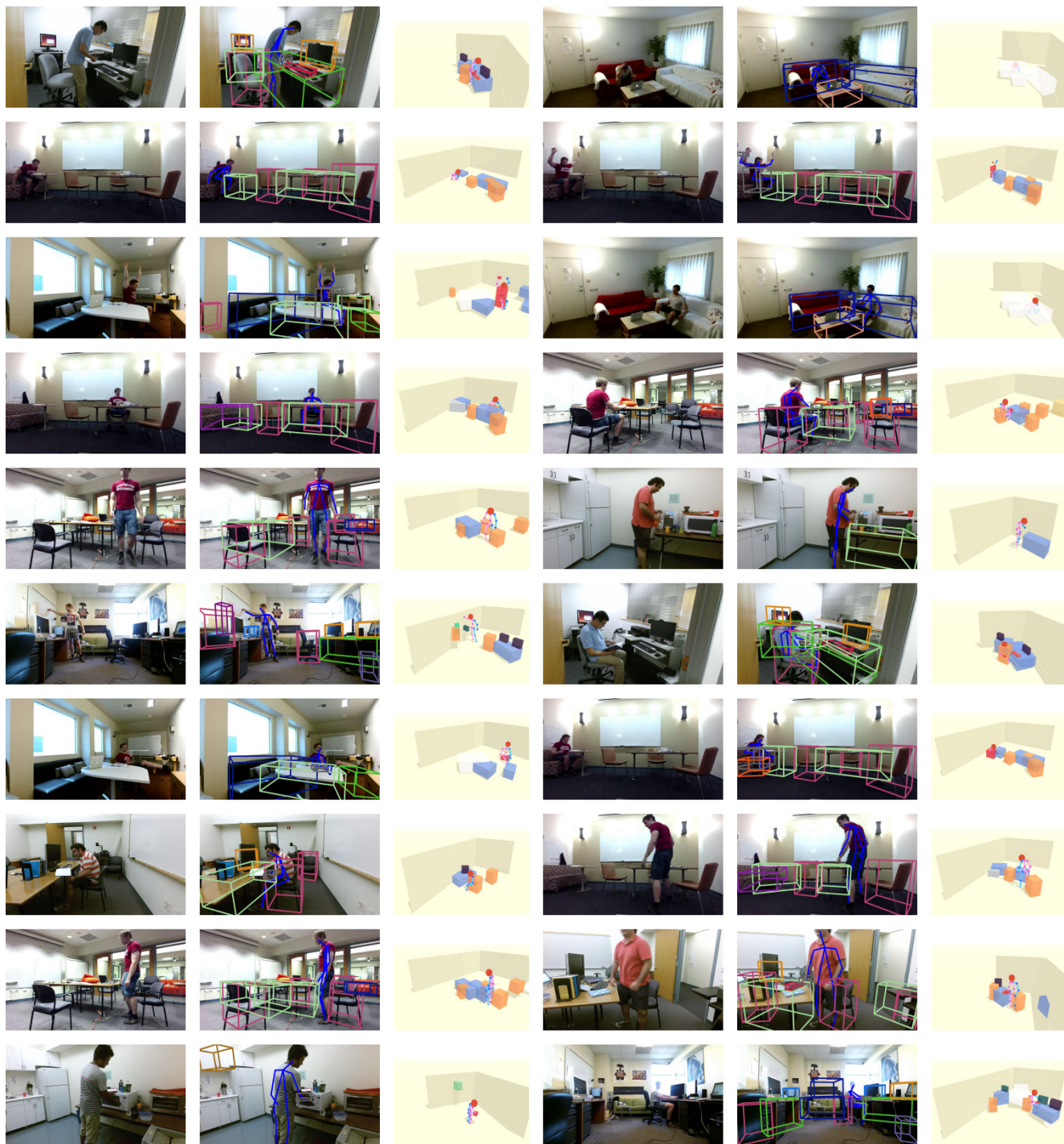


Figure 82. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

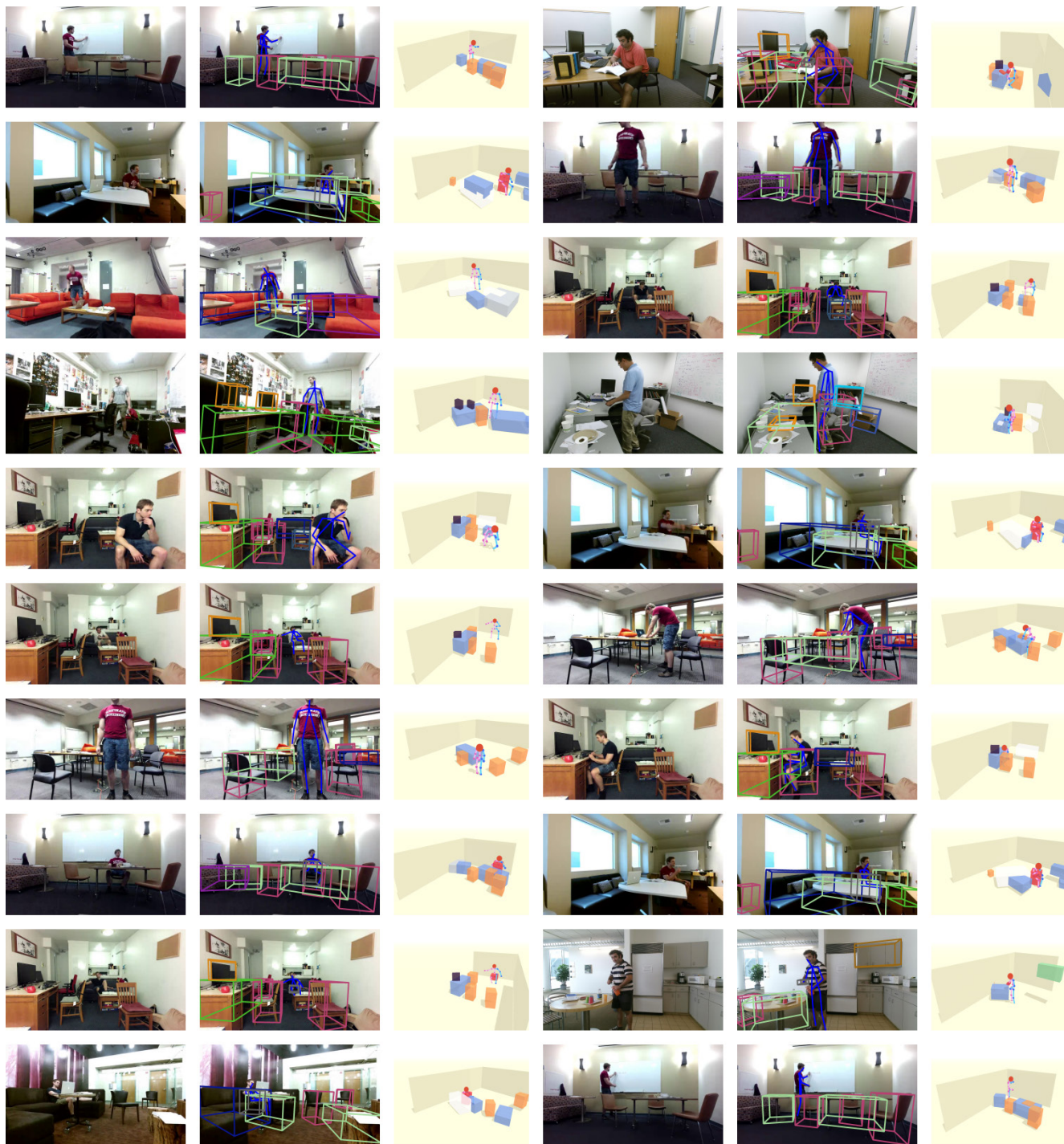


Figure 83. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

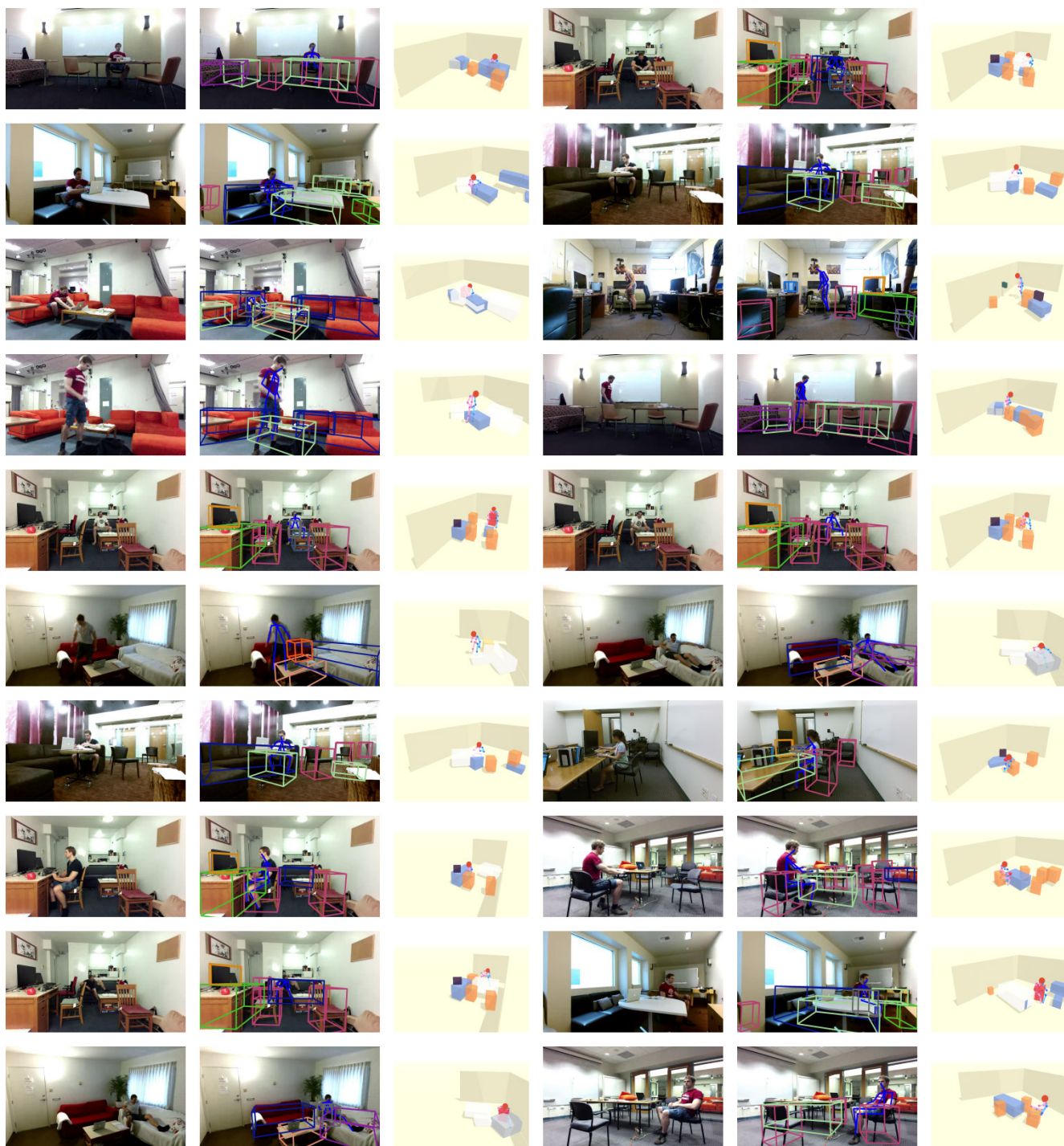


Figure 84. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

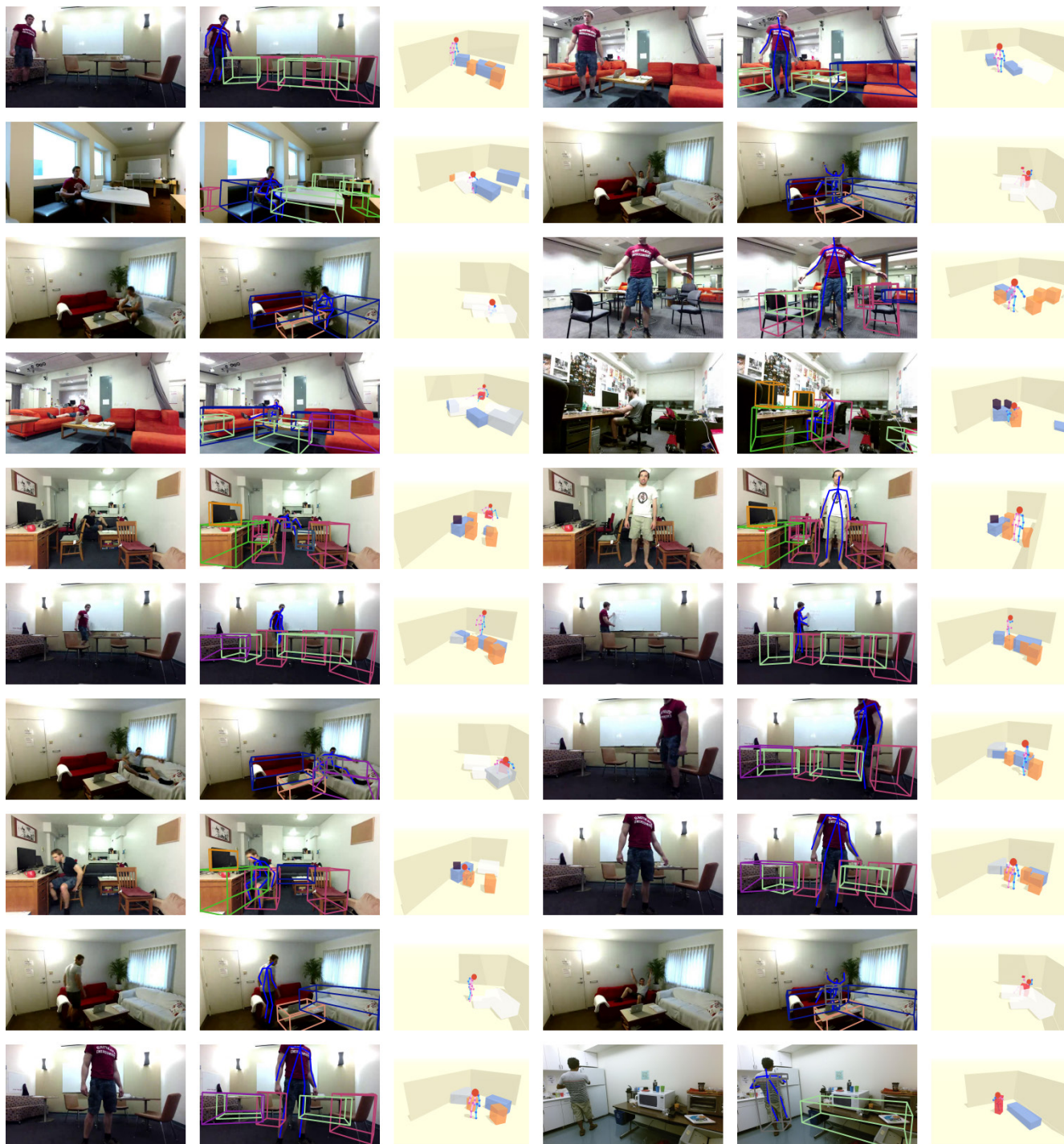


Figure 85. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

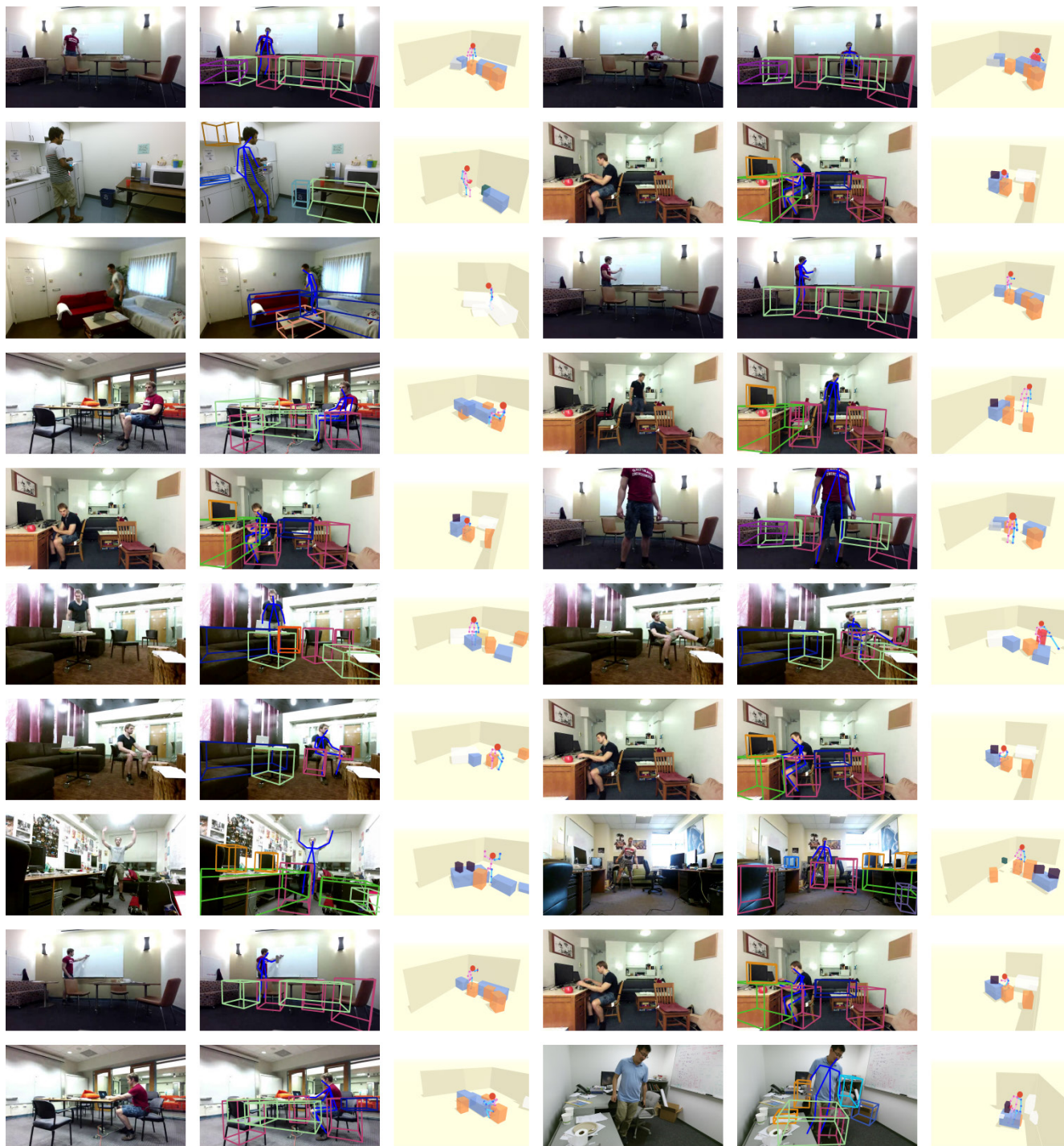


Figure 86. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

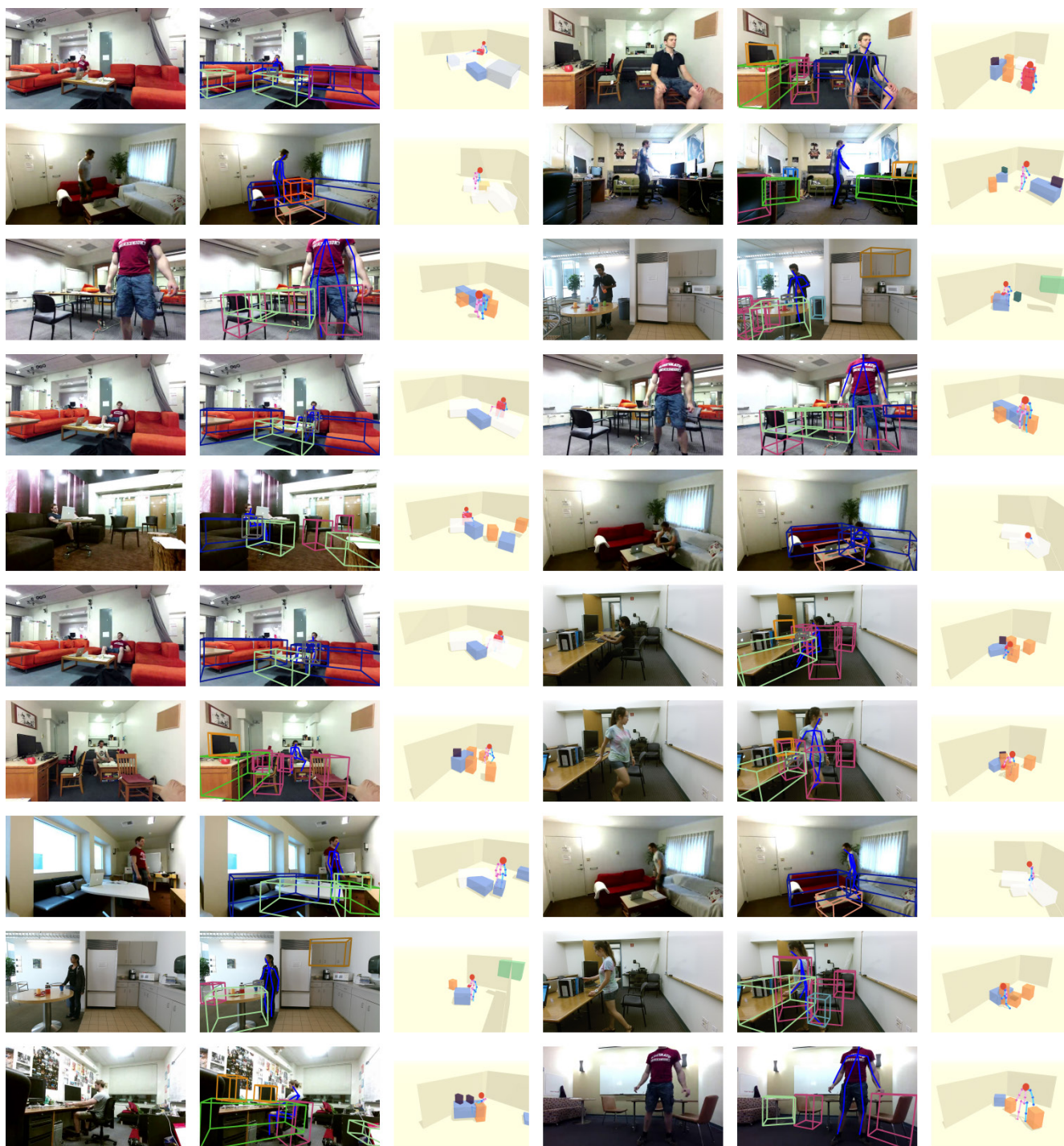


Figure 87. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

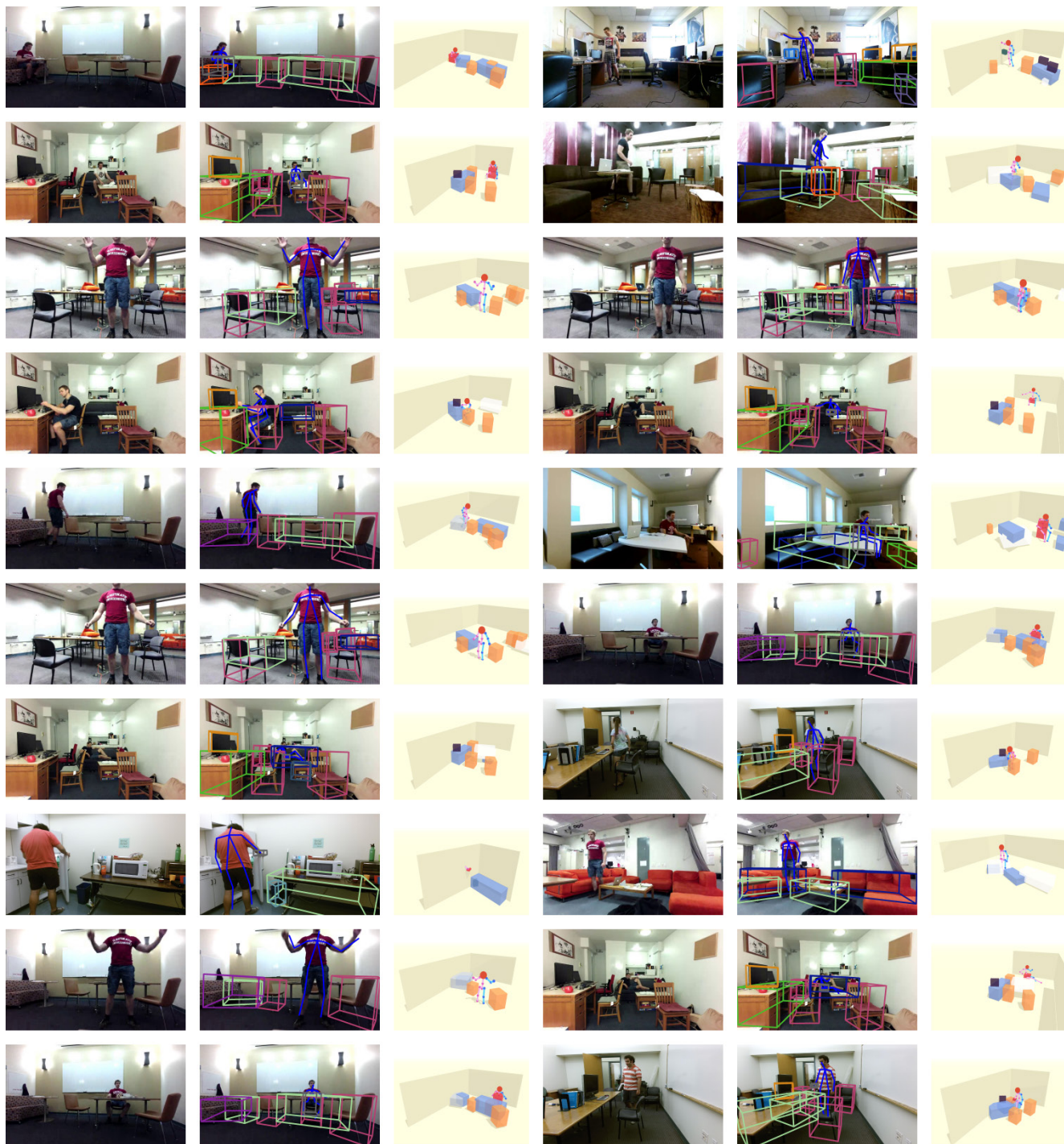


Figure 88. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

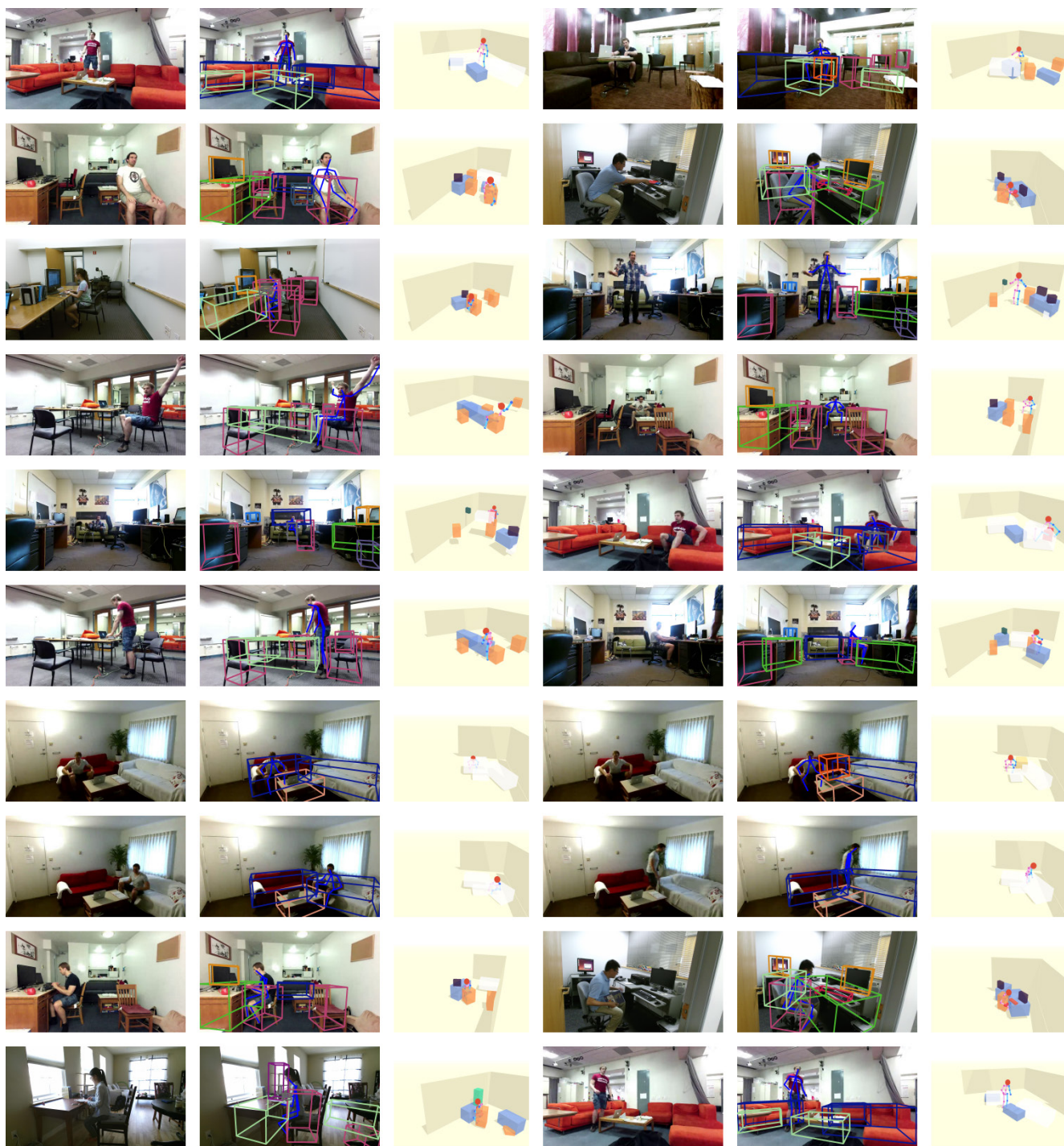


Figure 89. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

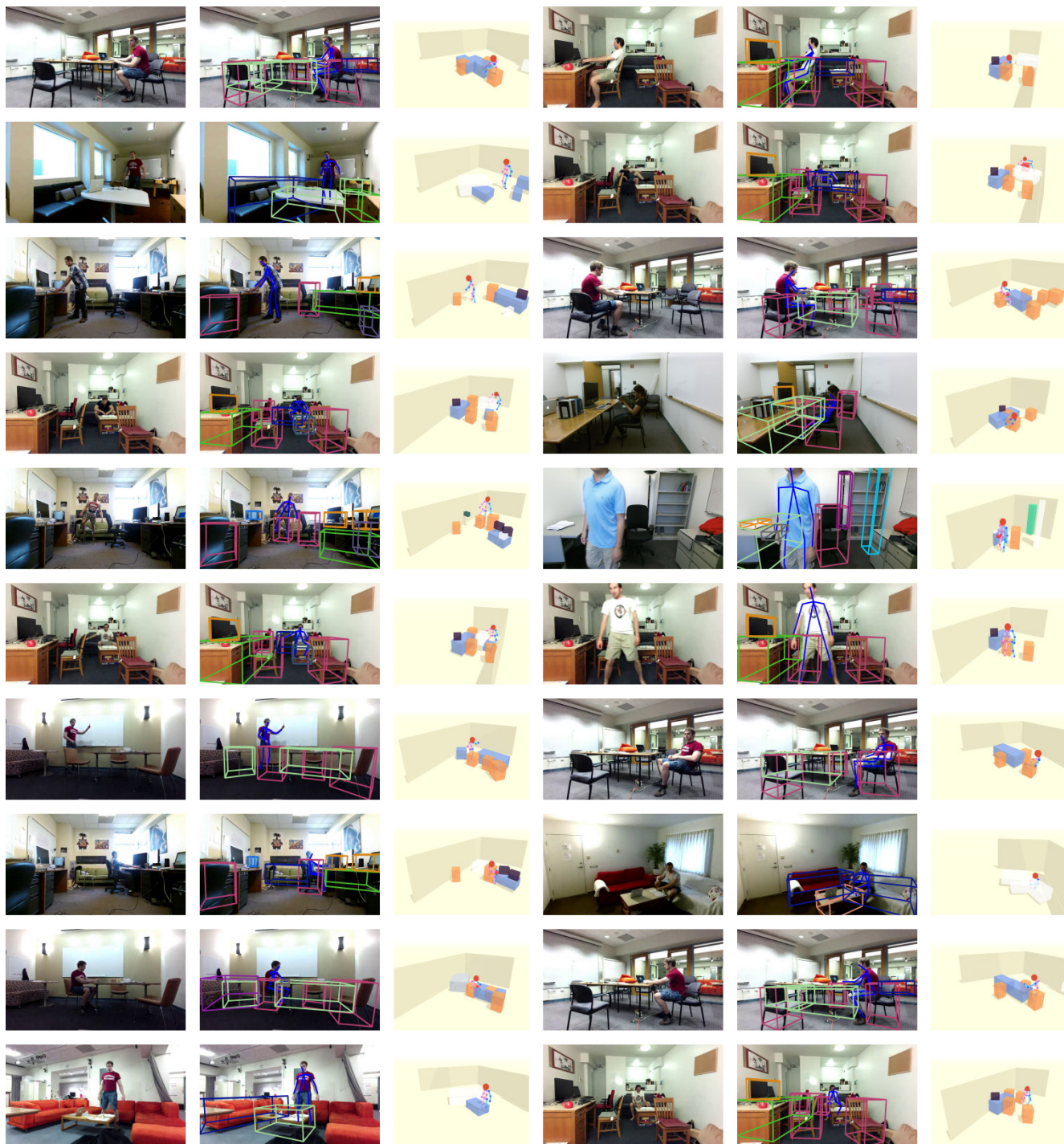
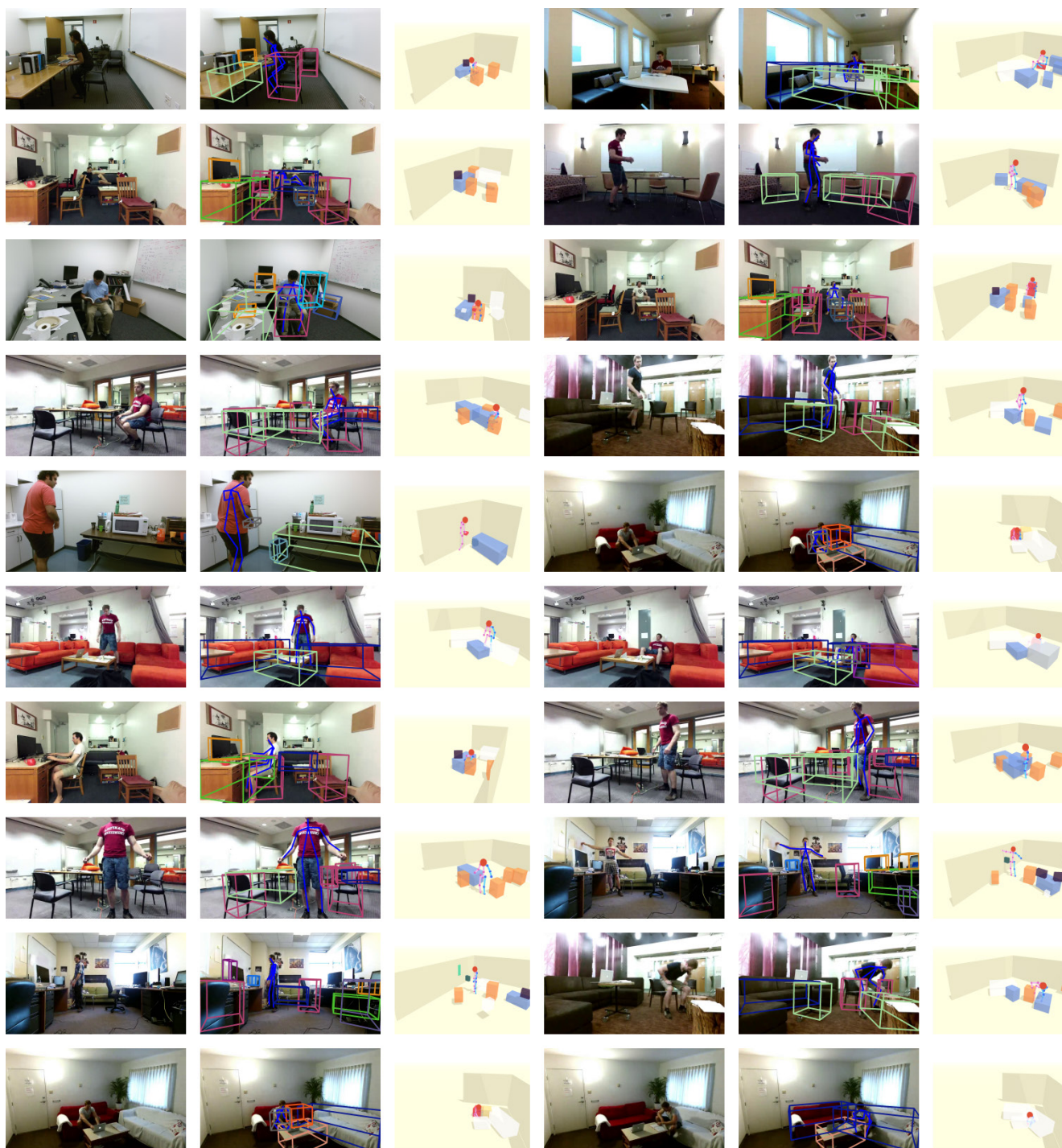


Figure 90. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.



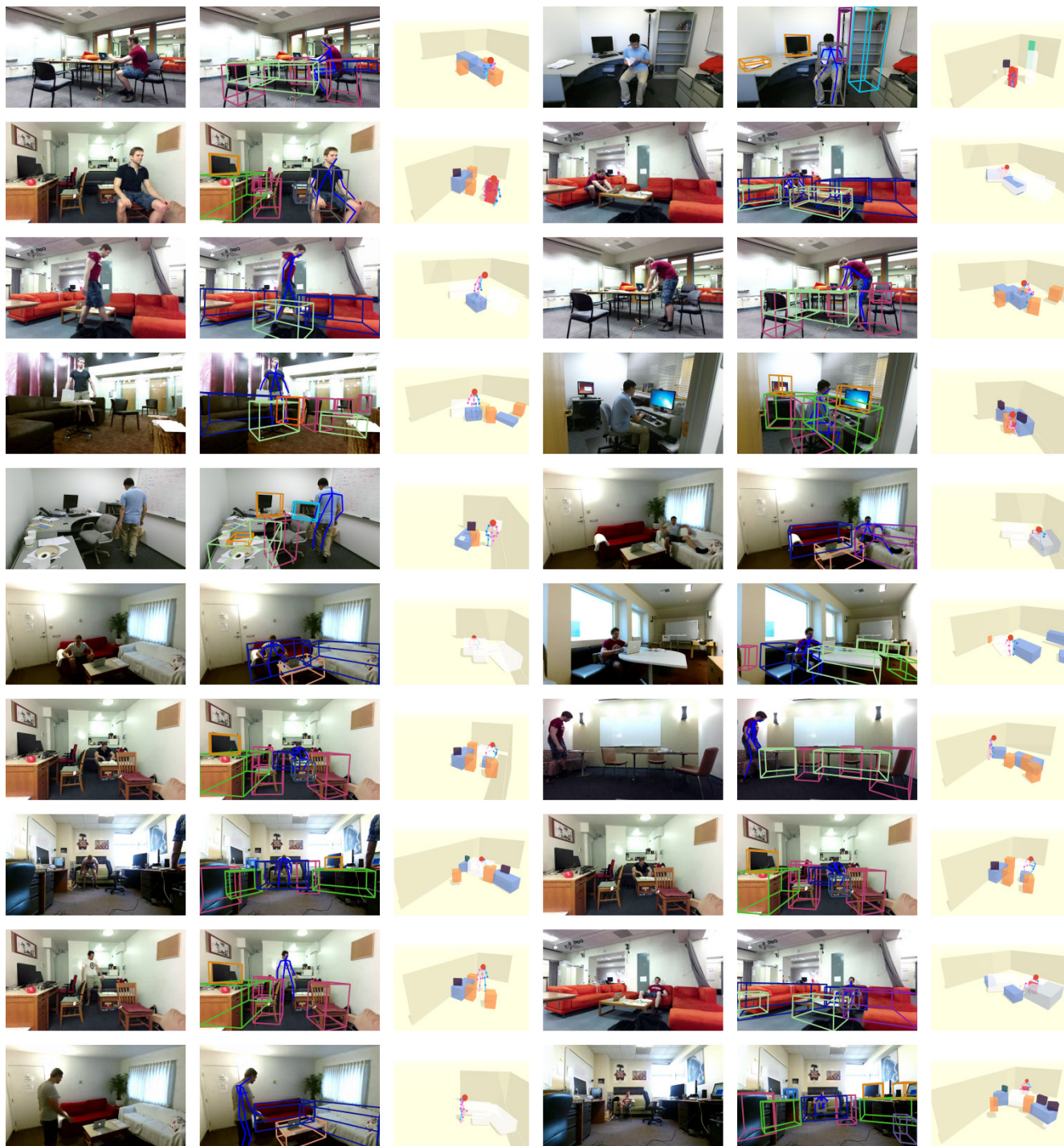


Figure 92. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

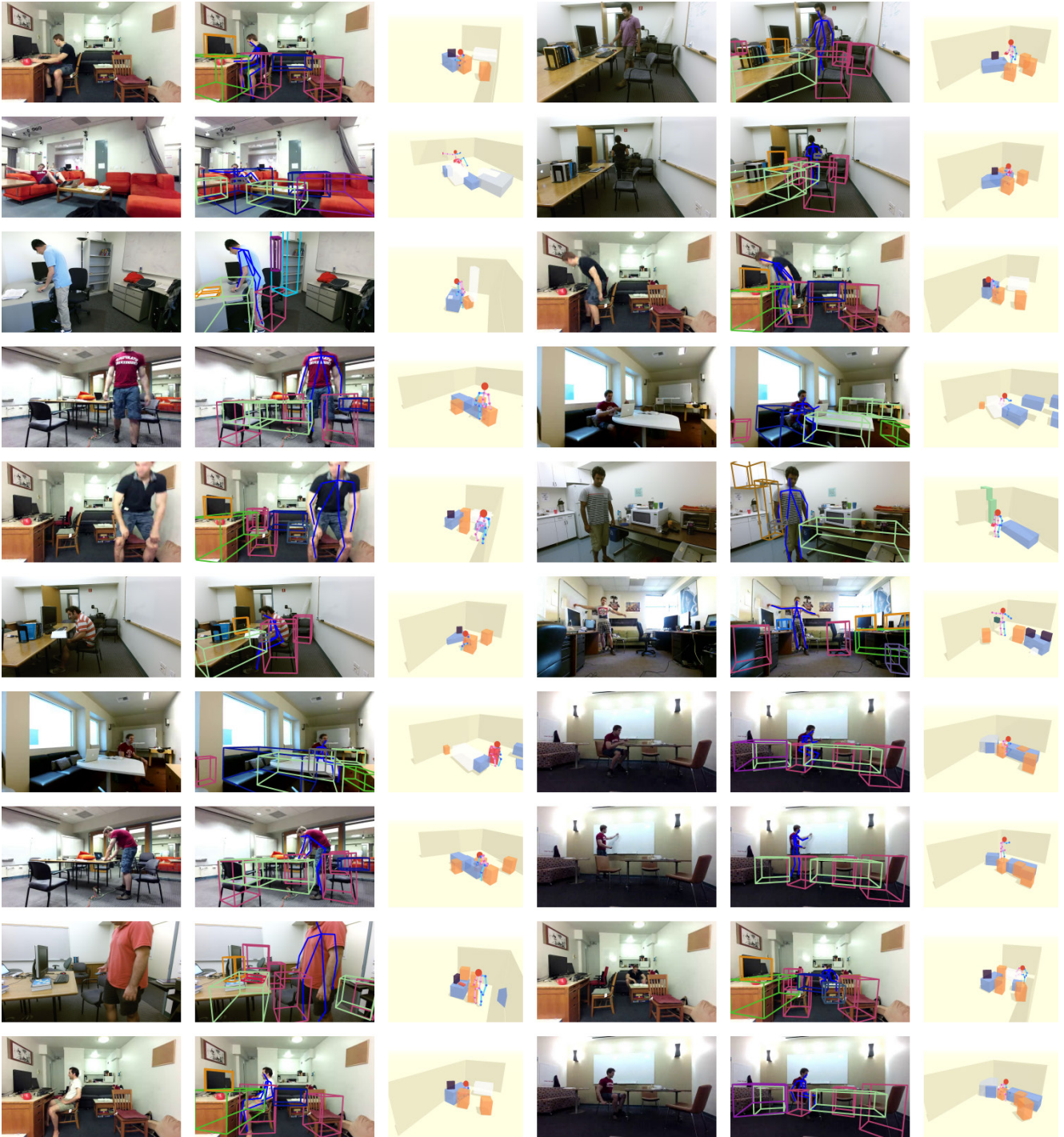


Figure 93. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

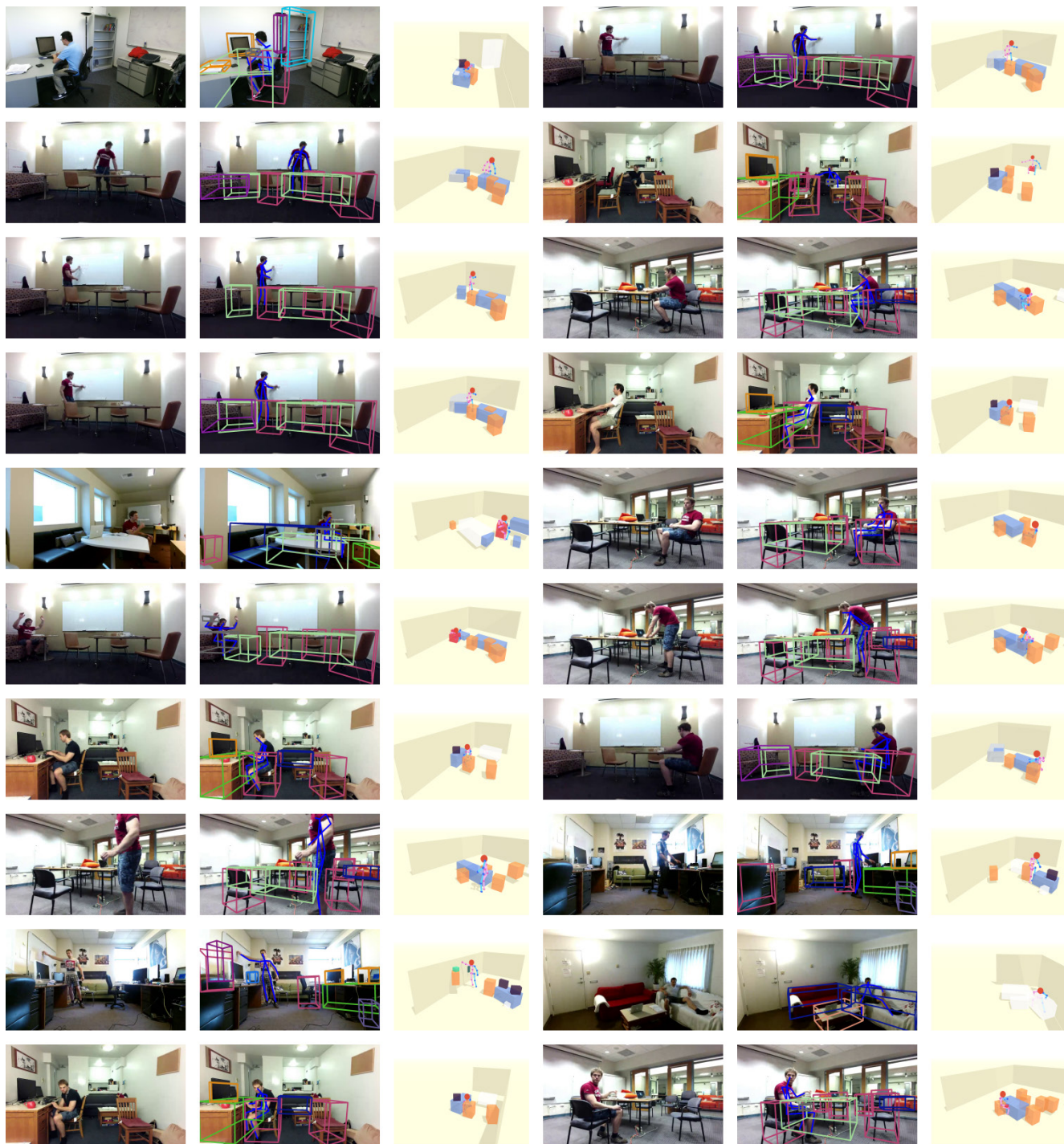


Figure 94. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

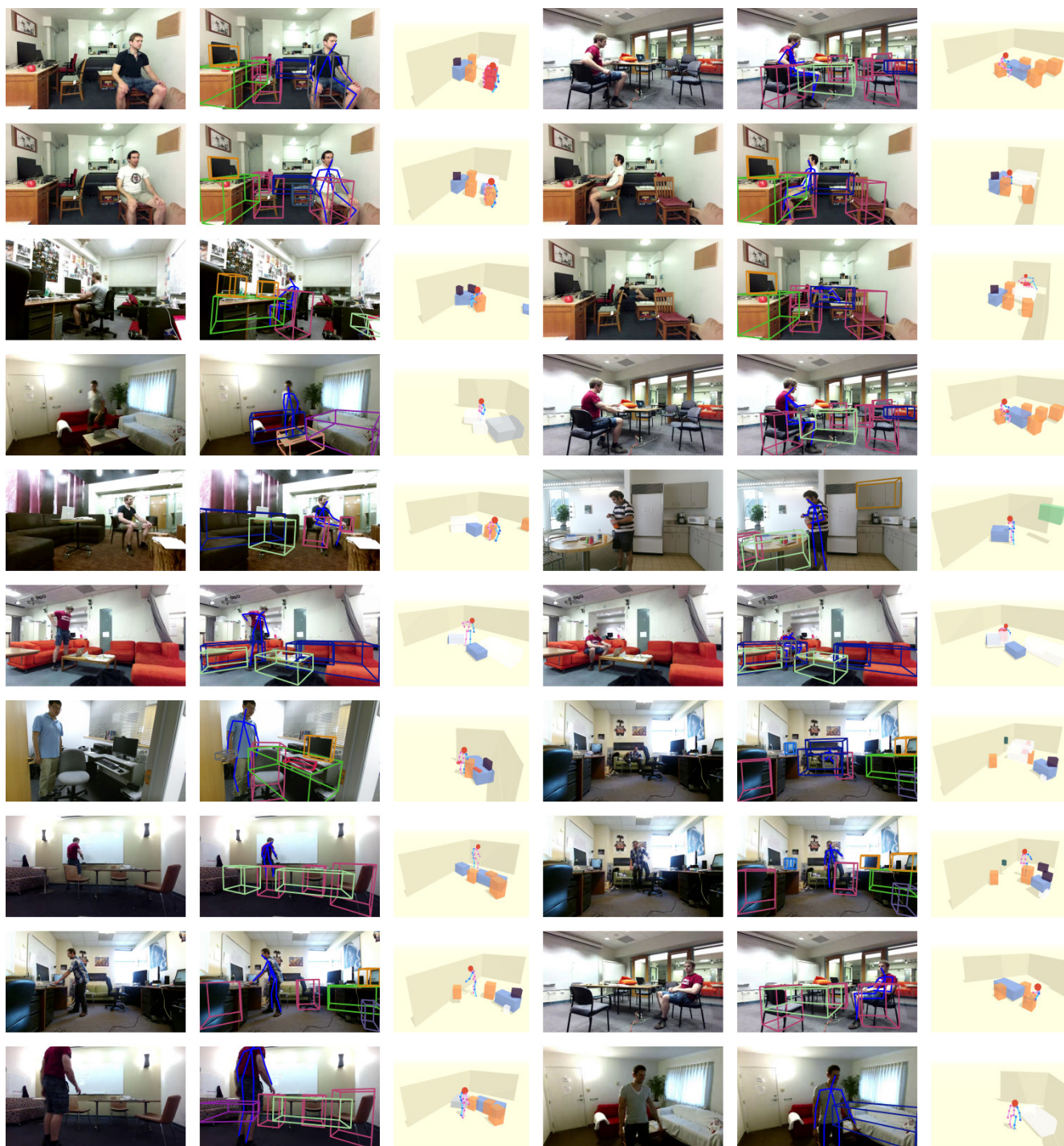


Figure 95. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

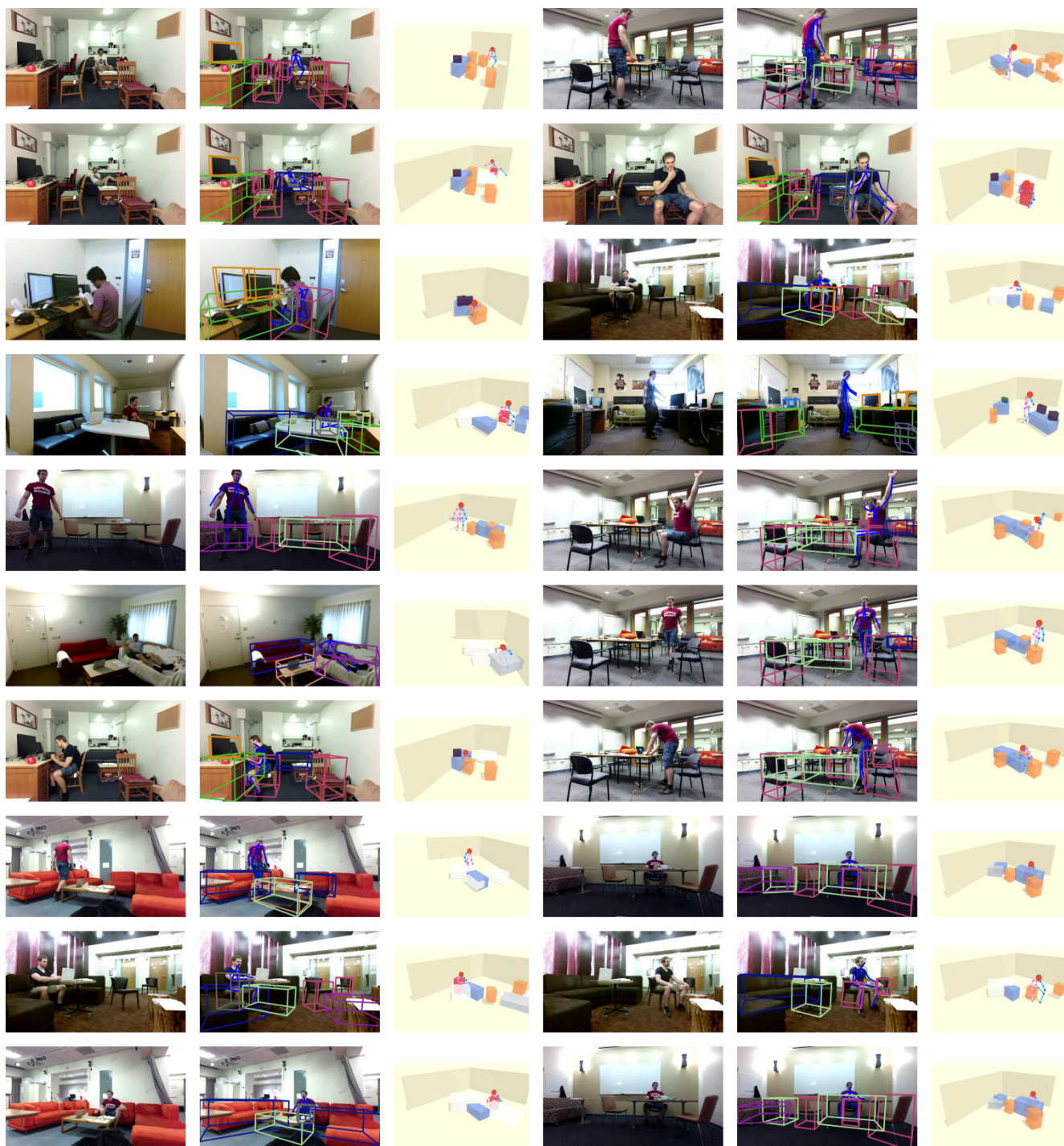


Figure 96. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

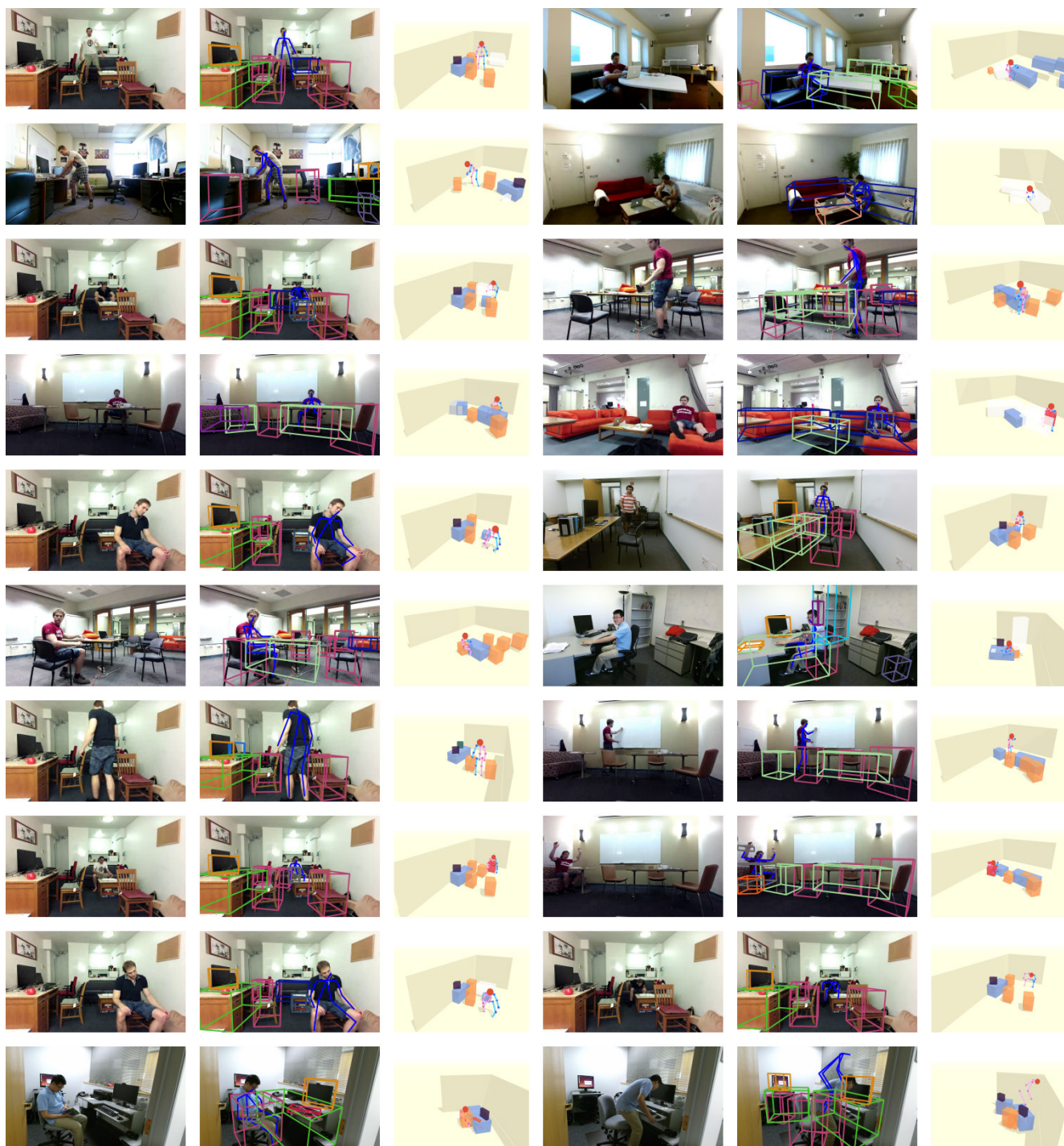


Figure 97. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

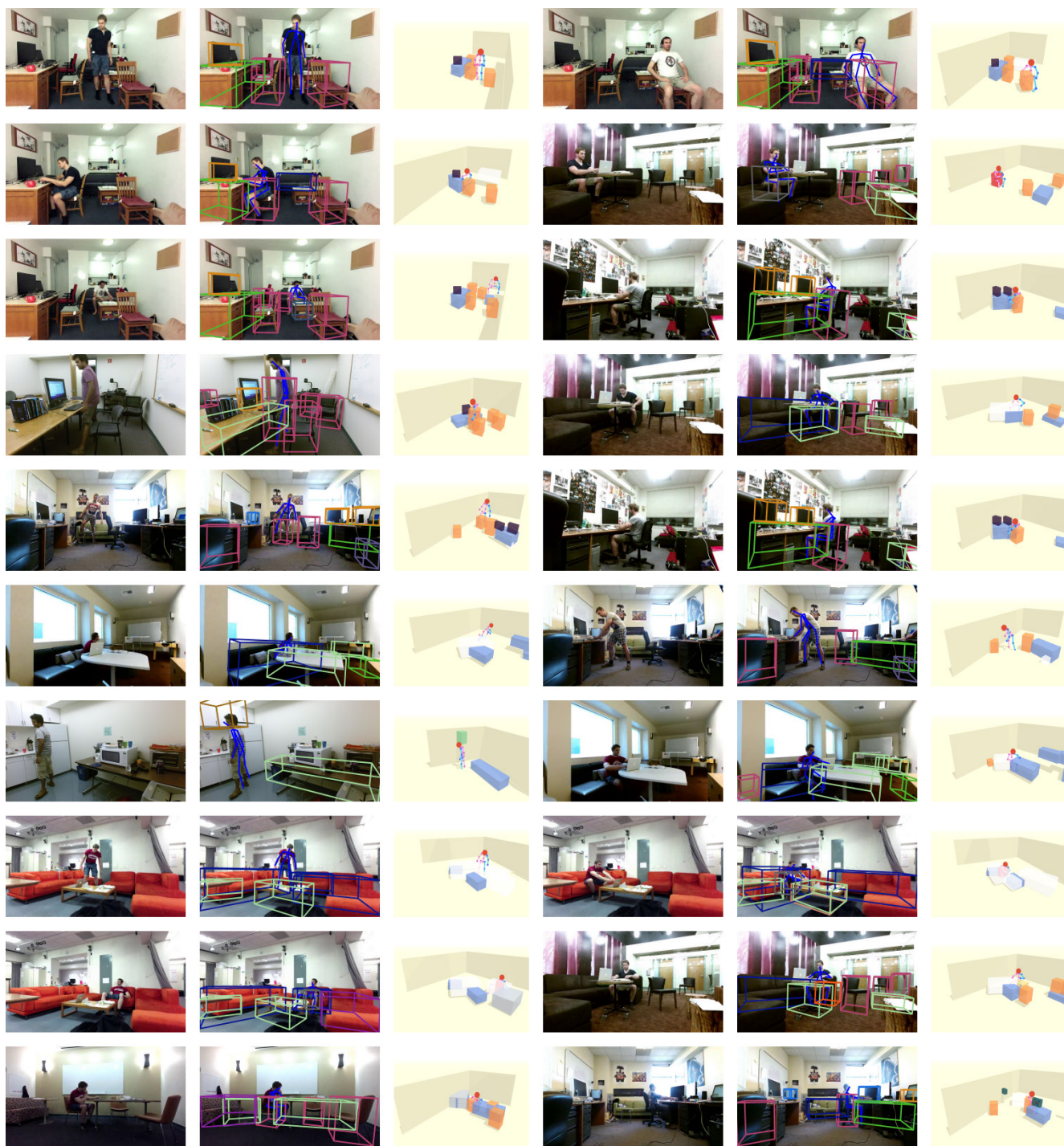


Figure 98. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

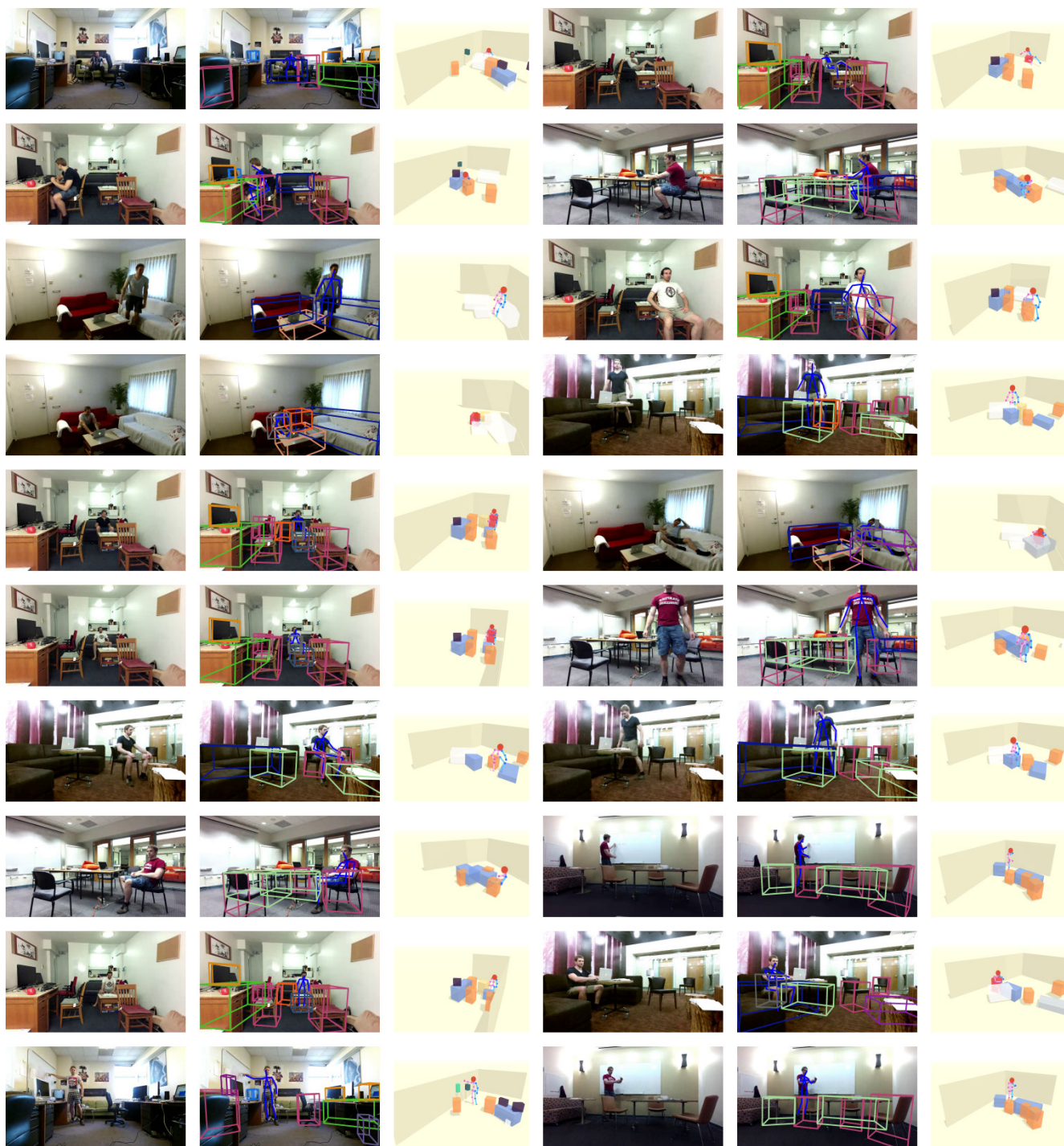


Figure 99. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

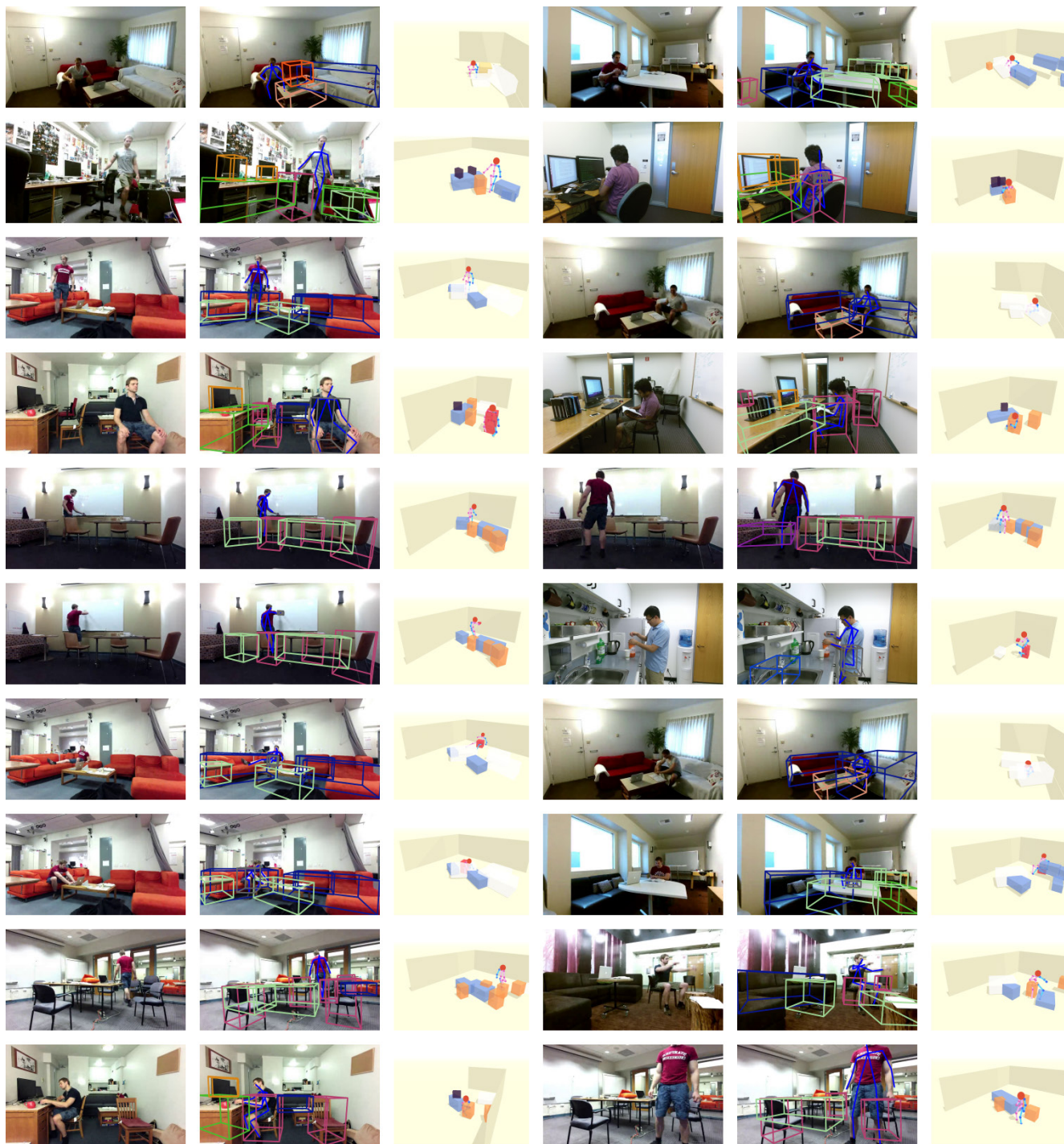


Figure 100. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.

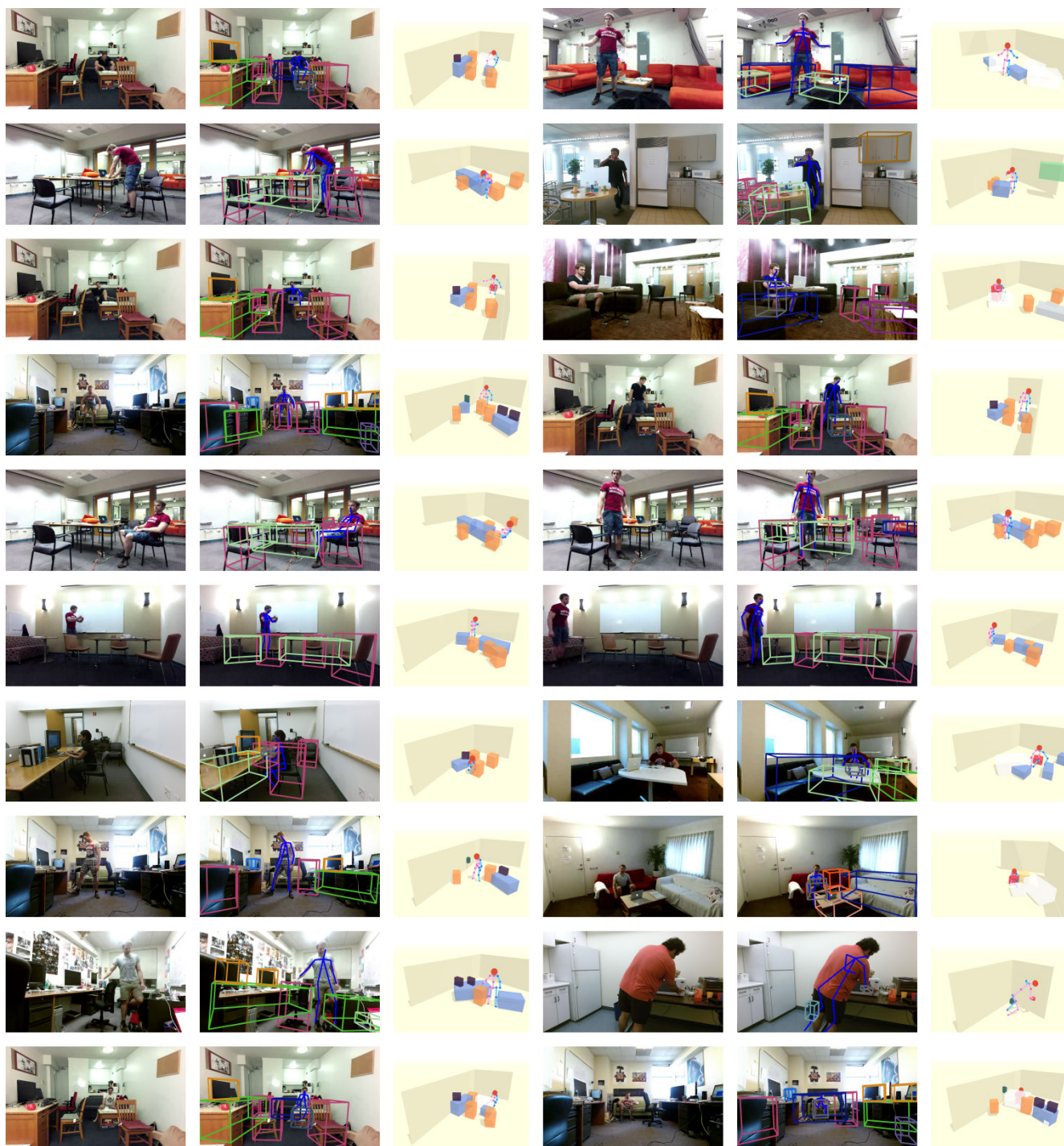


Figure 101. Qualitative results of the proposed method on Watch-n-Patch and PiGraphs dataset.