Surface prediction for a single image of urban scenes

Foat Akhmadeev

Kazan Federal University

Introduction

- Our goal
  - Reconstruct a 3D model of a scene from a single image of urban scenes.

Why we need that?

- Applications
  - Robot navigation.
  - Improvement of object detection and scene understanding.
  - Rendering synthetic objects into photos.
  - Image segmentation improvement.

- Basic assumption
  - An image should be in agreement with the "Manhattan" worlds assumption.
- Novelty
  - The algorithm recovers objects from different levels of a scene.
  - Works on both indoor and outdoor man-made scenes.
  - Combinations with other methods outperform several state-of-the-art approaches.
- Experimental approach
  - We test the orientation of surfaces on standard datasets.

Method

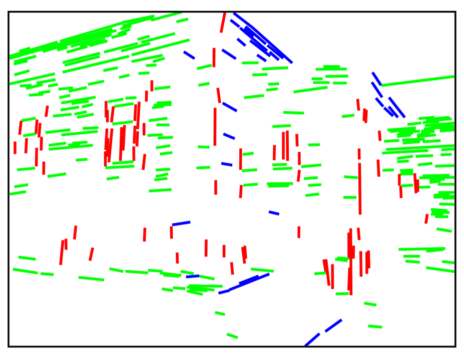

Fig. 2. Lines extracted from the image. Each group of lines form a vanishing point (*vp*).

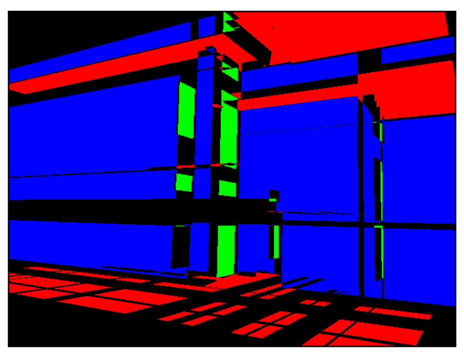

Fig. 3. Initial orientation map from [4]. Orientation map represents orientations of pixels on an image.

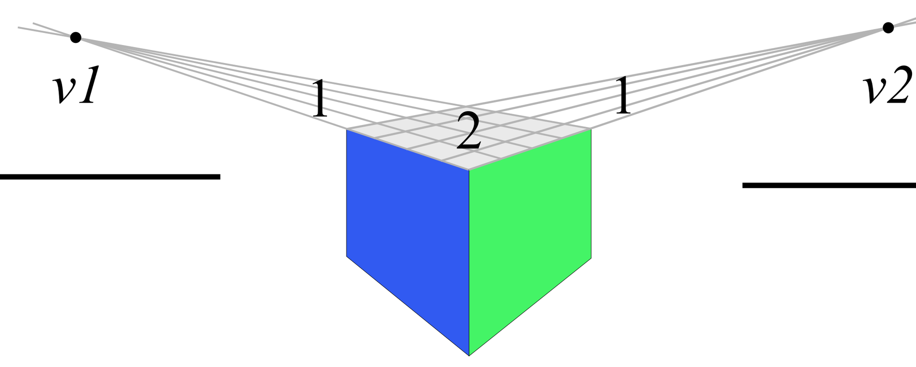

Fig. 4. Apply surface prediction approach. Each part of a plane boundary votes for neighbor surfaces using vanishing points. After voting we choose the best matching for surfaces orientation. Numbers indicate votes for surfaces. *v1* and *v2* are vanishing points.

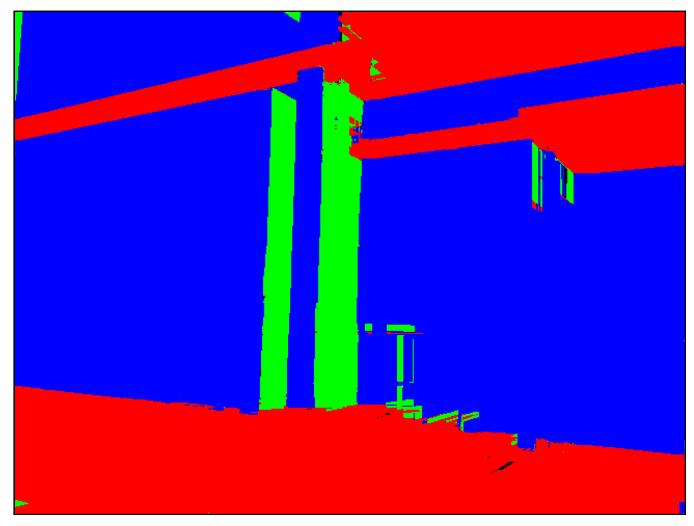

Fig. 5. Orientation map after surface prediction.

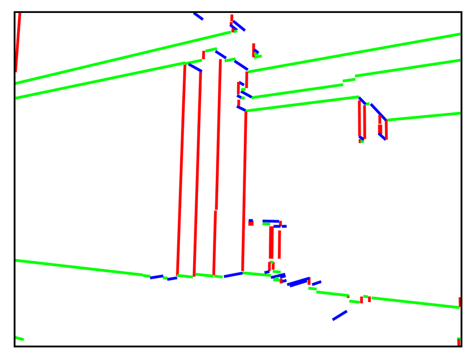

Fig. 6. Clean obtained orientation map with lines which belong to *vp*.

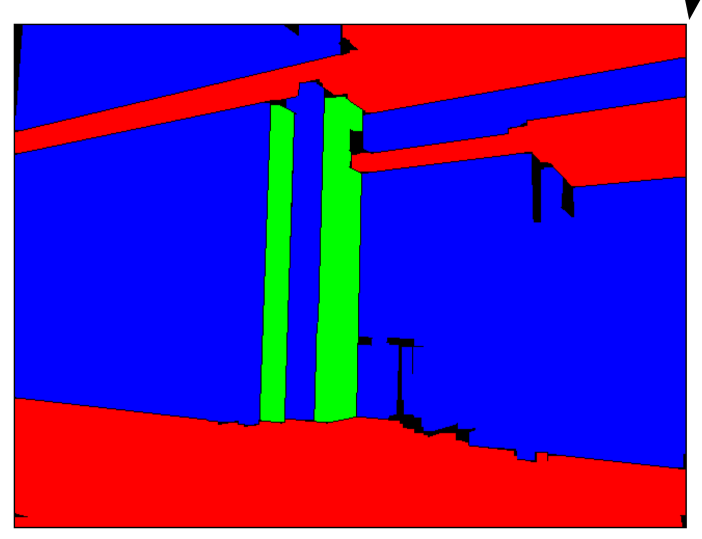

Fig. 7. Final orientation map.

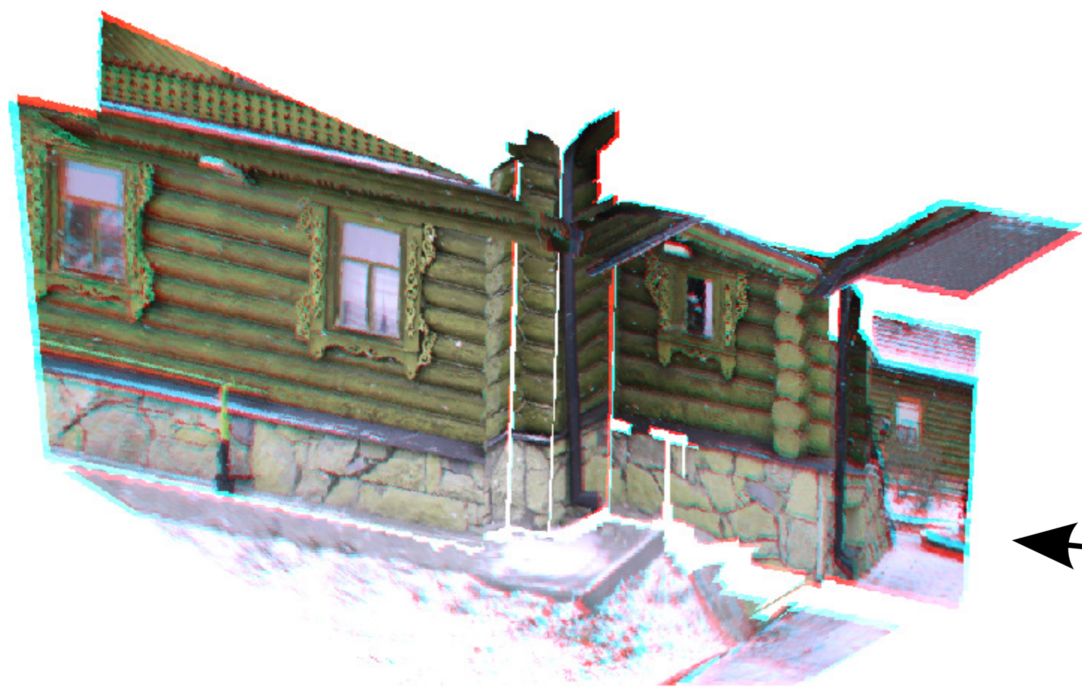

Fig. 8. Output 3D model (use anaglyph glasses). We build it iteratively, thus, algorithm recovers surfaces from different levels of a scene.

Fig. 1. Input image, can be either indoor or outdoor.

"Manhattan" worlds assumption states that most of horizontal lines in space are divided into two orthogonal groups.
Vanishing point — single point on the image plane which is created by intersection of projections of parallel lines in space on an image.

Results

- We used Delage et al. [1] (48 images of indoor scenes) and York Urban [2] (102 images of indoor and outdoor scenes) datasets.
- We tested the average percentage of pixels with the correct orientation.
- The best result is the combination of our approach with *geometric context* [3].
- It outperforms Lee et al. [4] approach for 8.9% and *geometric context* [3] for 1.9%.

Contacts
foat.me foat.akhmadeev@gmail.com

Conclusion

We introduced a new method for 3D surface reconstruction. By evaluation on popular image datasets, it was shown that it works on both urban scenes and indoor apartments. Next, we addressed the problem of restoring objects that are not connected to only one base plane (ground or floor). In addition, due to the voting process this approach is easily combined with other algorithms and obtained results outperformed several state-of-the-art techniques.

Acknowledgments to Jiri Matas for valuable comments on the paper and to Evgeny Stolov for the help during the research.

References

1. Delage et al. *CVPR* **2**. 2006, 2418
2. Denis et al. *ECCV*. 2008, 197
3. Hoiem et al. *IJCV* **75**. 2007, 151
4. Lee et al. *CVPR, IEEE*. 2009, 2136