Physics Inspired Optimization on Semantic Transfer Features: An Alternative Method for Room Layout Estimation

Wenjie Niu

June 4,2018

Abstract

In this paper, we propose an alternative method to estimate room layouts of cluttered indoor scenes. This method enjoys the benefits of two novel techniques. The first one is semantic transfer (ST), which is: (1) a formulation to integrate the relationship between scene clutter and room layout into convolutional neural networks; (2) an architecture that can be end-to-end trained; (3) a practical strategy to initialize weights for very deep networks under unbalanced training data distribution. ST allows us to extract highly robust features under various circumstances, and in order to address the computation redundance hidden in these features we develop a principled and efficient inference scheme named physics inspired optimization (PIO). PIOs basic idea is to formulate some phenomena observed in ST features into mechanics concepts. Evaluations on public datasets LSUN and Hedau show that the proposed method is more accurate than state-of-the-art methods. [10]

1. Introduction

Given an input RGB image, a room layout estimation algorithm should output all the wall-floor, wall-wall, and wall-ceiling edges (depicted by Fig 1). This is a fundamental indoor scene understanding task as it can provide a strong prior for other tasks like depth recovery from a single RGB image [3] [2] or indoor object pose estimation [8] [4] [9]. Besides, the room layout itself provides a high-level representation of an indoor scene for emerging applications like intelligent robots and augmented reality. This problem draws constant attention since the publication of the seminal work [5], and there are two lines of followers: (1) As the upper part of Fig 1 shows, conventional methods follow a proposing-ranking scheme. Typically, the proposing part consists of three sub-modules as edge detection, vanishing point voting and ray sampling. With hand-crafted features and structured inference techniques, the ranking part outputs the best layout proposal, sometimes along with a representation of the clutter. (2) Recent methods [6] [1] [7] achieve dramatic performance improvements via features produced by fully convolutional networks (FCNs). [6] [7] still follow the traditional proposing-ranking scheme. [1] is a proposal-free solution in which all those steps about proposal generation are eliminated. And instead of proposal ranking, in [1] inference is achieved through an optimization module.

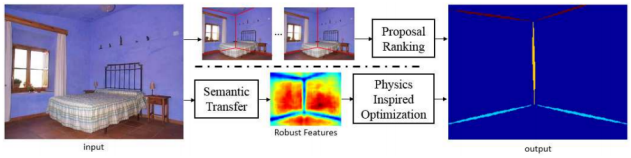

Figure 1. Above is the overview of conventional methods. Below is the overview of our method. Better viewed electronically.

References

[1] S. Dasgupta, K. Fang, K. Chen, and S. Savarese. Delay: Robust spatial layout estimation for cluttered indoor scenes. In *CVPR*, 2016. 1

[2] D. Eigen and R. Fergus. Predicting depth, surface normals and semantic labels with a common multi-scale convolutional architecture. In *CVPR*, 2014. 1

[3] D. Eigen, C. Puhrsch, and R. Fergus. Depth map prediction from a single image using a multi-scale deep network. In *NIPS*, 2014. 1

[4] S. Gupta, P. Arbelaez, R. Girshick, and J. Malik. Aligning 3D models to RGB-D images of cluttered scenes. In *CVPR*, 2015. 1

[5] V. Hedau, D. Hoiem, and D. Forsyth. Recovering the spatial layout of cluttered rooms. In *CVPR*, 2010. 1

[6] A. Mallya and S. Lazebnik. Learning informative edge maps for indoor scene layout prediction. In *ICCV*, 2015. 1

[7] Y. Ren, S. Li, C. Chen, and C. C. J. Kuo. A coarse-to-fine indoor layout estimation (CFILE) method. In *ACCV*, 2016. 1

[8] S. Song and J. Xiao. Sliding shapes for 3D object detection in depth images. In *ECCV*, 2014. 1

[9] S. Song and J. Xiao. Deep sliding shapes for amodal 3D object detection in RGB-D images. In *CVPR*, 2016. 1

[10] H. Zhao, M. Lu, A. Yao, Y. Guo, Y. Chen, and L. Zhang. Physics inspired optimization on semantic transfer features: An alternative method for room layout estimation. In *CVPR*, 2017. 1