Graph Cut Based approach in creating superpixel for Depth estimation

Vaibhav B Waman[1],Dr.Rahul M Mulajkar[2]

Department Of Electronics and Telecommunication,
Savitribai phule Pune University,Pune
Jaihind college of Engineering,kuran Pune

Abstract

The Target of estimating the depth of each pixel in a scene from a single image which is a monocular image using the monocular cues. In past research work approaches [1, 2], which attempt to map from appearance features to depth directly, we first perform a generate segmentation (creation of Superpixels) of the scene and use the semantic labels to guide the 3D reconstruction. This approach provides several advantages: By knowing the segmentation class of a pixel or region, depth and geometry constraints can be easily enforced (which gives idea about "sky" is far away and "ground" location of scene). In addition, depth can be more readily predicted by measuring the difference in appearance with respect to a given segmentation class. For example, a tree will have more uniform appearance in the distance than it does close up.by means of which, the incorporation of segmentation features allows us to achieve state-of-the-art results with a significantly simple model than previous research works.various models have been used to convert the monocular image into the 3d reconstruction model.

Keywords-*Superpixels,Graphcut,depth Estimation,3d Reconstruction,segment-ation*

I. INTRODUCTION

Superpixels (image segments) can provide powerful approach to grouping cues to guide segmentation, where superpixels can be collected easily by (over)-segmenting the image using any reasonable existing segmentation algorithms. Image segmentation is a fundamental low-level vision problem with a great potential in applications. While human can parse an image into coherent regions easily, it is found rather difficult for automatic vision systems. Despite a variety of segmentation techniques have been proposed, it remains challenging for any single method to do segmentation successfully due to the broad diversity and ambiguity of visual patterns in a natural image.

II. RELATED WORK

In previous research work, Michels, Saxena and Ng [4] used supervised learning for estimation of 1-D distances to obstacles for the application of automatic driving of a remote control car. Nagai et al. [5] performed surface reconstruction from single images on known fixed objects such as hands and faces.Using single camera vision, Gini & Marchi [6] drove an indoor robot, which relied heavily on known ground colors and textures. Shape from shading [7] offers another method for monocular depth reconstruction, but is difficult to apply to scenes that do not have fairly uniform color and texture. Hoiem, Efros and Herbert (personal communication) also considered monocular 3-D reconstruction, but focused on generating 3-D graphical images rather than accurate metric depthmaps. In this paper, we address the task of learning full depthmaps from single images of unconstrained environments.

III. MONOCULAR CUES

Humans finds easy to measure depth from single monocular images. [12] This is done using monocular cues such as texture variations, texture gradients, occlusion, known object sizes, haze, defocus, etc. [4, 13, 14] For example, many objects texture will look different at different distances from the viewer. Texture gradients captures the distribution of the direction of edges and helps to indicate depth. Haze is another depth cue which is caused by atmospheric light scattering. Most of these monocular cues are 'contextual information' in the sense that they are global properties of an image and cannot be inferred from small image patches. For example, occlusion cannot be determined if we look at a small portion of an occluded object. Local information like texture and color of a patch can give some information about its depth which is usually not enough to determine its absolute depth accurately. For more clarification, if we look at a patch of a clear blue sky, it is not easy to tell either patch is infinitely far away (sky) or it is part of a blue object. To avoid ambiguities, one should look at the *overall* organization of the image to determine depths.

IV. GRAPH CUT SEGMENTATION

Process of the Graph cut partitioning a directed or undirected graph into disjoint sets. The concept of optimizing of such cuts is usually introduced by associating energy to each cut. This kind of issues have been studied within the field of graph theory but it works for graphs with more than only a few nodes. Nevertheless, ever since it became apparent that many low-level vision problems can be posed as finding energy minimizing cuts in graphs. In computer vision community, these techniques have received a lot of attention. For successful application on stereo, image restoration, texture synthesis and image segmentation, graph cut methods have been very much useful. In the figure below, we have given a brief overview of graph cut methods for image segmentation and introduction to some basic definitions as well.

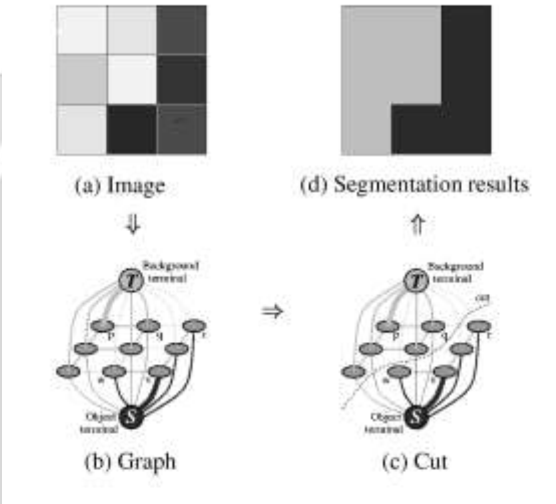

Fig 1 is an example of segmentation of a very simple 3-by-3 image. Edge thickness corresponds to the associated edge weight.

The min-cut of the resulting graph will be the segmentation of the image. This segmentation should be a partition owing to the definition of image-pixel resemblance and similar pixels close to each other will belong to the same partition. Also due to the result of the terminal weights, pixels should also be segmented in such a manner that they end up in the similar partition like the terminal node corresponding to the model (object or background).

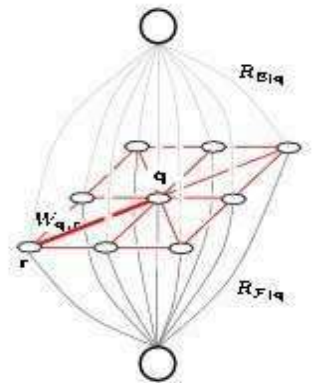

Fig 2.Graph Cut Representation

RESULT

We have tested this model on real world images of forests (containing trees, bushes, etc.), campus areas (buildings, people, corridors).

Fig 3.Superpixel output of sample image

Fig.4 Graph cut superpixel output

ACKNOWLEDGMENT

With deep sense of gratitude, we sincerely acknowledge all who gave us the idea, guidance, important suggestions & encouragement during the progress of our project and without this completion of this project would have been difficult task.

DISSCUSSION

The Target of depth perception is Important to computer vision, one that has enjoyed the attention of many researchers and seen significant progress in the past few decades. However, the vast majority of this work, such as stereopsis, used multiple image geometric cues to infer depth. In contrast, single-image cues offer a largely orthogonal source of information, one that has heretofore been relatively underexploited.also at different stages of depth estimation like segmentation or creating superpixel the accurate method is much important for better results.

REFERENCES

[1] D. Scharstein and R. Szeliski. A taxonomy and evaluation of dense two-frame stereo correspondence algorithms. *Int'l Journal of Computer Vision*, 47:7–42, 2002.

[2] David A. Forsyth and Jean Ponce. *Computer Vision : A Modern Approach*. Prentice Hall, 2003.

[3] S. Das and N. Ahuja. Performance analysis of stereo, vergence, and focus as depth cues foractive vision. *IEEE Trans Pattern Analysis & Machine Intelligence*, 17:1213–1219, 1995.

[4] J. Michels, A. Saxena, and A.Y. Ng. High speed obstacle avoidance using monocular vision and reinforcement learning. In *ICML*, 2005.

[5] T. Nagai, T. Naruse, M. Ikehara, and A. Kurematsu. Hmm-based surface reconstruction from single images. In *Proc IEEE Int'l Conf Image Processing*, volume 2, 2002.

[6] G. Gini and A. Marchi. Indoor robot navigation with single camera vision. In *PRIS*, 2002.

[7] M. Shao, T. Simchony, and R. Chellappa. New algorithms from reconstruction of a 3-d depth map from one or more images. In *Proc IEEE CVPR*, 1988.

[8] J. Lafferty, A. McCallum, and F. Pereira. Discriminative fields for modeling spatial dependencies in natural images. In *ICML*, 2001.

[9] K. Murphy, A. Torralba, and W.T. Freeman. Using the forest to see the trees: A graphical model relating features, objects, and scenes. In *NIPS 16*, 2003.

[10] Xuming He, Richard S. Zemel, and Miguel A. Carreira-Perpinan. Multiscale conditional random fields for image labeling. In *proc. CVPR*, 2004.

[11] S. Kumar and M. Hebert. Discri- minative fields for modeling spatial dependencies in natural images. In *NIPS 16*, 2003.

[12] J.M. Loomis. Looking down is looking up. *Nature News and Views*, 414:155–156, 2001.

[13] B. Wu, T.L. Ooi, and Z.J. He. Percei- ving distance accurately by a directional process of integrating ground information. *Letters to Nature*, 428:73–77, 2004.

[14] P. Sinha I. Blthoff, H. Blthoff. Top-down influences on stereoscopic depth-perception. *Nature Neuroscience*, 1:254–257, 1998.