ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

Point Cloud Based 3D Reconstruction from Single Image

Sagi Venkata Priyanka, Prof. T. Ramashri, P.Janardhan Sai Kumar

M.Tech Student, Dept. of ECE, S.V.U. College of Engineering, Tirupati, Andhra Pradesh, India

Professor, Dept. of ECE, S.V.U. College of Engineering, Tirupati, Andhra Pradesh, India

Research Scholar, Dept. of ECE, S.V.U. College of Engineering, Tirupati, Andhra Pradesh, India

ABSTRACT: Holistic scene understanding is a major goal in recent research of computer vision. To deal with this task, reasoning the 3D relationship of components in a scene is identified as one of the key problems. The problem of estimating detailed 3-d structure from a single image of an unstructured environment is considered. The goal is to create 3-d models which are both quantitatively accurate as well as visually pleasing. Consequently, scene categorization is idetified as the first step towards robust and efficient depth estimation from single images. Stage information serves as the first approximation of global depth, narrowing down the search space in depth estimation and object localization. Classification results demonstrate that stages can be efficiently learned from low-dimensional image representations. An extension method which gives better results based on point cloud concept is proposed. Here, the first and second images of a single image are considered and 3d layout is constructed. Experiments show that the results of our model outperform the state-of-the-art methods for 3D structure classification.

KEYWORDS: 3d structure, stage classification.

I. INTRODUCTION

Visual perception is the process of inferring world structure from image structure. Although the projection of the physical 3D scene onto the image plane carries ambiguities as to which physical configuration gave rise to the depiction, human observers can effortlessly derive an impression of scene depth from a single image. This is because the world around us behaves regularly, and because structural regularities are directly reflected in the 2D image of some world scene. The aim of this paper is to identify these regularities in 2D images and exploit them for the purpose of scene geometry reconstruction of accurate 3d image. The notion of a stage from the observation that objects of the world act in relatively stable, recurring geometrical environments is derived.

The objects come with almost infinite variation in appearance, as well as geometry. Scenes, on the other hand, show a much more regular pattern. One is quickly led to conclude that a vast majority of images depicts the scene with only a few different geometry types if one considers only a subset of given constraints. Namely, straight image lines always converge at vanishing points and the horizon; walls are almost always perpendicular to the ground surface; the camera is almost always at approximately two meters height, etc.

The important conclusion is that whereas a precise geometry may be requested for the object, for many applications it suffices to build a rough model for the geometry of the scene. In this paper, the aim is to discover the stages as models of scene geometry derived from a single image.

There are many advantages of knowing just the stage type. Stages represent the general 3D geometry structure of scenes such as sky-background-ground,

Sky-ground, and box,. Apart from identification of the scene configuration, the stage may reveal to the observer, the information about the semantic context of the scene, the identities of scene elements, as well as relative depth order. Recognition of the stage thus serves as the entry point of a detailed depth analysis, object localization and object recognition.

The rest of this paper is organized as follows. The existing method is described in detail. The proposed method is illustrated later. Then experimental results are reported to demonstrate the superior performance of the framework. Finally, conclusions are presented.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

II. RELATED WORK

The existing method consists of mainly three main contributions: Segmentation is exploited to obtain robust stage classification. An algorithm is proposed using stage classification to generate 3D layouts. No need for pixel-wise labeled training data and applicable to generic images. A step by step algorithm for the existing method is illustrated below.

Stage classification:

1. For stage classification, a graphical model is proposed to learn a mapping from image features to a stage.
2. The scenes are parsed by a fixed set of pre defined templates. These templates capture the basic geometry of stages, such as sky-background-ground and box. Here, 'component' is used to denote a sub-region of a single template. For example, certain Template contains three components: the upper, middle and bottom part. Instead of extracting features from each component of these templates directly, we apply segmentation based on the image templates. Then the image features are extracted from the segment of each component of the templates.

Template based feature extraction:
Template based soft segmentation:
Soft segmentation is based on the figure-ground segmentation method. Foreground seeds are uniformly placed on a grid and the background seeds are placed on the borders of the image. Foreground segments are generated from each seed at different scales by solving the max-flow energy optimization. Next, we keep the obtained segments which are highly intra-consistent (i.e. segments generated several times by different seeds).
In this way, hundreds of segments are generated for each image. For each component of the templates, the segment which has the largest overlap-to-union score is selected.

Template based hard segmentation:
Since the soft segmentation is not always reliable, hard segmentation is used as compensation. Hard segmentation is predefined by the template. For example,,hard segmentation parses the image uniformly into three sub-regions: top, middle and bottom.

Feature representation:
The two segmentation methods are used to generate two segments at each component of the templates. Features (e.g. HOG) are extracted from both of the segments, and these features are concatenated into a single feature vector.

Mapping:
Mapping of input features to a stage are done by Model formulation, Objective function, Inference and Learning.

Random walk segmentation:
Random walk segmentation method provides a natural way of using the stage prior which can be exploited in two ways: seeds and edges. Seeds are placed in the image regions corresponding to the geometric planes, i.e. the seeds of the sky in the class *sky-background-ground* are placed at the top of the image while the seeds of ground are placed at the bottom.
In this way, the global structure obtained by the stage classification is used as a prior. Segmentation is generated based on the new weight map and seeds.
In this structure, the ground truth of the pixel-level design is required.
Also, since the mapping capacity of the image to 3D format is found out from neighborhood sections, abnormal state setting of the geometry is difficult to misuse. For instance, for the class box (e.g. hallway), it is difficult to separate the geometry by watching a solitary section. In any case, in our structure, the scene class is initially predicated by considering the entirety image and the interdependency of the sub regions.
At that point, the image level 3D format is gathered by considering this earlier data. Along these lines, we consolidate base up division with the worldwide connection of the image. At the point, when the stage order gives poor results, the calculation sections the picture utilizing most likely false earlier data.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

For this situation, the format may not be obvious. Really, the blunder relies on upon the degree of perplexity between the two classes. For instance, if sky-ground is anticipated into foundation ground, then the calculation can even now effectively section the picture into two sections.

Be that as it may, if sky-ground is anticipated as a container, then the calculation fragments the picture into five sections also, will bring about an extensive blunder. From the disarray grid, it is demonstrated that the most befuddled stages are foundation ground what more, sky-foundation ground is.

Note that the ground-truth is only used for evaluating the 3D layout algorithm. Our approach only needs

Image-level label ground-truth for training. When this happens, the algorithm segments the background-ground images into three parts. The top part of the buildings is assigned to be sky this will result in error.

III. PROPOSED ALGORITHM

In this work, a novel method is proposed. Surface growing based segmentation approach is implemented on the basis of bottom-up strategy. It means that procedure starts from the some selective seed points and the segments are grown based on pre-defined similarity criterion. The selection of the seed points is one of the important issues in this process that can affect directly on the final results.

Segmentation is carried out in both object space and image space to extract building roofs and streets. Firstly, it is applied in the object space based on surface growing-based segmentation by using surface normal that are computed from X-Y-Z image for each valid image pixel. In object space, adjacent point clouds are grouped based on similarity measures like direction of locally estimated surface normal and distances of the point clouds to the best fitted plane.

Afterwards, segmentation proceeds in image space by the usage of segmented regions from the object space and grouping adjacent image pixels based on similarity of gray or color value of neighboring image pixels. Therefore, final segmentation results in extraction of building roofs and streets that extracted regions significantly differ from each other due to having different surface slopes or being in different elevations. A point cloud is a set of data points in some coordinate system.

In a three-dimensional coordinate system, these points are usually defined by *X*, *Y*, and *Z* coordinates, and often are intended to represent the external surface of an object. Point clouds may be created by 3D scanners. These devices measure a large number of points on an object's surface and often output a point cloud as a data file. The point cloud represents the set of points that the device has measured.

The 3D data sets of this work were provided from pixel-wise image matching. Firstly, parallax is generated as a result of image matching and consequently 3D coordinates are produced for each image pixel. Parallax is represented by so-called disparity image and 3D coordinate is stored in X-Y-Z image. Thus, corresponding X-Y-Z image is adapted to range image segmentation problem3D point clouds are transformed into the raster domain to be adapted for range data segmentation problem that causes loss of information. However, generated X-Y-Z image provides us 3D point cloud for each pixel that is 3D structured data and no need of the transformation of 3D data from 3D space. Moreover, additional information like gray or color value from gray image or RGB image can be utilized to interpret each pixel as well.

The feature extraction problem is closely related to surface reconstruction, which has important applications in laser range scanning, scientific, computing, computer vision, medical imaging, and computer assisted surgical planning. Our algorithm addresses some of the problems that arise in surface reconstruction of datasets that contain creases and corners.

For these approaches it is very helpful to extract the feature lines and junctions beforehand.

Furthermore, valid or invalid image pixels are the result of matching process and 3D coordinates are merely available for valid image pixels. Hence, validity of 3D coordinate for each X-Y-Z image pixel can be checked by another image that is so-called valid image.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

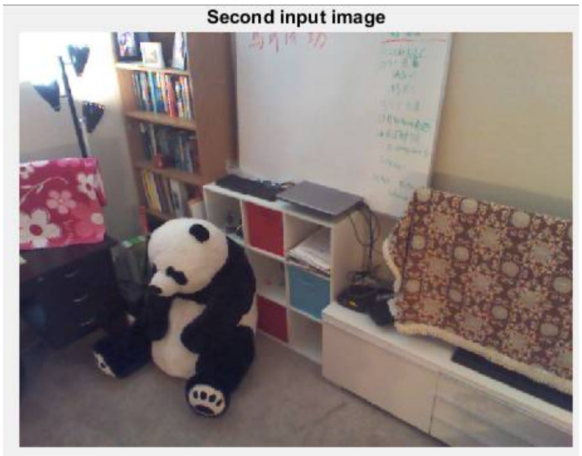
ALGORITHM:

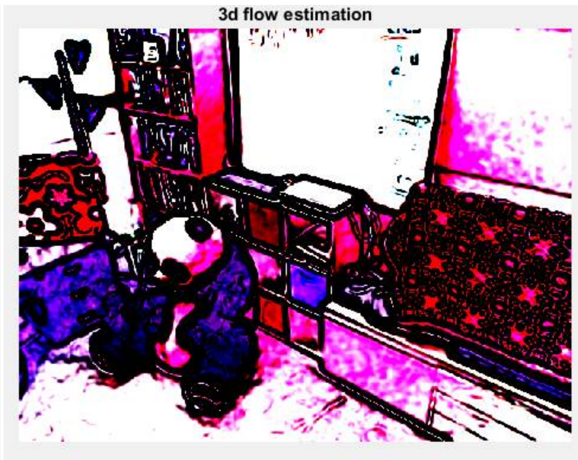
The following steps give the information of the proposed method:

1. The first and second images are taken from the original image.
2. Then strongest features are taken from the image.
3. Then the features are tracked.
4. Segmentation process is done based on features.
5. Then for each data point its 3d point is extracted.
6 .Then all the 3d data points are pooled together to form the 3d image.
7. Finally the 3d layout is extracted from the 2d image.

By using this method, the 3d layout is extracted from the 2d image very accurate than the existing methods this is very useful in scene understanding, construction of buildings etc. This gives better results compared to the existing methods

IV. SIMULATION RESULTS





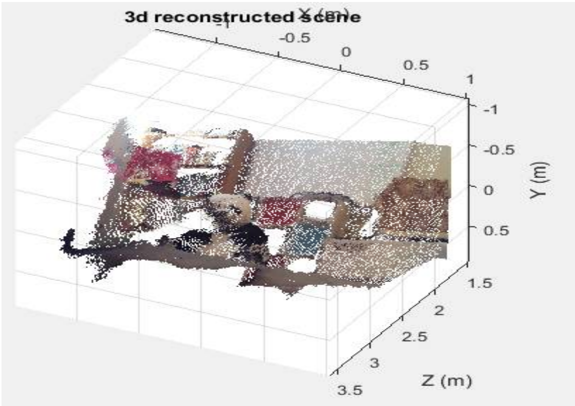
 DOI: 10.15680/IJIRCCE.2016. 0411222

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016



Parameters:
Accuracy: Total number of correctly predicted samples divided by the total number of test samples.
Precision: The precision Pr(i) of the ith category is defined as the number of correct predictions of the ith category divided by the total number of samples which are predicted to be in the ith category.
Recall: The recall Re(i) of the ith category is defined as the number of correct predictions for the ith category divided by the total number of the ith category in the test data.
F-Score: F-score F(i) is defined as the harmonic mean of precision and recall.
$F(i)= 2*Pr(i)*Re(i) / (Pr(i)+Re(i))$

Parameters	Existing method	Proposed method
Accuracy	8.32	68.47
Precision	8.8327	43.0817
Recall	58.98	76.38
F-score	16.39	63.79

V. CONCLUSION AND FUTURE WORK

In this paper, a probabilistic model for inferring the rough geometric structure of a single image is presented. A novel method is used to segment the image based on point cloud. Furthermore, the method exploits an image-level 3D structure prior to infer a pixel-wise 3D layout. Experimental results show that our approach outperforms state-of-the-art methods in the context of stage classification. By using the stage as prior information, it is shown that pixel-wise 3D layout can be generated in a proper way. Finally, we are able to obtain the pixel-level 3D layout from a single 2Dimage without pixel-wise ground truth label for training.
The existing 3D scene layout methods, which mainly focus on indoor scenes, are limited in outdoor applications. In the future work is performed to develop improved manifold sorting algorithm in sketch retrieval method, which can get the 3D models rapidly. Secondly, according to particular properties of outdoor architectures, specialized energy constraints are used, which define the energy function that meets the functional and aesthetic needs. Thirdly, the scene achieves automatic layout by adopting simulated annealing algorithm.

REFERENCES

[1] W. Richards, A. Jepson, and J. Feldman, “Priors, preferences and categorical percepts,” in Perception as Bayesian Inference, D. Knill and W. Richards, Eds. Cambridge University Press, 1996, ch. 3, pp. 80–111.
[2] B. Horn and M. Brooks, Shape from Shading. MIT Press, 1989.
[3] L. Matthies, T. Kanade, and R. Szeliski, “Kalman filter-based algorithms for estimating depth from image sequences,” IJCV, vol. 3(3), pp. 209– 238, 1989.
[4] S. Barnard, “A stochastic approach to stereo vision,” in 5th National Conference on Artificial Intelligence, 1986.
[5] D. Hoiem, A. A. Efros, and M. Hebert, “Recovering surface layout from an image,” IJCV, vol. 75(1), pp. 151–172, 2007.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 4, Issue 11, November 2016

[6] D. Hoiem, A. Stein, A. A. Efros, and M. Hebert, “Recovering occlusion boundaries from a single image,” ICCV, 2007.
[7] A. Torralba and A. Oliva, “Depth estimation from image structure,” IEEE PAMI, vol. 24(9), pp. 1226–1238, 2002
[8] V. Nedovic, A. W. M. Smeulders, A. Redert, and J. M. Geusebroek, ´ “Depth information by stage classification,” ICCV, 2007.
[9] V. Hedau, D. Hoiem, and D. Forsyth, “Thinking inside the box: Using appearance models and context based on room geometry,” in *Proc. 11thECCV*, 2010, pp. 224–237.
[8] V. Nedovic, A. W. M. Smeulders, A. Redert, and J. M. Geusebroek, ´ “Depth information by stage classification,” ICCV, 2007.
[9] V. Hedau, D. Hoiem, and D. Forsyth, “Thinking inside the box: Using appearance models and context based on room geometry,” in *Proc. 11thECCV*, 2010, pp. 224–237.
[10] A. Saxena, M. Sun, and A. Y. Ng, “Make3D: Learning 3D scene structure from a single still image,” *IEEE Trans. Pattern Anal. Mach.Intell.*, vol. 31, no. 5, pp. 824–840, May 2008.
[11] D. Hoiem, A. A. Efros, and M. Hebert, “Recovering surface layout from an image,” *Int. J. Comput. Vis.*, vol. 75, no. 1, pp. 151–172, 2007.
[12] D. C. Lee, M. Hebert, and T. Kanade, “Geometric reasoning for single image structure recovery,” in *Proc. IEEE Conf. CVPR*, Jun. 2009,pp. 2136–2143.
[13] E. Delage, H. Lee, and A. Y. Ng, “A dynamic Bayesian network model for autonomous 3D reconstruction from a single indoor image,” in *Proc.IEEE Comput. Soc. Conf. CVPR*, Jun. 2006, pp. 2418–2428.
[14] O. Barinova, V. Konushin, A. Yakubenko, K. Lee, H. Lim, and A. Konushin, “Fast automatic single-view 3D reconstruction of urban scenes,” in *Proc. 10th Eur. Conf. Comput. Vis. (ECCV)*, Marseille, France, Oct. 2008, pp. 100–113.
[15] S. Lazebnik, C. Schmid, and J. Ponce, “Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories,” in *Proc.IEEEComput. Soc. Conf. CVPR*, Jun. 2006, pp. 2169–2178.