© 2015 IJEDR | Volume 3, Issue 2 | ISSN: 2321-9939

Depth Estimation Using Collection of Models

Karankumar Thakkar[1], Viral Borisagar[2]
PG scholar[1], Assistant Professor[2]
Computer Science & Engineering Department [1, 2]
Government Engineering College, Sector 28, Gandhinagar, India[1, 2]

Abstract **- Today images are easy to acquire, view, publish, and share however they lack critical depth information as the images usually are projected views of a 3-D scene. This makes severe restrictions for many image manipulation, editing, and retrieval tasks. Depth estimation is a technique that aims to retrieve depth information either based on depth cues such as texture, focus and shading or using collection of 3D models. With the recent considerable interest in 3D image analysis, estimating depth information has become a rapidly evolving topic in computer vision research. Also it finds applications in various imaging applications including depth-enhanced image editing, novel view generation etc. Hence, we have strong motivation to consider the problem of adding depth to an image of an object and provide a basis for 3D reconstruction. In this paper, we present an automatic method to find depth information of single image by employing collection of models of the same object class. The key advantage of this method is that even if the dataset does not contain the exact 3D model of the imaged object, it will characterize shape components and generate its depth map. We apply our method on various indoor objects like lamp, chair, cup and car and obtain plausible depth maps.**
Keywords – Depth estimation, Collection of models, Pointcloud, depth map, 3D Reconstruction

I. INTRODUCTION

Images have been the most popular medium to visualize the information. Images contain rich visual detail of a structure and are easy to obtain, view, publish and share. However, they lack critical depth information as they are the projected view of a 3D scene and might occlude important parts of objects and cannot contain depth data. This makes severe restrictions to applications like image recognition, manipulation and depth-enhanced image editing [1]. It would be an advantage if we can recover the lost information while acquiring images. Therefore, we have a strong motivation to find depth information of an object. Estimating depth becomes extremely difficult when we have only one image as input.

Hence, we formulate the problem of estimating the depth information of single image segmented from its background by means of collection of 3D models of the same class. The proposed approach takes a single segmented image and a collection of 3D models of the same image class to find the depth of input image hence it requires prior knowledge of object class. The method is fully automatic and will estimate the depth even if the dataset does not contain exact match of 3D model. The algorithm matches the image with the 3D dataset and extracts similar shapes. After extracting the similar shapes, the algorithm deforms the extracted models using embedded deformation model [8] and generates pointcloud for each shape. Using this pointclouds, the algorithm provides the z-coordinates to our image and generates the final depth map.

II. LITERATURE REVIEW

Hao proposes a method to estimate depth of a single segmented image by incorporating a modest sized dataset of 3d models which aims to identify depth information in three steps preprocessing, reconstruction and optimization [1]. This technique works fine with human made objects but cannot be applied to outdoor objects of high variability (i.e. trees) and high articulation (i.e. animals). Additionally, this method performs the deformation on each model so it takes too much time to process collection of models.

Changhwan proposes another method for depth estimation that estimates depth from a single image of the indoor scene by identifying floor and ceiling parts of image via nonlinear diffusion and image segmentation techniques [2]. A major limitation of this algorithm is that it does not work for man-made objects like chair, table, car, lamp etc. Also, the method is much slower for real time implementation.

Saeed proposes another depth estimation method using in-focused and de-focused images. It uses the focus/blur as a depth cue to find depth information [3]. This method estimates depth based on focus/blur amount, hence the result cannot be as accurate as data-driven methods. Also this method is not suitable for indoor human made objects like chair, lamp, car etc.

Saxena proposes learning based and supervised approach to find depth information from single image [4]. This method fails for estimating depth of indoor man-made objects and the major limitation is that it is not fully automatic as it is supervised and requires instance specific learning.

Akimov proposes depth estimation method using edge blur information. This method calculates the blur amount of edge and propagates the knowledge to find depth of the given image [5]. The results from this method are less accurate as compared to data-driven methods. Also this method is not suitable for indoor human made objects.

III. PROPOSED METHOD

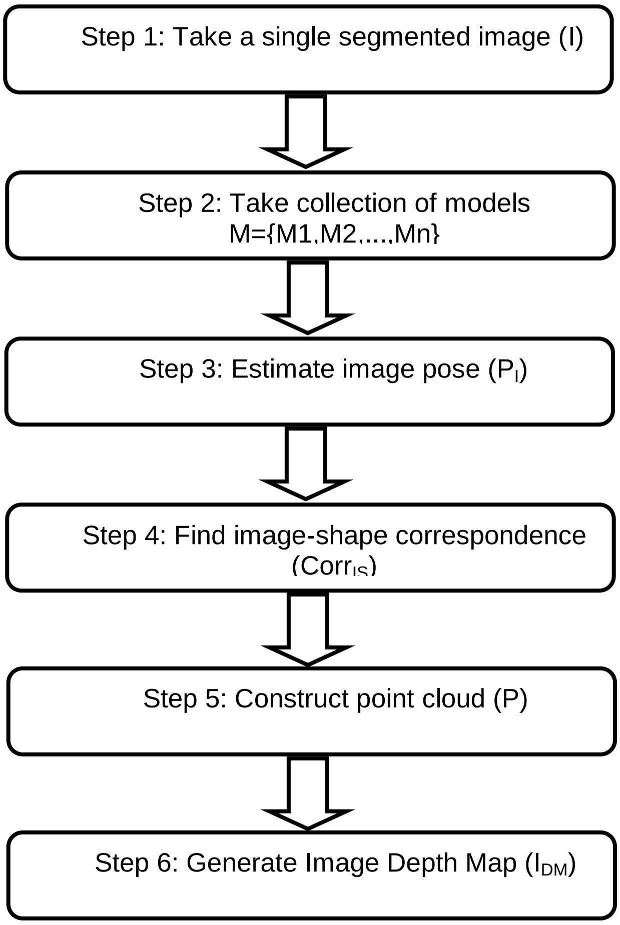


Fig.1 Flow of proposed work

Step 1: Take a single segmented image (I)

In the first step, we take a single image segmented from its background for which we are willing to find the depth information. Acquiring segmented image from its background is easy these days as many ecommerce websites uses them.

Step 2: Take collection of models M= {M_1,M_2,...,M_n}

In this step, we supply the large set of 3D models of the given input object class of the input image (I). The model collection contains large and diverse set of 3D models that guides the depth estimation process and enable us with the z-coordinates for given image. This step requires prior knowledge of object class of given input image.

Step 3: Estimate image pose (P_I)

In the pose estimation step, we initialize camera configuration parameters and estimate the exact pose of given image. The pose estimation provides base for further steps as it is used to fetch similar models from the collection. It finds local image features, such as points, lines, contours or the curve segments.

Step 4: Find image-shape correspondence ($Corr_{IS}$)

In image-shape correspondence step, we find the correspondence between image and extracted models [$Corr_{IS}$] and construct the individual pointcloud of each extracted model [7].

Step 5: Construct pointcloud (P)

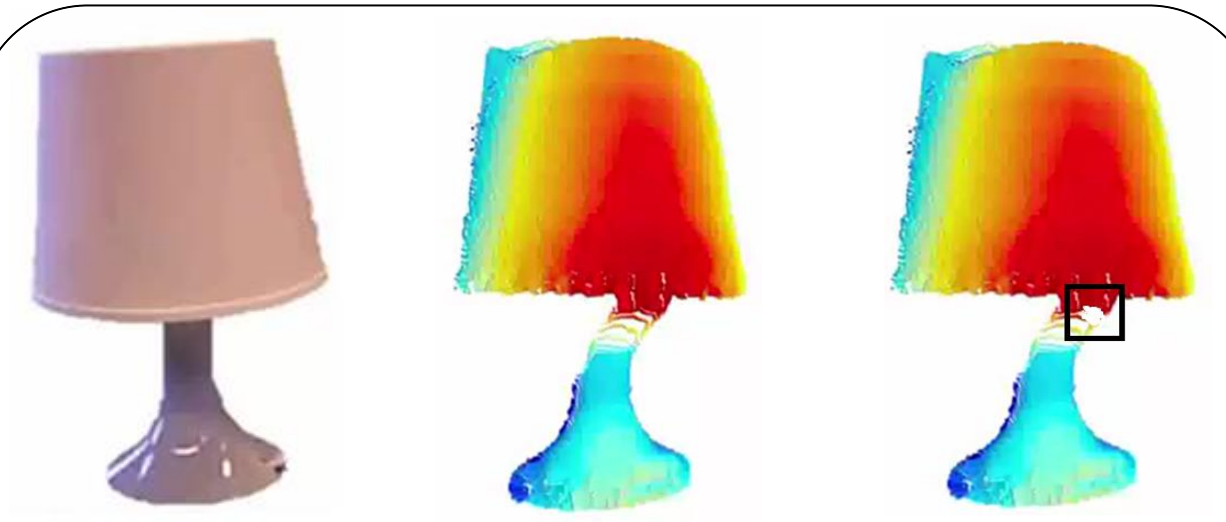
In this step, we deform each extracted shapes and generate pointclouds by rendering mesh of each extracted shape [8]. The median of all pointcloud is considered as the resulting pointcloud.

Step 6: Generate Image Depth Map (I_{DM})

At last, we transform the resulting pointcloud to the depth map in order to achieve the final image depth map (I_{DM}). Thus, we can obtain the resulting depth map of the given image

IV. EXPERIMENTAL RESULTS

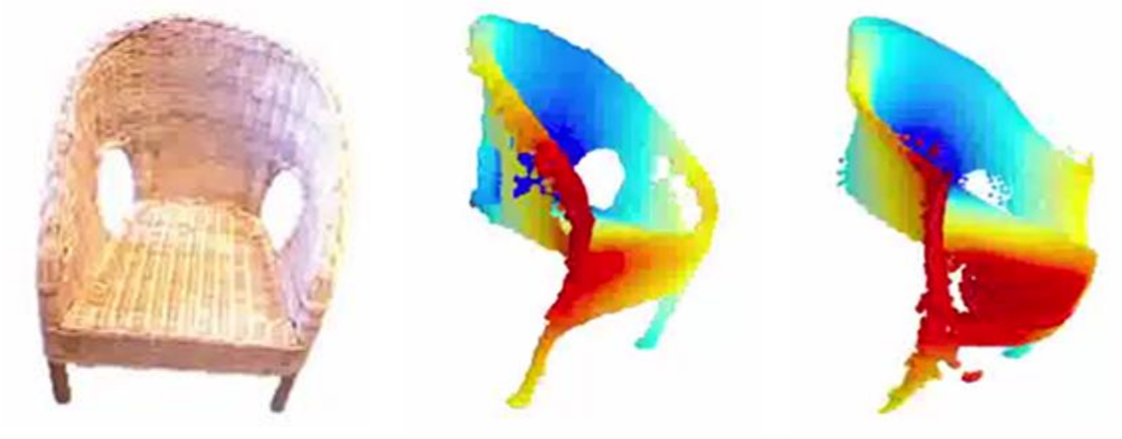
1. LAMP DATASET



Input Image | Depth Recovery [Hao's Method] | Depth recovery [Proposed Method]

Fig. 2 Depth recovery of a single segmented image of lamp

2. CHAIR DATASET



Input Image | Depth Recovery [Hao's Method] | Depth recovery [Proposed Method]

Fig. 3 Depth recovery of a single segmented image of chair

3. CUP DATASET

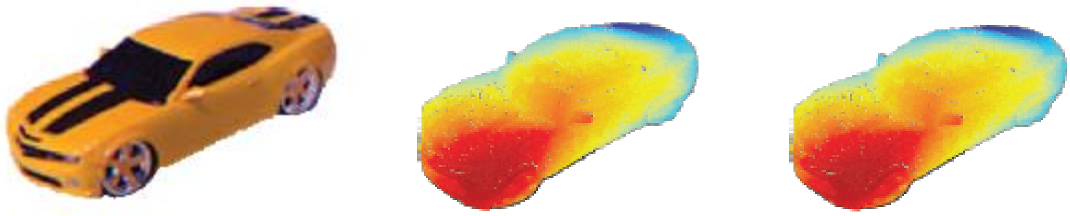
Input Image

Depth Recovery
[Hao's Method]

Depth recovery
[Proposed Method]

Fig. 4 Depth recovery of a single segmented image of cup

4. CAR DATASET



Input Image

Depth Recovery
[Hao's Method]

Depth recovery
[Proposed Method]

Fig. 5 Depth recovery of a single segmented image of car

V. Results comparison

Model Class	Number of Models	Processing Time [In seconds]	
		Hao's method[1]	Proposed method
Lamp	2.0 K	1811	1659
Chair	7.3 K	4807	4734
Cup	1.1 K	688	602
Car	1.7 K	1593	1476

Configuration setup [2.4 GHz Core 2 Duo 8GB RAM]

Dataset: Trimble 3D Warehouse

VI. Discussion

We have applied our algorithm to standard dataset from Trimble 3D warehouse and the results are shown above.

The lamp dataset contains 2000 models from sketchup. The total time for processing lamp dataset and finding depth consumes 1659s which is much faster than the Hao's method [1811s] and the generated depth map is also quite similar to the results obtained using Hao's method and obviously better than the kinect scan results.

Similarly, the algorithm was applied on chair dataset of 7300 models from sketchup and it consumes 4734s which is much faster than Hao's method [4807s] and the resulting depth map is also quite similar to the results obtained using Hao's method and are better than the kinect scan results.

Then, we applied our algorithm on cup dataset of 1100 models from sketchup and it consumes 602s which is much faster than Hao's method [688s] and the resulting depth map is also quite similar to the results obtained using Hao's method and are better than the kinect scan results.

At last, we applied our algorithm on car dataset of 1700 models from sketchup and it consumes 1476s which is much faster than Hao's method [1593s] and the resulting depth map is same as the results obtained using Hao's method [1] and are far better than the kinect scan results.

VII. Conclusion

In this paper, we have presented a model-based technique that adds depth knowledge to the input image object by transferring z-coordinates from input collection of models of same class to the image. The proposed method takes a single image of an object separated from its background and collection of models of similar object class and identifies depth information automatically. The proposed method was applied on standard dataset of lamp, chair, cup and car and it gives plausible results and takes much less time to process collection of models and generate depth map as compared to Hao's method.

VIII. References

[1] Hao Su, Qixing Huang, Niloy J. Mitra, Yangyan Li and Leonidas Guibas "Estimating Image Depth using Shape Collections" july-2014, ACM ToG

[2] Changhwan Chun, Dongjin Park, Wonjun Kim, and Changick Kim – SAMSUNG "Floor detection based depth estimation from a single indoor scene", 2013 IEEE

[3] Saeed Mahmoudpour and Manbae Kim, "A Novel Depth Estimation Method Using Infocused and Defocused Images", 2014 IEEE

[4] Ashutosh Saxena, Sung H. Chung and Andrew Y. Ng, "Learning Depth from Single Monocular Images"

[5] Dmitry Akimov, Dmitry Vatolin, Maxim Smirnov "Single-Image Depth Map Estimation Using Blur Information" 2011, GraphiCon International Conference

[6] AVERKIOU, M., KIM, V., ZHENG, Y., AND MITRA, N. J. 2014. "Shapesynth: Parameterizing model collections for coupled shape exploration and synthesis". CGF

[7] Vladimir G. Kim, Wilmot Li, Niloy J. Mitra, Stephen DiVerdi, Thomas Funkhouser "Exploring Collections of 3D Models using Fuzzy Correspondences" July 2012, ACM ToG

[8] SUMNER, R. W., SCHMID, J., AND PAULY, M. 2007. "Embedded deformation for shape manipulation". ACM ToG 26, 3 (July)