First International Conference on Computer Engineering
International Journal of Scientific Research in Science and Technology
Print ISSN: 2395-6011 | Online ISSN: 2395-602X (www.ijsrst.com)
Volume 5 Issue 8, November-December-2020

A Survey on 3D Model Generation from Images

Akash Chaudhari[1], Aditya Deo[1], Mahesh Badhe[1], Ritesh Patidar[1], Dr. Sunil Rathod[2]

[1]Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra India

[2]Professor, Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegaon, Maharashtra India

ABSTRACT

3D models are used for a variety of domains including video games, movies, architecture, illustration, engineering, and commercial advertising. We have seen significant progress in 3D model generation and reconstruction in recent years. There are multiple approaches, or method to do it. We discuss about different approaches in this paper, such as 3D Generative Adversarial Network (GAN), Differentiable Interpolation-based Renderer (DIB-R), Hierarchical Surface Prediction. We also discuss advantages, and limitations of these approaches. In the end it shows the results produced by them.

Keywords: 3D Model, GAN, DIB-R, Image, Neural Net

I. INTRODUCTION

A. What is a 3D Modeling ? [9]

3D modelling is the process of using software to create a mathematical representation of a 3-dimensional object or shape. The created object is termed a 3D model and these 3-dimensional models are employed in a spread of industries. The 3D modelling process produces a digital object capable of being fully animated, making it an important process for character animation and lighting tricks. The core of a model is the mesh which is a collection of points in space. These points are mapped into a 3D grid and joined together as polygonal shapes, usually triangles or quads. Each point or vertex has its own position on the grid and by combining these points into shapes, the surface of an object is created.

B. 3D Models generation from Images

Images are 2D, and they do not have complex structure, they are just 2D matrices of pixel values. Images often contain objects from real world like a car, table, or a bird. These real-world objects are 3D, and can be represented in 3D graphics also, as we see in animated movies, and games. Creating 3D model manually is very time-consuming task, therefore automation is must. Hence, 3D Model generation from images(2D). We in this survey paper try to explain the progress being made in this area.

C. What is GAN ? [9]

A Generative Adversarial Network (GAN) has two neural net (Generator and Discriminator): The generator learns to generate plausible data. The generated instances become negative training

Copyright: © the author(s), publisher and licensee Technoscience Academy. This is an open-access article distributed under the terms of the Creative Commons Attribution Non-Commercial License, which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited

examples for the discriminator The discriminator learns to differentiate the generator's fake data from real data. The discriminator penalizes the generator for producing implausible results. And in the end we can use that generator alone to generate our own examples/data which never existed i.e. 3D model in this case.

II. LITERATURE SURVEY

A. 3D Reconstruction from single 2D Image [2]

This was published by Deepu R, and Murali S in 2016. It uses corners of objects in an image, and then calculate depth, and length of objects. This method requires only one image. In some images, corners detection-based approach cannot give enough depth information to reconstruct 3D model. As this is not a complete ML based approach to generate 3D models, it can fail often on real world objects.

B. Hierarchical Surface Prediction for 3D Object Reconstruction [7]

Convolutional Neural Networks have shown promising results for 3D geometry prediction. They can make predictions from very little input data such as a single-color image. A major limitation of such approaches is that they only predict a coarse resolution voxel grid, which does not capture the surface of the objects well. This proposes a general framework, called hierarchical surface prediction (HSP), which facilitates prediction of high-resolution voxel grids. The main insight is that it is sufficient to predict high resolution voxels around the predicted surfaces. The exterior and interior of the objects can be represented with coarse resolution voxels. This approach is not dependent on a specific input type. It shows results for geometry prediction from color images, depth images and shape completion from partial voxel grids. Our analysis shows that high resolution predictions are more accurate than low resolution predictions in this method.

C. 3D GAN [1]

3D-GAN applied GAN in learning latent 3D space, and it can generate 3D voxel models from the latent space by extending 2D convolution into 3D convolution. Combining 3D-GAN with WGAN-GP, 3D-IWGAN can generate high-quality 3D models with a more stable training process. Wang et al. utilized an encoder–decoder as generator of the adversarial network to address 3D shape inpainting. Then a long-term recurrent convolutional network (LRCN) was employed to refine the generated results to obtain more complete 3D models in higher resolution. Chen et al. proposed text2shape system which combined 3D generation with natural language processing. The network encoded the text, then regarded the results as a condition, and utilized WGAN to decode it into a 3D model related to input text.

D. Learning to Predict 3D Objects with an Interpolation-based Differentiable Renderer (DIB-R) [5]

Unlike other approaches, this uses a whole new way to generate 3D models. It introduces a new renderer called DIB-R. DIB-R gives control to the end of how pixel value gets calculated in the rendering process of a 3D model. It introduces a new rendering pipeline just like OpenGL, and DirectX have their own pipelines. The difference is DIB-R has differentiable pipeline, which means it open doors to Machine Learning in rendering process. A GAN trained using the method proposed in this paper can predict 3D model with lighting, and texture.

The overall finding of the above discussion is given below as Table-1

© 2020 IJSRST | Print ISSN: 2395-6011 | Online ISSN: 2395-602X | Published : 18/12/2020 - 273-277

Table -1: Critical Review of Research Papers

Sr. No.	Paper Name	Advantages	Limitations
1.	Learning to Predict 3D Objects with an Interpolation-based Differentiable Renderer	It can generate 3D model with texture, and lighting with complete control of rendering	- Too much computational power requirement - Model resolution can be improved
2.	3D Reconstruction from single 2D Image	It uses corners of objects in an image, and then calculate depth, and length of objects. It requires only one image.	In some images, corners detection-based approach cannot give enough depth information to reconstruct 3D model.
3.	Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modeling. (3D-GAN, 3D VAE-GAN)	ML based approach, learns latent space (space where similar points are closer) from training data. It can give good results even with less than 25 training examples per class.	The model learns latent space, but cannot find relationship between different groups of similar points in the space, which can give better results.
4.	Hierarchical Surface Prediction for 3D Object Reconstruction	- Faster than GAN which uses differential rendering. - Can generate model when single input image given	- No texture, lighting with 3D model - Model is not smooth

III. TAXONOMY CHART

Attributes → Approaches ↓	3D Model Quality	Is ML based approach	Type (Image or Model)	Resources required	Hardware Required
DIB-R	HIGH	YES	Images & Models	25-30 Models 25-30 Images	GPU
3D Reconstruction from single 2D Image	LOW	NO	Image	1 Image 1 Depth Map	Basic CPU
3D GAN, and 3D VAE GAN	MEDIUM	YES	Models & Images	25 Models 25 Images	GPU
Hierarchical Surface Prediction for 3D Object Reconstruction	HIGH	YES	Image	25-30 Models 25-30 Images	GPU

© 2020 IJSRST | Print ISSN: 2395-6011 | Online ISSN: 2395-602X | Published : 18/12/2020 - 273-277

IV. RESULTS

Results of various method or approaches are shown in below Figure-1.

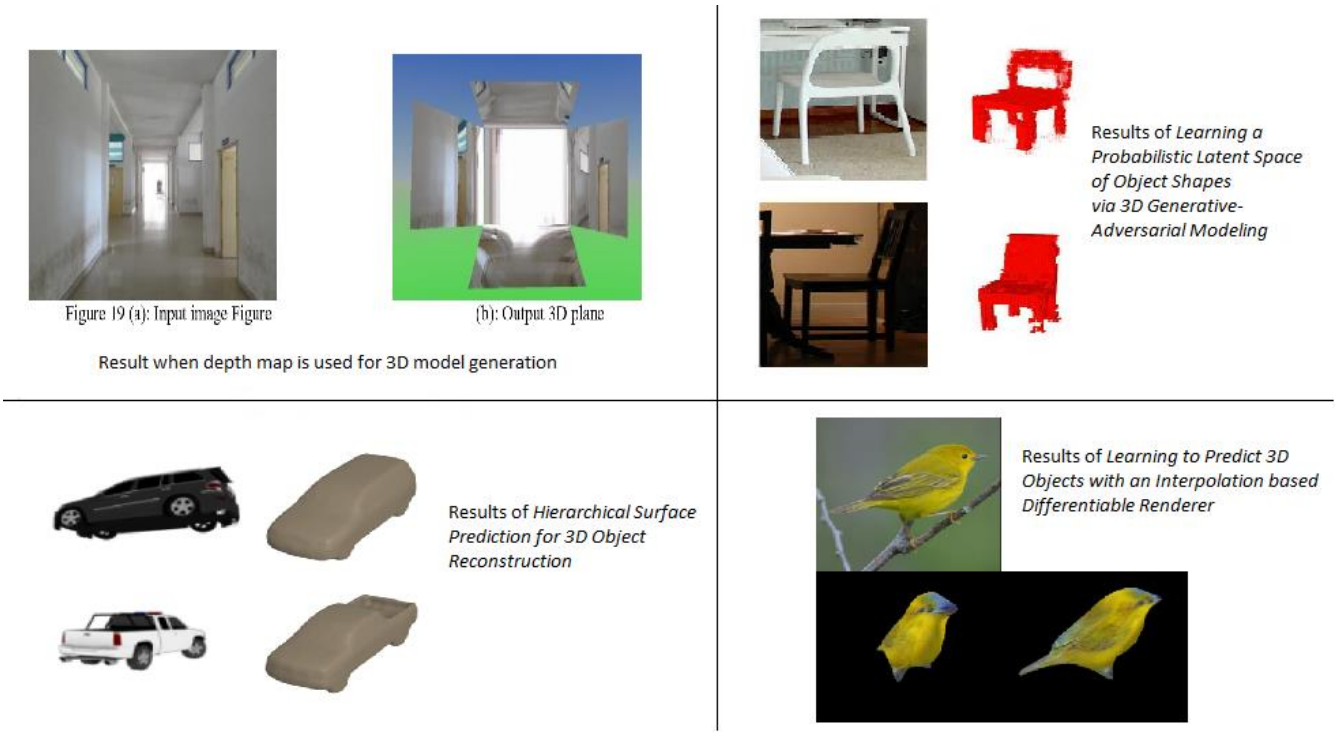

Figure-1: Results of various research papers

V. CONCLUSION

We observed that a Model should find relationship between different groups of similar points in the space, and reduce time-consumption for output generation [1]. Finding depth (z index) with a corner detection-based approach cannot give good results [2]. We also observed that if we make rendering pipeline differentiable, then it opens possibility for Machine Learning to control rendering. Taking rendering into consideration also opens possibility to predict textures, and lighting. Some approaches are very close to predict perfect 3D models which can be used in production, but still there is still need of improvement. Computational power, and data required are two main factors which need to considered.

VI. ACKNOWLEDGEMENT

It gives us a great pleasure in presenting the paper on "A SURVEY ON 3D MODEL GENERATION FROM IMAGES". We would like to take this opportunity to thank Dr. Pankaj Agarkar, Head of Computer Engineering Department, DYPSOE, Pune for giving us all the help and support we need during course of the Paper writing work. We are really grateful to him for giving an opportunity to work with R&D cell of our department. Our special thanks to Dr. Ashok Kasnale, Principal DYPSOE who motivated us and created a healthy environment for us to learn in the best possible way. We also thank all the staff members of our college for their support and guidance.

VII. REFERENCES

[1]. Learning a Probabilistic Latent Space of Object Shapes via 3D Generative-Adversarial Modelling, Jiajun Wu*, Chengkai Zhang*, Tianfan Xue*, William T. Freeman*, Joshua B. Tenenbaum* (2016)

[2]. 3D Reconstruction from Single 2D Image, Department of CSE, Maharaja Institute of Technology Mysore, India, January 2016

[3]. Image Generation with Disentangled 3D representation Jun Yan Zhu* (6 Dec 2018)

[4]. Unsupervised learning of 3D representations from natural images, Thu Nguyen-Phuoc*(1 Oct 2019)

[5]. Learning to predict 3D objects with an interpolation- based differentiable renderer. (2019)

[6]. https://en.wikipedia.org/wiki/3D_rendering

[7]. Hierarchical Surface Prediction for 3D Object Reconstruction Christian Häne, Shubham Tulsiani, Jitendra Malik (3 Apr 2017)