IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Research on monocular depth estimation and image refocusing

Sudarshan Gampa[1], Varun Andharia[2], Mr.Ashraf Siddiqui[3]

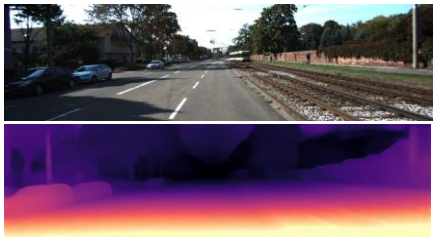
[1]*Mumbai India*
[2]*Mumbai India*
[3]Assistant Professor Department of Computer Engineering, Universal College of Engineering, Vasai, Maharashtra, India

Abstract: *In this paper we use a monocular depth estimation method to estimate depth in a given input image.This is performed by using a unsupervised deep learning model to estimate depth and get the depth image of the given input image using this depth image we are going to perform image processing on the input image to give bokeh effect to the input image (portrait image) this is one application of the depth information we get from the depth map . Main aim is to use the depth map which contains the depth information or the information of the object distance from the camera lens, to get a 3d representation of the image and perform various applications using depth information.*

***Key Words*: Depth estimation, , Machine Learning, Deep Learning, Portrait image , Bokeh Effect, Image processing**

1.INTRODUCTION

We will be taking a single color image as an input in our model to get a depth image so that the process of refocusing to get the output image as portrait image.
Estimating depth is crucial step towards inferring scene geometry from 2D images.The goal is to predict the depth value of each pixel in a given only a single RGB image as in input.Since the same input image can project to multiple plausible depth estimation.There are many supervised learning methods to predict depth,this supervised learning [1] method use partial ground truth depth as an input to the model while training, this ground truth depth is hard to obtain because of costly hardware.To overcome this we have used an unsupervised monocular depth estimation method which only uses two input images while training.



The applications of depth estimation are self-driving car ,parking car sensors, image-editing, AR-composting,etc, The application of depth estimation in our project is to get a portrait image by our method where we will take an input image to get a depth image and then use it for performing the refocusing by the bokeh effect.

2. METHODOLOGY

The first part of this thesis is mostly a **literary review** of recent and relevant research in the field of computer vision, specifically the task of depth estimation[1][2] from both stereo and single images[1][2]. All emphasis was made on studies and research papers that treated this problem with machine or deep learning. We have used the method and model based on paper Digging Into Self-Supervised Monocular Depth Estimation.[2]. They train a machine learning model to generate a depth map using just a single image.The link to their Paper is.[2].

Our Method consists of the following steps:

1. We will select an input image for performing the refocusing or applying the bokeh effect[8] to the input image.
2. Input image will be used to predict the depth image and this depth image will be used in the image processing task to apply the bokeh effect[8] to the input image.

3. WORKFLOW OF PROJECT

The Design of the Gui is to first take the input image then use the monodepth2 [2] model to infer a dense depth image[1][2][3][4] and save it.After that use this depth image and its npy file to perform image processing and apply bokeh effect to the input image and get the resultant portrait image.
The operations are divided into 3 parts:
1. Selecting the input Image from the image folder .
2. Then select the continue button which will use the model to predict the depth map and save the depth map in the folder.
3.Clicking a point on the input image so that bokeh effect can be applied to the pixel surrounding the point or object . Finally the portrait image is
Displayed and saved in the folder.



Input image **Depth image** **Portrait image**

4. EXPERIMENTAL STUDY

In this, the application will be tested with different inputs to check whether the platform provides desired and relevant results as expected to happen.

The table below shows the results as follows :

Input no.	Input type	Depth map output	Portrait image output
1	Human 1st Sample image	Predicted successfully	Applied successfully
2	Human 2nd Sample image	Predicted successfully	Applied successfully
3	Human 3rd Sample image	Predicted successfully	Applied successfully
4	Butterfly 4th Sample image	Predicted successfully	Partially applied
5	Flower 5th Sample image	Predicted successfully	Partially applied
6	Multiple flower 6th " "	Predicted successfully	Not applied
7	Tree bark 7th " "	Predicted successfully	Not applied

5. CONCLUSIONS

Thus with the help of all these unsupervised learning based methods[1][2][3][4] our main aim is to predict depth and use the most accurate algorithm Instead of using aligned ground truth depth data, which is both rare and costly for supervised algorithms and use this depth information [2]in many applications in real time Image processing , scene reconstruction , self driving cars, AR composition and many other applications.

ACKNOWLEDGEMENT

Our model is based on the paper"Digging Into Self-Supervised Monocular Depth Estimation, by C. Godard, O. Mac Aodha, M. Firman, G. Brostow, arXiv:1806.01260". We thank the above authors for providing us with their amazing model and knowledge to use in our paper.

REFERENCES

[1] Clement Godard ,Oisin Mac Aodha, Gabriel J. Brostow, Unsupervised Monocular Depth Estimation with Left-Right Consistency, source https://arxiv.org/abs/1609.03677

[2] Clément Godard, Oisin Mac Aodha, Michael Firman, Gabriel Brostow,Digging Into Self-Supervised Monocular Depth Estimation https://arxiv.org/abs/1806.01260

[3] Xiaohan Tu; Cheng Xu; Siping Liu; Guoqi Xie; Renfa , Real-Time Depth Estimation with an Optimized Encoder-Decoder Architecture on Embedded Devices Lihttps://ieeexplore.ieee.org/abstract/document/8855335

[4] KITTI Single Depth Evaluation Server. http://www.cvlibs.net/datasets/kitti/eval depth.php? benchmark=depth prediction. 2017.

[5] Clement Godard, Oisin Mac Aodha, and Gabriel J Bros- ´ tow. Unsupervised monocular depth estimation with left right consistency. In CVPR, 2017.

[6] Nikolaus Mayer, Eddy Ilg, P. Häusser, Philipp Fischer, D. Cremers, Alexey Dosovitskiy, Thomas Brox ,IEEE International Conference on Computer Vision and Pattern Recognition (CVPR), 2016, https://lmb.informatik.uni-freiburg.de/Publications/2016/MIFDB16/

[7] Nikolaus Mayer, Eddy Ilg, Philip Hausser, Philipp Fischer, Daniel Cremers, Alexey Dosovitskiy, and Thomas Brox. A large dataset to train convolutional networks for disparity, optical flow, and scene flow estimation. In CVPR, 2016.

[8] Muneeb ul Hassan ,Pix2Pix – Image-to-Image Translation Neural Network,Source: https://arxiv.org/pdf/1611.07004.pdf

[9] Bo Li, Yuchao Dai, and Mingyi He. Monocular depth estimation with hierarchical fusion of dilated cnns and soft weighted-sum inference. Pattern Recognition, 2018.

[10] Huaizu Jiang, Erik Learned-Miller, Gustav Larsson, Michael Maire, and Greg Shakhnarovich. Self-supervised relative depth learning for urban scene understanding. In ECCV, 2018