ISSN No:-2456-2165

An Efficient Approach to Single-Image Refocusing

Chandana N, Tejaswini R Pattange, Kavya Rani S R, Soundarya B C
Department of Computer Science and Engineering,
Vidyavardhaka College of Engineering,
Mysuru, Karnataka, India

Abstract: Post-capture image refocusing is a high-demand application, where certain portions of the image are refocused to attract the viewers' attention towards it. Focusing on a specific element of the image is achievable by calculating the depth of the image. Objects are at similar depth during the capture, which brings in the depth error, and the desired object is out of focus. During such circumstances, professionals post-capture refocus the image to obtain the desired output. This paper gives an analysis of calculating the depth of the image required to synthetically blur non-preferred portions and different techniques used to deblur and produce the desired output image.

Keywords:- *Bokeh, depth map, Depth of field, Synthetic refocus in.*

I. INTRODUCTION

Bokeh, a term forged by Japanese photographers' is the aesthetic quality of the blur in an out-of-focus image. To naturally produce the bokeh effect in the captured image, it is necessary to use high-end professional photography equipment. And is highly connected to the depth of field of the lenses used in this equipment. Since most of the images captured in the present day are with conventional devices like mobile phones, which generally have a narrow aperture it is difficult to capture images with a shallow depth of field. But as the bokeh effect is highly sought-after in modern digital photography, it is necessary to have tools to produce bokeh images from a single all-in-focus deep-focused image. But synthetically mimicking this effect can be quite a tedious task as the images might have been captured with an optical system with alowaperture where all the elements of the image are at the same depth.

In this paper, we discuss synthetic refocusing, a technique for producing an image with a different depth of field with a single image, thus creating a bokeh effect. To achieve this depth map must be generated calculating the distance of the element from the point of focus. This depth map then, along with the all-in-focus camera shot is fed into a refocusing system to produce a bokeh image. But various challenges can be experienced while estimating the depth map.

II. LITERATURESURVEY

Photography is the art of capturing light with a camera. Photographers use various techniques to display their ideology to the world. One of the most valued effects is the bokeh effect. It helps photographers to attract the viewer's attention to a particular section of the image, by sharpening the element and blurring the rest of the images. Depth of field is an essential component that refers to the range of depths in the scene. This range is primarily determined by the aperture of the camera lens, to get shallow depth-of-field wide aperture is used whereas to generate wide depth-of-field narrow aperture is used. In bokeh images, there is an effective background blur and shallow depth of field. This effect is best generated by a digital single-length reflex camera (DSLR) which has a wider aperture[1]. The operation of DSLR is complex and hence photographers need additional training to manage various settings to get the desired effect. Since these cameras have expensive hardware, not all common people can afford them.

III. DEPTH MAPESTIMATION

In today's world smart phones are capable of mimicking complex software. One of the most recent efforts is to generate high-quality bokeh images in smart phones. As mentioned above this effect requires a wide aperture in contrast smart phone camera lenses have a narrow aperture that can only generate a wide depth-of-field. To overcome this challenge smart phones have introduced dual cameras instead of one. Smartphone manufacturers are aiming to generate shallow depth-of-field images. One of the most used techniques is to inculcate dual cameras [2] instead of single ones. The images captured by two cameras are given as input to Stereo Algorithms which compute depth maps. This depth map is used to blur the background of one of the images. This results in the second camera as an additional component that increases cost and power consumption during use. This also causes issues with space complexity in the mobile. Few manufacturers instead prefer including special sensors like time-of-flight and structured-light depth sensors. These sensors make the system expensive, higher power consumption, and in addition, do not function well outdoors. In recent years researchers have come up with advanced technology of lens blur method [4] which enables it to generate shallow depth images without any additional hardware. But this technique demands the user to move the phone during capture to introduce parallax. If the photographer fails to progress the camera at the correct trajectory and speed or if the object in the scene moves, this can result in the negative experience of the photographer and fail to capture desired bokeh image. To overcome these obstacles researchers are working on generating post-capture bokeh images In recent years, there has been substantial research work being done for calculating the depth information from a single monocular image. Through depth information, we get various helpful clues such as horizontal boundary, vanishing point location, etc. Hence depth estimation is considered to be one of the primary and prominent functions in computer vision. This process acts as a prerequisite for many advanced operations such as 3D scene modeling and reconstruction, simultaneous localization and mapping (SLAM), navigation, object detection, and semantic segmentation and also image refocusing, object detection, scene parsing, pose estimation, visual tracking, semantic segmentation, shape analysis, image-based rendering, and 3D reconstruction, etc.

- **Sensor-based :** The most commonly used method for depth estimation is using a specific type of depth estimating sensor. These devices are used to generate 3D reconstruction which helps to avoid running complex structures from motion (SFM) routines to produce high-quality models. RGB-D sensors are one of these which work at the pixel level in association with RGB sensor cameras to generate dense depthmaps [5]. Many RGB-D sensors such as Microsoft Kinect, Structure IO, ASUS Xtion Pro, and Intel RealSense are currently available. But this method does not guarantee high-quality depth images as it has a limited range and is outdoor sunlight sensitive. Besides, LiDAR(Light Detection and Ranging), the remote sensing device commonly used in self-driving cars [6], estimating crop density [7] and measuring the distance of objects on the surface of the earth, is also used to generate accurate 3D information about an article with relation to its environment. It measures the time taken by the laser pulse to travel from one object's position to another surface. This time is further used to calculate the depth map. This technology is best suited for large-size and power-consuming devices. That makes these sensors inconvenient for low-cost small size and wide applications of monocular depth estimation.

- **Stereo Vision**: Stereo vision matching aims at imitating human eyes to generate 3D structures of a scene using two viewpoints of the same scene. It uses two cameras to calculate the disparity maps of images through a cost function. (Benjamin Busam el al) [3] Introduced SteReFo pipeline utilizing this method to leverage a state-of-art to obtain a high-quality disparity map. The pipeline consists of two Siamese towers to draw out features of the deep image at 0.125 of the stereo pair resolution. The recovery of high-frequency detail is performed by residual refinement. Although this method successfully calculates depth values it depends on image pairs or sequence of images. The challenge of generating depth maps from a single image persists.

- **Deep Learning:** Deep learning or deep neural networks has been a major contributor in the field of image processing, object detection, classification, and semantic segmentation, etc. In the field of depth map estimation, an end-to-end manner-based deep learning algorithm is exploited to produce a pixel-level depth map from a single image. Various neural networks are implemented to address the challenge of monocular depth map estimation, a few of which are recurrent neural networks (RNNs), convolutional neural networks (CNNs), variational auto-encoders (VAEs), and generative adversarial networks (GANs). This method achieves a strong expressive ability to calculate single-view image depth maps but its reliance on the result is questionable. And its optical flow algorithms become excessively complex for smart phones to accommodate.(MINSOO SONG1) [8].

- **Unsupervised Learning:** Supervised learning techniques to generate depth information require a large quantity of ground truth depth data as training data set. The generation of high-quality data sets of a range of environments brings challenges. Cl`ement Godard, OisinMac Aodha and Gabriel J. Brostow introduced the [9] Unsupervised monocular depth estimation enables CNN to train in the absence of ground truth depth data. This method during its training considers depth estimation problems as an image reconstruction that might yield inconsistent depths in case of spectacular and transparent surfaces. A reconstruction function is used to reconstruct one image from another to observe information of the imaged 3D shape of the scene. The depth correspondence field is found, which can be applied to the left image to reconstruct the right image. The requirement of rectified and aligned stereo pairs for training purposes makes it currently not feasible to use the single view data set that is existing. Chuanxue Song, Chunyang Qi, Shixin Song and Feng Xiao [10] propose an idea based on Uncertainty Analysis and Retinex Algorithm. Uncertainty analysis helps to deal with the challenges of less reliable outputs and low prediction accuracy by providing improved accuracy and quantified output results. Coupled with Retinex lighting theory, interference problems caused by dynamic objects can be resolved as it is utilized to construct photometric loss functions.

IV. REFOCUSING

In this rapid advancement of technology various fields such as video gaming, image processing demand for efficient algorithms to post-process the image by refocusing the image to look similar to natural capture. Refocusing algorithms are provided with the depth map of a scene to generate the refocused image. Utilizing different depths of the image (Benjamin Wiberg)[11] proposed a method to refocus a post capture image using convolution with a blur kernel whose size depends on that pixel. This method involves four steps which are arranging the images in the focal stack, creating a sharp image to calculate the depth map using $I^2 + I^2$,where I_x represents a convolved image with a horizontal Sobel kernel and $I_y$image convolved with a vertical Sobel kernel. To estimate the final depth map, a near focused depth map and far focused depth map are combined ($I_{depthnear}$+(max($I_{depthfar}$) - $I_{depthfar}$).

The final step of computing refocused image, the amount of blur is calculated using the following equation $I_{blur} = |I_{depth} - d|$. The sharp$_x$ image$_y$ is convolved with a Gaussian kernel whose size varies spatially based on the blur amount I_{blur}. Edges of the objects can also be used in estimating the blur as explained by (Wei Zhang, Wai-Kuen Cham) [12] where they first identify the edges and its blurriness depicted using a parametric model to estimate the depth map. To refocus the image a single-image blind deconvolution framework (SBD), with both global and local constraints is used to model prior image. The blurry edges are sharped by the parametric edge model and to ensure the sharpness of the image queues are rendered as local prior. This method has been proved to have low computational complexity and also user initialization can be exempted. To achieve the aim of refocusing Yosuke Bando and Tomoyuki Nishita[13] put forward a method to refocus a single input image which contains prior defocus blur. Theyfirst estimate the existing defocused blur by deconvoluting differently blurred images that are generated with each predetermined blur radii to remove spatially variant defocus blur. Among these deblurred images ``best" images are selected and using sectional method latent images are obtained. The latent image undergoes convolution with new desired blur radius which is obtained by interactive stage followed by linear interpolation for new blur radius and blurred images.At the final step once again sectional method is employed to stitch the desired refocused image. A similar method was suggested by Parikshit Sakurikar, Ishit Mehta, Vineeth N. Balasubramanian and P. J. Narayanan [14]. RefocusGAN which uses deblur then re-blur approach. The first step deblurs the input image to obtain the radiance of the scene. The proposed generator network consists of two strided convolution blocks , two transposed convolution blocks and nine residual blocks which is based on ResBlock architecture. The wide aperture image produced along with computed radiance, the refocused image is developed based on a refocus control parameter δ. blur-and-compare framework is employed to compute the per-pixel blur radius between the input image G1 and the radiance image ˆGr. The best defocus radius $\sigma$is estimated which is then inverted as $\sigma' = \sigma_{max} - \sigma$to re-blur the radiance image and generate desired refocused image.

V. CONCLUSION

In this analysis we aim to contribute in the field of image processing. We discuss solutions to address one of the major requirements of computer graphics, video gaming, professional photography to post-process the image. The images are desired to refocus as per requirement. This maily demands generation of accurate depth map which then coupled with refocusing algorithms to achieve refoucing. We have surveyed techniques to compute depth maps which include sensor based , stereo vision and deep learning methods. Refocusing algorithms employ various techniques such as using different angles and distance of the image , considering the edge blurriness and proficiency to initially deblur later perform rebluring using the blur radius. This survey would assist in development of efficient techniques and algorithms to face this challenge of artificially refocusing a single image presented as a natural capture.

REFERENCES

[1.] Nealwadhwa, Rahul Garg, David E. Jacobs, Bryan E. Feldman, Nori Kanazawa, Robert Carroll, Yair Movshovitz-attias, Jonathan Barron, Yael Pritch, and Marc Levoy "Synthetic Depth-of-Field with a Single-Camera Mobile Phone" Google Research, 2016.

[2.] Idris Abdul Ghaffar, Mohd Norzali Haji Mohd "Selection of Android Smartphones with Built-in Dual Lens Camera for Stereo Vision Android App Development" University Tun Hussien Onn Malaysia, 2018.

[3.] Benjamin Busam, Matthieu Hog, Steven McDonagh and Gregory Slabaugh, "SteReFo: Efficient Image Refocusing with Stereo Vision", Huawei Noah's Ark Lab and TechnicalUniversity of Munich.

[4.] Yuejiao Su, Yuan Yuan, Zhiyu Jiang (School of Computer Science and School of Artificial Intelligence, Optics and Electronics (iOPEN), Northwestern Polytechnical University, Xi'an 710072, P.R. China), "DeepFeature Selection-and-fusion for RGB-d Semantic Segmentation",2021.

[5.] Yan Wang, Wei-Lun Chao, Divyansh Garg, Bharath Hariharan, Mark Campbell, and Kilian Weinberger, "Pseudo-LiDAR from Visual Depth Estimation: Bridging the Gap in 3D Object Detection for Autonomous Driving" ,Cornell University, Ithaca, NY,2018.

[6.] "Estimation of the crop density of small grains using LiDAR sensors"

[7.] Minsoo Song and Wonjun Kim, (Member, IEEE) "Depth Estimation from a Single Image Using Guided Deep Network", 2019.

[8.] Cl´ement Godard, OisinMac Aodha andGabrieJ. Brostow "Unsupervised Monocular Depth Estimation withLeft-Right Consistency",University College London, 2017.

[9.] Chuanxue Song, Chunyang Qi, Shixin Song and Feng Xiao, "Unsupervised Monocular Depth Estimation Method Based on Uncertainty Analysis and Retinex Algorithm",2020.

[10.] Benjamin Wiberg, "A Method for Refocusing Photos using Depth from Defocus", Department of Electrical Engineering Stanford University.

[11.] Wei Zhang, Member, IEEE, and Wai-Kuen Cham, Senior Member, IEEE, "Single-Image Refocusing and Defocusing" IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 21, NO. 2, FEBRUARY2012.

[12.] Yosuke Bando and Tomoyuki Nishita, "Towards Digital Refocusing from a Single Photograph"

[13.] Carlos Hernández, "Lens Blur in the new Google Camera app", 2014

[14.] Saikat Duttaa , "Depth-aware Blending of Smoothed Images for Bokeh Effect Generation" 2020

[15.] Francesc Moreno-Noguer, Peter N.Belhumeur and Shree K. Nayar ,"Active Refocusing of Images and Videos"

[16.] Chaoqiang Zhao, Qiyu Sun, Chongzhen Zhang, Yang Tang and Feng Qian "Monocular Depth Estimation Based On Deep Learning: An Overview"

[17.] Simon Evain and Christine Guillemot "A Neural Network with Adversarial Loss for Light Field Synthesis from a Single Image", 2021

[18.] Lijun Wang, Jianming Zhang, Oliver Wang, Zhe Lin, and Huchuan Lu, "SDC-Depth: Semantic Divide-and-Conquer Network for Monocular Depth Estimation", 2020.