Post-processing Spatially-Varying Blur Effects with Learned 6D Blur Fields

Zhecheng Wang, Esther Lin
{zhecheng, lin}@cs.toronto.edu

Motivation

The point spread function (PSF) is the response of an optical imaging system to a point light source. At its essence, the PSF dictates a camera's rendering of a scene by determining the mapping of any point in 3D space onto the captured image. Consequently, accurately modeling an imaging system's PSF is akin to extracting its operational behavior, lending itself immediately to depth estimation and deblurring applications.

There is ongoing work at the University of Toronto on estimating and representing PSFs. They introduce the novel use of a coordinate-based neural network for representing the PSF. This representation is faithful to the underlying physics: the PSF has a multidimensional dependence on a laundry list of spatial variables, for example, location on the 2D image plane and the distance between the sensor and the lens, and the proposed method is the first to encompass these parameters holistically with the highest accuracy.

While these realistic PSFs can immediately elevate existing depth estimation and deblurring applications, there is another area that remains unexplored: rendering. In this project, we propose the use of these PSFs for two applications:

1. Creating photorealistic renderings that capture the aberrations of a real imaging system.
2. Synthesizing cinematic effects through realistic synthetic defocusing.

Both of these applications have the potential to be game-changing. First, photorealistic rendering is a battle that many professionals face. Existing rendering engines such as Blender only support the simple thin lens model, which cannot capture the spatially-varying aberrations natural imaging systems inject into an image. Utilizing these PSFs in the rendering pipeline will add realism to the results. Second, lens-adjustment-based techniques such as Dolly zoom and focus adjustment are commonly used in cinematography for dramatic effect or guiding audience attention. However, these effects are expensive in terms of manpower, time, and prior training. Our PSFs can create the same effects digitally through synthetic defocus--at a much lower cost.

Related Work

There have been multiple works trying to post-process spatial blur effects on images/videos. Among different blur effects including defocusing/refocusing/portrait mode (*Gao et al. 2020*), dolly zoom/vertigo effects (*Liang* et al. 2020) are the most challenging and artistically pleasing one as it involves constant changes of lens position and FoV. In a recent attempt by Gao *et al. 2021*, the authors jointly trained a time-invariant static NeRF and a time-varying dynamic NeRF and could achieve dolly zoom effects. However, these works do not capture the spatially-varying nuances a real camera can do in general as the PSF information is lost when capturing the RGB image.

Project Overview

This project will explore the possibility of applying coordinate-network-based PSFs to rendering. Specifically, we propose the following deliverables, ordered from most critical to least:

1. *A Blender plugin that integrates our PSFs in post-processing*: the camera module in Blender simplistically models the lens with a thin-lens model with no abbreviation and has simple f-stops and aperture sizes. In comparison, using a learned 6D blur field to apply spatially varying blurring would achieve photorealistic.
2. *A Blender UI that supports user interactivity for choosing lens model and defocus position*: This will consist of a dropdown for lens selection (i.e. iPhone 12 Wide vs. Telephoto) and an adjustable slider for focus position.
3. *Focus adjustment* [Esther]: Given an all-in-focus image of a scene, a depth map, and two points A and B on the image, generate a video where the focus is adjusted continuously from A to B.
4. *Dolly zoom*: Given an all-in-focus image of a scene, a depth map, a camera translation distance D, and a pair of focus at positions {A, B}, generate a video where the focus is shifted from focus A to B while the camera is simultaneously being pulled away from the scene by a distance D.

Milestones, Timeline & Goals

11/18 Post-process DoF on rendered images using a consistent blur kernel in Blender
11/21 Post-process DoF on rendered images using a learned 6D blur field in Blender
11/25 Complete and test the Blender add-ons on user ends
12/02 Apply learned 6D blur fields to generate dolly zoom/vertigo effects with real/synthetic images.
12/05 Apply learned 6D blur fields on more cinematic effects if time permits

Reference

1. C. Gao, A. Saraf, J. Kopf and J. -B. Huang, "Dynamic View Synthesis from Dynamic Monocular Video," 2021 IEEE/CVF International Conference on Computer Vision (ICCV), 2021, pp. 5692-5701, doi: 10.1109/ICCV48922.2021.00566.
2. Y. Liang, R. Ranade, S. Wang, D. Bai and J. Lee, "The "Vertigo Effect" on Your Smartphone: Dolly Zoom via Single Shot View Synthesis," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, pp. 1407-1415, doi: 10.1109/CVPRW50498.2020.00180.
3. C. Gao, Y. Shih, W. -S. Lai, C. -K. Liang, and J. -B. Huang, “Portrait Neural Radiance Fields from a Single Image”, arXiv:2012.05903, 2020.