Smart Focusing:
A brand new method that allows photo shooting device to focus intelligently

Author/Lecturer: Cheng-Yuan Lee

Date: Apr 22nd

Introduction:

Ever since the invention of the camera, there was one question that kept being asked: What do we do to make the photo shooted better? Better sensor, better lens,... better hardware units in general, would definitely increase the overall quality of the photo. That's one thing. The other one that most people do neglect is the software, or, the algorithms, rather, has been indispensable when it comes to taking pictures, which could be ranging from converting input light signals to denoising the synthesized result. It's undoubtedly a huge topic, so I'm going to focus on one of the fundamental problems: how do we find the focus quickly and accurately ?

Previously, big companies like Canon, Nikon and Sony have widely adopted a technique called "Phase Detection Autofocus", or, PD AF, which includes continuous focusing mode and focus-locking mode. Firstly, let's introduce what is PF first:

Phase detection: In each figure, the area within the purple circle represents the object to be focused on, the red and green lines represent light rays passing through apertures at the opposite sides of the lens, the yellow rectangle represents sensor arrays (one for each aperture), and the graph represents the intensity profile as seen by each sensor array. Figures 1 to 4 represent conditions where the lens is focused (1) too near, (2) correctly, (3) too far and (4) much too far. The phase difference between the two profiles can be used to determine in which direction and how much to move the lens to achieve optimal focus.
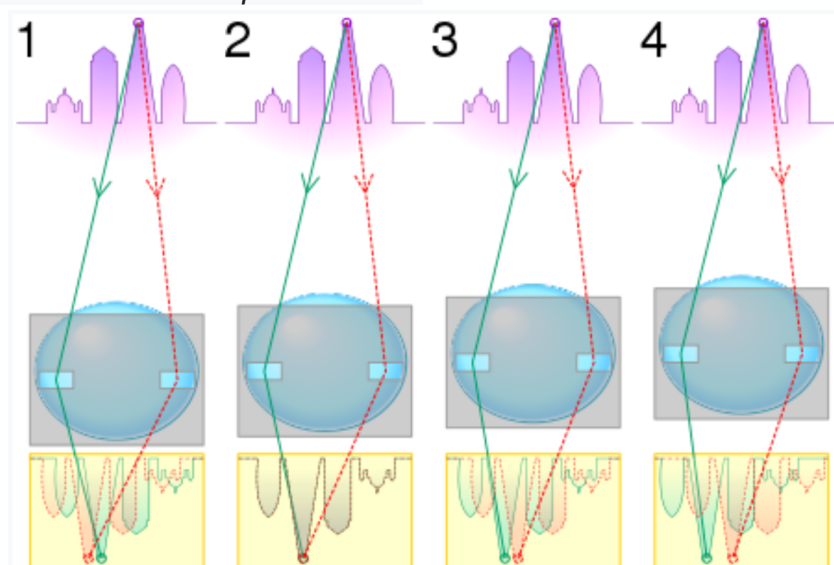
However, this approach requires continuously sampling, which would be affected a lot when the photo-shooting condition is not quite stable (e.g., fast moving object, not fixed enough place that the camera sits on, and so on).

The latest approach is mentioned in [2] (by a research team lead by Marc Levo at Google), where the authors build a customized algorithm based on the important mobile ISP function, which could continuously adjust the exposure time, gain, focus and white balance as the user aims the camera. The algorithm contains 3 steps:
1. deliberately underexposure: so that fewer pixels saturated,
2. capture multiple frames: to reduce noise in the shadows,
3. compress the dynamic range: using local tone mapping

The aforementioned ones are mostly parameter-free approach, however, I believe that we could adjust this problem by brand new approach: parametrize the input matrix (image) and map it into latent space, calculate the predicted distance (or, focus), then we tune the hyper-parameters until the the parameters suits the most of the scenarios [3]. Despite its heavy workload when tuning the hyper-parameters, it would suit most of the scene without having to worry about the potential defect in the closed form formulas like most of the approaches nowadays.

The main contribution of this proposed solution would be to assist the process of finding the focus much more intelligently, and thus improve both efficiency and efficacy of the resulting photo.

Methodology

1. Calculate the distance from the object of interest to the camera
2. Deduce the aforementioned distance via PD AF method, Burst Photography method, and the proposed one.
3. Compare the result (accuracy of the focus, latency when shooting photo) and evaluate the overall performance of these methods.

I first propose a model of parameters with randomly initialized weights, and tune the parameters via feeding the photos to it, and adjust the parameters accordingly through each epoch.
Next, re-implement the methods from the literature and use these to evaluate the resulting photo and index of interest (focus, or distance in this case)

Tasks

Shot (or process) the photos based on naive approach (PD AF), the proposed approach (mine), and the Burst photography approach (the Google's one)

Compare the extent of blur of processed photos of the raw photo, and the ones processed via the mentioned ones.

Build a customized camera that runs these algorithms to prove its accessibility to the general public if time permits.

The predicted result of the proposed approach would be better than PD AF, but not necessarily beats the burst photography approach by a huge amount,

Facilities

I have found an awesome project[1] where the author builds a customized camera via raspberry pi board only, which customized algorithms could be run on. If time permits, I will build my own camera based on this project, which is used to verify the experiment we mentioned above.
Price: <3000 NT$

Otherwise, I'll just run the algorithms based on photos from existing (public recognized) image datasets.

Reference

[1] https://github.com/penk/ruha.camera
[2] Burst photography for high dynamic range and low-light imaging on mobile cameras
[3] Anastasis Kratsios. Universal Approximation Theorems. ResearchGate. Published October 8, 2019. Accessed April 22, 2022.