Depth Estimation Using Segmentation in Natural Images

Shrayan Roy, Roll No.: MD2220

Project Supervisor: Dr. Deepayan Sarkar, ISI Delhi

Traditional photographs are 2-dimensional projections of a 3-dimensional scene. The third dimension is Depth, which represents the distance between the camera and the objects in the image. Estimating the depth of objects from a single photograph is an intriguing problem with many applications [1], such as post-capture image refocusing [2], but is difficult without special hardware.

Levin et al [3] proposed a simple depth estimation method that requires a modified camera with a special coded aperture. The premise of their method is that a conventional camera captures blurred versions of scene information away from the plane of focus, and by making use of a simple gradient prior on natural images, the amount of blur can estimated as a surrogate for depth. This method is used to estimate a layered depth map which can be used for automatic scene segmentation, post-capture refocusing of blurred areas, or re-rendering of the scene from an alternate viewpoint.

Our project aims to explore whether similar tasks can be performed using images from standard (unmodified) cameras, making use of modern high-performance segmentation algorithms such as Segment Anything [4]. Our hope is that by starting from a pre-segmented image, it will be easier to estimate the blur kernel without a special coded aperture, using the dependent image gradient prior proposed in [5]. In addition to the non-parametric blur kernel estimation method proposed, we plan to formulate and use simple parametric models for depth blurring, which we hope will lead to more efficient estimates.

References:

1. Alican Mertan et al. *"Single Image Depth Estimation : An Overview"*. In: arXiv:2104.06456 [cs.CV]. URL: https://arxiv.org/abs/2104.06456
2. Yosuke Bando & Tomoyuki Nishita. *"Towards Digital Refocusing from a Single Photograph"*. In: 15th Pacific Conference on Computer Graphics and Applications (Pacific Graphics 2007), pp. 363-372, 2007. URL: https://web.media.mit.edu/~bandy/refocus/
3. Anat Levin et al. *"Image and Depth from a Conventional Camera with a Coded Aperture"*. In: ACM Transactions on Graphics,Volume 26,Issue-3. URL: https://doi.org/10.1145/1276377.1276464.
4. Alexander Kirillov et al. *"Segment Anything"*. In: arXiv:2304.02643 [cs.CV]. URL: https://arxiv.org/abs/2304.02643
5. K.Nandy & D.Sarkar. *"Bayesian Image Reconstuction Using Locally Dependent Gradient Priors and Efficient Sparse Matrix Computations"*. URL: https://www.isid.ac.in/~deepayan/pubs/preprints/brecon.pdf