Using Light Fields to Enable Deep Monocular Depth Estimation

Niall Hunt, Master of Computer Science

University of Dublin, Trinity College, 2021

Supervisors: Dr. Martin Alain and Prof. Aljosa Smolic

Knowing the depth of a scene is a vital part of many computer vision applications, ranging from autonomous driving to augmented reality. Monocular depth estimation is the technique of predicting the depth of a scene given a single image. This ill-posed problem is increasingly being tackled with end-to-end neural networks thanks to advancements in machine learning. Deep learning techniques require large datasets with accurate ground truth depth maps for training. There are many features within an image that are used to decipher depth, these include: texture, object size and defocused blur. The large datasets that are provided for monocular depth estimation by nature do not include all of these features.

This project studies the state of the art in monocular depth estimation and their ability to generalise to an unseen dataset that includes images with features across all-in-focus images and images with a defocus blur. This dataset consists of various four-dimensional light fields which unlike traditional two-dimensional images, capture both the spatial intensity and angular direction of light rays. We use these light fields to create three distinct datasets (all-in-focus, front focus and back focus), using a shift-and-sum algorithm. These datasets are used to study and evaluate the performance of the chosen monocular depth estimation techniques.

In this dissertation we present the findings of the study, showing that despite the advancements made in monocular depth estimation, there is still a large gap in performance in comparison to other depth estimation techniques and that the performance of the state of the art cannot be predicted across the three datasets.

The code that accompanies this work is available at:
`https://github.com/NiallEHunt/MonocularDepth-Using-LightFields`