Using Neural Radiance Fields and Gaussian Splatting for 3D reconstruction of aircraft inspections

Roos Bottema

The field of computer vision is an area where multiple technological disciplines are combined. One such field is machine learning. Rapid recent development of machine learning techniques, have revolutionized the area of computer vision. One example is the rendering of novel views. Rendering novel views means that based on a set of images that have captured a scene creating the views of the scene that are not exactly captured by a camera. For example at an angle between two adjacent images.

This Master's thesis has focused on two different ways of novel view synthesis: one machine learning method called Neural Radiance Fields (NeRFs) and one optimization method called Gaussian Splatting. Whereas NeRFs consist of a machine learning architecture that is trained to represent a scene in a continuous manner, Gaussian Splatting consists of 3D geometric shapes that are optimized to represent the scene.

During this project, the area of application was aircraft inspection performed by autonomous drones. The drones fly a predetermined path around the aircraft and take images. Because of the localization feature of the drone, the exact location of the drone, and thus the camera, is known at each instance an image is taken. Recreating novel views can be a way of optimizing these types of inspections.

Neural Radiance Fields use a neural network architecture to create novel views from input images. The input to these networks are the images and the positions of the camera. The output of the network are color and opacity for each point in space. This data is then used to create new images from not before seen angles.

Gaussian splatting consists of 3D Gaussians, which can be seen as 3D ellipsoids that are defined by a mean (the centre) and a standard deviation in 3 directions (how much the Gaussian is spread out in space). During optimization, the Gaussians are initialized and moved around until they represent the input images in a correct manner. In order to produce novel views, these 3D ellipsoids are reprojected down to a 2D plane. Their colours are accumulated and rendered into an individual pixel.

In order to analyze the results of NeRF and Gaussian Splatting, three metrics as well as visual inspection were used during this project. This way, the output of different models could be compared.

The results of this project have shown promising results for NeRF when a notion of depth was added to the MLP. This was done by computing the distance from the drone to the aircraft based on the pre-computed path the drone would fly. Including more accurate data such as LiDAR data could have the potential of improving the results even further. Even Gaussian Splatting showed results that with further tuning of optimization parameters and testing different ways of initializing the Gaussians can lead to accurate novel view synthesis. However, none of the methods tested in this projected showed results that were detailed enough to be used during actual inspections.

As future recommendations it is highly encouraged to continue with depth-supervised NeRFs and to process the images further by masking backgrounds to only train on the aircraft. Moreover, it can be interesting to look into the use of vision transformers as they have shown great potential on various computer vision tasks.