Multi-view Camera synthesis using Convolutional Neural Network

Valeria Olyunina, Master of Science in Computer Science

University of Dublin, Trinity College, 2019

Supervisor: Matthew Moynihan, Prof Aljosa Smolic

This dissertation trained a neural network capable of producing an intermediate image between two spatially distributed images - effectively creating a novel point of view. This research is based on the article by Niklaus et al, 2017 "Video Frame Interpolation via Adaptive Separable Convolution" where a convolutional neural network is deployed to generate interpolated frames in a *video sequence*. The same approach is successfully applied in this research to *multi-view camera images*. The neural network is re-trained on a dataset of synthetically produced multi-view camera images. The resulting images are evaluated both for their quality in 2D and as a tool for improving the photogrammetery method of Shape-from-Silhouette in 3D reconstruction. The neural network trained on multi-view camera images produced by this research can generate visually correct interpolated multi-view images. When compared to ground truth, PSNR of these images is above 40 and SSIM is above 92% for the distance between multi-view cameras of less than 60cm (distance from camera to subject between 3-5m) when tested on a synthetic test set. This is higher than the corresponding results for the original video interpolation article. For 3D reconstruction, the cameras needed to be further apart (1-2m) and the silhouettes were not always pixel-accurate. Within 60cm only 1% of pixels were lost, however at distance between cameras over 1 m over 3% of pixels are lost, resulting in loss of voxels in extremities.