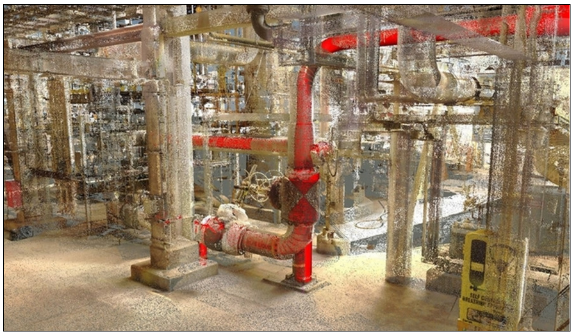

Simon
Walser

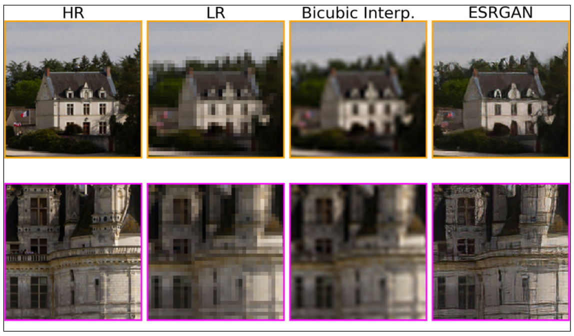
Student	Simon Walser
Examiner	Prof. Dr. Guido Schuster
Subject Area	Software and Systems

Single image super-resolution for point cloud colorization

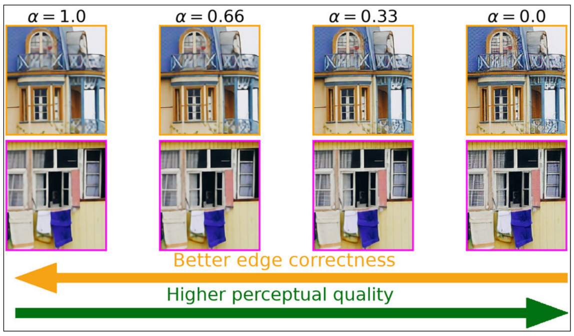
Image upsampling based on convolutional neural networks with emphasis on edge correctness



3D-model: It consists of distance data as a point cloud which has been colorized according to the color images.
Source: kleinfelder.com 3D Reality Capture Solutions



Comparison image versions: ESRGAN-image (right) is almost indistinguishable from HR-image (left). [4x upscaling]
Own presentment



Interpolated ESRGAN: A greater alpha results in better edge correctness, a smaller alpha results in higher quality.
Own presentment

Introduction: Due to the current development of ever cheaper computing power and constantly improving algorithms, expensive high quality sensors are replaced by cheaper sensors in combination with an enhancement algorithm. Images no longer need to be high-resolution and distortion-free if a piece of intelligent software can correct these shortcomings.

This also holds true for reality capture devices which use distance and image sensors to acquire a high-precision 3D model of a scene (see example on the left). For the 3D model to be accurate and consistent, the color images of the scene must have a sufficiently high resolution. Consequently, a single scan results in a considerable amount of data, which must be processed and stored for later use. Hence, it would be desirable if these color images were acquired and stored in a lower resolution and only needed to be enlarged for later processing. The aim of this thesis is to find a method to enhance lower resolution images without compromising geometric accuracy or image quality.

Approach: In the last ten years, a significant progress has been made in the field of image resolution enhancement, also referred to as "single image super-resolution". Modern super-resolution generators are based on fully convolutional neural networks (FCNN) which take a low-resolution image as input and output an enlarged and enhanced version by inserting details and textures that the generator has learned to be suitable (see example on the left). Since the geometrical correctness of the images is particularly important for colorization of high-precision point clouds, this thesis will pay special attention to this aspect of super-resolution.
This task is approached by modifying state-of-the-art super-resolution networks in such a way that the spatial edge correctness of the generated images can be ensured without compromising on the visual appearance.

Result: The edge correspondence and perceptual quality are the objectives that guided the search for a solution to the problem at hand. If the neural networks are optimized for edge correspondence only, the visual appearance of the image suffers and vice versa. This dilemma was overcome by optimizing two super-resolution networks, one for each sub-target.
In a next step, these two baseline models were combined and merged together with the aim of finding a good compromise between visual appearance and edge correspondence. Ultimately, this led to the main result of this thesis, the interpolated ESRGAN. This interpolated ESRGAN is obtained by a simple interpolation of the trainable weights of these two baseline models (see illustration on the left). This allows the two requirements to be mixed and matched at will.
Using the solution proposed in this thesis, the resolution is increased by a factor of four per spatial dimension without compromising on the quality of the 3D model.