

Posudek diplomové práce

Matematicko-fyzikální fakulta Univerzity Karlovy

Autor práce Jan Špaček

Název práce Generation of realistic skydome images

Rok odevzdání 2020

Studijní program Informatika **Studijní obor** Umělá inteligence

Autor posudku Alexander Wilkie **Role** vedoucí

Pracoviště CGG KSVI MFF UK

Text posudku:

The text of the thesis is well written and easy to comprehend. The section describing the method used is very detailed and includes snippets of code which contain important pieces of the puzzle that allowed the author to produce good-looking results.

One of the goals of the thesis was to generate images in a higher resolution than was possible before: low resolution output was a key disadvantage of the initial technique developed by Stepan Hojdar. The end result of this work is a model that is able to produce clean 512x512 skydome images in fisheye projection. The author also shows several examples from a run trained for 1024x1024 resolution and discusses the limitations of the method as presented. Instead of continuing to try to scale up the output resolution which runs into limitations of available GPU memory, the author proposes the idea of using patch-GAN super-resolution and presents preliminary results after applying this technique. Because of the lack of real data at this resolution, the author uses a combination of artificial loss functions which implement the upsampling in practice. The results are promising, but would clearly need more work to be able to scale all the way up to the resolutions required by CGI applications. So significant progress was made regarding the resolution issue, but the problem is not entirely solved yet.

Regarding the quality of the sky dome appearance that the technique is capable of generating, the resulting model also has some way to go before it can be used in a production renderer: the end result still exhibits artefacts, and is not robust enough yet. However, the work of the student presented in this thesis was still a significant step forward, as output quality was considerably improved compared to the previous implementation presented by Hojdar. While that work was technically able to produce vaguely similar images, they were of much poorer quality, and exhibited artefacts typical for badly chosen network architectures. In order to improve model performance, the author carried over most of the basic ideas from StyleGAN (which is an orthogonal research project at ČVUT that deals with style transfer, and which has a considerably more sophisticated

structure than the work of Hojdar) to this project, and was able to identify a subset of hyper-parameters for which the new system works particularly well. The author also showed that while the models trained by Hojdar mostly reproduced individual data points from the training set, the results of this new work lead to properly synthesised images that visually capture the essence of the original dataset, but which also exhibit nontrivial deviations from e.g. nearest neighbours. This is in itself a significant step forward.

A definite issue with the thesis is that the author could have presented more results of his improved method, to showcase that he did, in fact, manage to make his model perform a lot better than the first attempt. The author did lots of work, achieved tangible improvements over the state of the art, and then did not actually show much of this in practice (for whatever reason). In spite of this shortcoming, the thesis presents solid work, and I recommend it for acceptance.

Práci doporučuji k obhajobě.

Práci nenavrhuji na zvláštní ocenění.

V Praze dne 1. 7. 2020

Podpis: