

Bc. Michal Kučera Master Thesis Review

Qualitative Comparison of Methods for Example-based Style Transfer

By Ing. Michal Lukáč, PhD. – Adobe Research – lukac@adobe.com

The candidate was charged with a straightforward, yet non-trivial task of running a perceptual evaluation of a group of methods for example-based stylization. Example-based stylization is currently a hot topic in research circles, attracting commercial applications such as *Prisma* or *DeepArt.io* and consequently the attention of numerous research teams which try to tackle the problem both from the more traditional patch-based direction as well as from the perspective of currently booming neural generative algorithms. Many of these methods make claims to visual quality of their results in their respective papers, and this thesis puts those to the test.

In the thesis, the candidate limits the problem to stylization of human face portraits, justifying it both by limitations of some of the tested methods (FaceStyle and Selim et al. in particular being limited to this domain) and by the expected ease of evaluation on this domain by lay users. While a comparison of general-purpose stylization methods would be more beneficial, I consider this limitation well-justified and reasonable, as facial stylization is an active enough area of research in its own right.

Following the description of the methods and their underlying principles in Chapter 2, the candidate elaborates on the test design in Chapter 3, providing important information on how the method results used in the test were generated. There is a detailed discussion of parameter selection in calibration of the implemented methods, which contributes to transparency of the test design and shows that the author attempted to test each of the methods at its best. Based on my own experience, I must point out that the task of getting five unrelated methods running and producing their optimal results represents a substantial amount of work.

The author then designs and runs a two-stage test, first with a smaller user group to verify experiment design, and next with a large body of online responders. The primary questions asked in these tests are on how well each method reproduces the artistic style, and how well the method preserves the identity of the subject in the stylization.

Results of the testing are overwhelmingly conclusive; FaceStyle does consistently best in preserving the artistic style, with Selim et al leading the neural methods. In terms of identity preservation, the results are almost perfectly inversely correlated to style quality, which however seems to be more of an effect of how the responders interpreted the questions than actual identity recognition.

Overall, the thesis confirms what we have suspected in terms of stylization quality, backing it with hard numbers. The author compares the various methods side-to-side and does a good job of analyzing the differences detected in user testing. In light of this, I recommend the thesis for defence and grade it **A – excellent**.

I suggest the following question for discussion:

How would you propose to restructure the experiment to measure identity preservation in terms of recognizability of the subject?

In San Jose, CA, 14th June 2018

