TopoSketch: Drawing in Latent Space

Ian Loh* and **Tom White***
School of Design
Victoria University of Wellington
Wellington, New Zealand
{lohjun@myvuw.ac.nz, tom.white@vuw.ac.nz}

Abstract

We review a sketch based interface for prototyping animations from the latent space of generative neural networks. The capability of our tool to support visual exploration and communication is demonstrated within the context of facial images, though our framework is domain independent. The motion of sketched gestures is utilised as a means for generating animations, and serves as a reusable visual encapsulation of an animation's semantic operation.
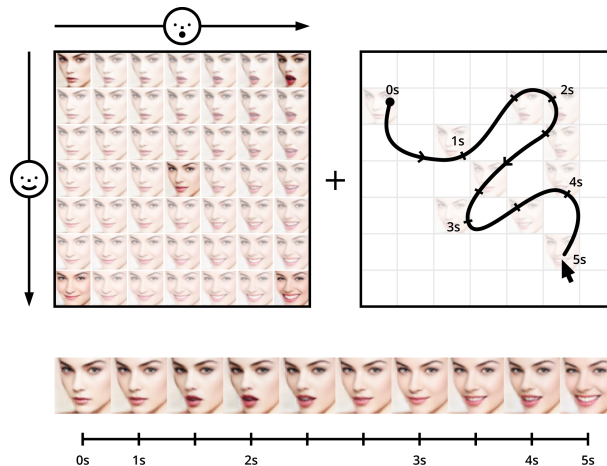
1 Introduction

Figure 1: TopoSketch animation process

In animation, gestural sketches may be used as visual guides to study and prototype movement. We developed TopoSketch, a tool for prototyping the animation of faces by sketching gestures using a tablet pen or mouse (Figure 1). Gestures generate a path through a latent vector space of faces that abstracts the complex task of posing faces into a simple animation control grid. TopoSketch animations are generated by reflexively exploring and charting a path through this pre-defined topographic latent space. Additionally the sketches become reusable components of the overall workflow. The sketch's appearance serves as a visual mnemonic to suggest the types of animations that will result when used as an operation in the latent space, and sketched gestures can be reused to achieve similar animation styles on different sets of faces.

---

*equal contribution

2 TopoSketch

TopoSketch is a sketch based facial animation tool that uses neural networks to navigate a plausible animation manifold. Posing and animating a believable face is a complex process, due to the interrelation of different facial features (eg: the eyes narrowing during a smile). While typical animation systems may enable individual adjustment of these facial features, TopoSketch utilities a high level control grid that enables changes within the context of operations such as 'smiling', or 'opening mouth'. These high dimension operations represent changes of many separate facial features simultaneously, allowing remapping of time-based gestures to facial animation.

The current TopoSketch prototype is based on a VAE model described in (Lamb 2016). Our model is initially trained in an unsupervised fashion on images from the CelebA training set (Liu 2015) resized to 256x256. After training, concept vectors are built using attributes from both the CelebA and the Radboud Faces Database (Langner 2010). We have extended the techniques of constructing concept vectors as described in (White 2016) by also using vectors orthogonal to an SVM hyperplane in latent space, which we have found gives superior results when given sufficient training data.

Extended information covering our current TopoSketch implementation including a working demo and videos are available online.[2]

2.1 Workspace

The TopoSketch workspace consists of two adjacent areas: the animation and preview window (Figure 2). An image is loaded into TopoSketch, processed into a grid of generated faces, and then displayed in the animation window. Each grid axis represents a different operation applied to the input face, increasing in effect along the axis. E.g. the mouth widens along the X axis while the smile gets larger along the Y axis, with each face on the grid representing a combination of the two. Animations are created by drawing gestures within the animation window, which are previewed in realtime by sampling the nearest grid face to the path. Playback is controlled with a timeline beneath both windows. Additional buttons provide functions for exporting, loading and editing animations or grids. Exported animations are rendered at higher temporal fidelity by decoding the final path through the latent space of the model.
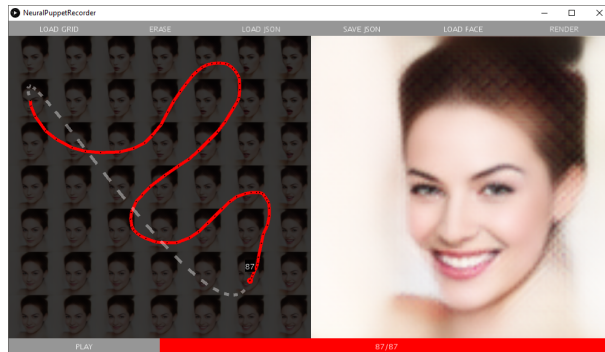

Figure 2: TopoSketch in use

2.2 Discussion

TopoSketch proposes a method of creating facial animations through a simple, high level, sketch based interface. Neural network generated latent spaces provide an underlying "intuition" which allows users to transfer expressive and reflexive qualities of inputs like sketched gestures to more complex animations. The combination of quick low precision gestures, simple representation, and low investment of face posing creates an environment that supports weak filters for quality, encouraging experimentation (Kim 2015). This approach to visual exporation is being refined based on TopoSketch user feedback. We are also working to generalize this style of tool with prototypes in other domains outside of facial animation.

---

[2]https://vusd.github.io/toposketch/

References

Kim, J.; Bagla, A.; and Bernstein, M. S. (2015) Designing creativity support tools for failure. *Proceedings of the 2015 ACM SIGCHI Conference on Creativity and Cognition*, pp. 157–160. New York, NY, USA: ACM.

Lamb, A.; Dumoulin, V.; and Courville, A. (2016) Discriminative Regularization for Generative Models. *ArXiv e-print* https://arxiv.org/abs/1602.03220

Langner, O.; Dotsch, R.; Bijlstra, G.; Wigboldus, D.; Hawk, S.; and van Knippenberg, A. (2010) Presentation and validation of the radboud faces database. *Cognition and Emotion*

Liu, Z.; Luo, P.; Wang, X.; and Tang, X. (2015) Deep learning face attributes in the wild. it Proceedings of International Conference on Computer Vision (ICCV)

White, T. (2016) Sampling Generative Networks. *ArXiv e-print* https://arxiv.org/abs/1609.04468
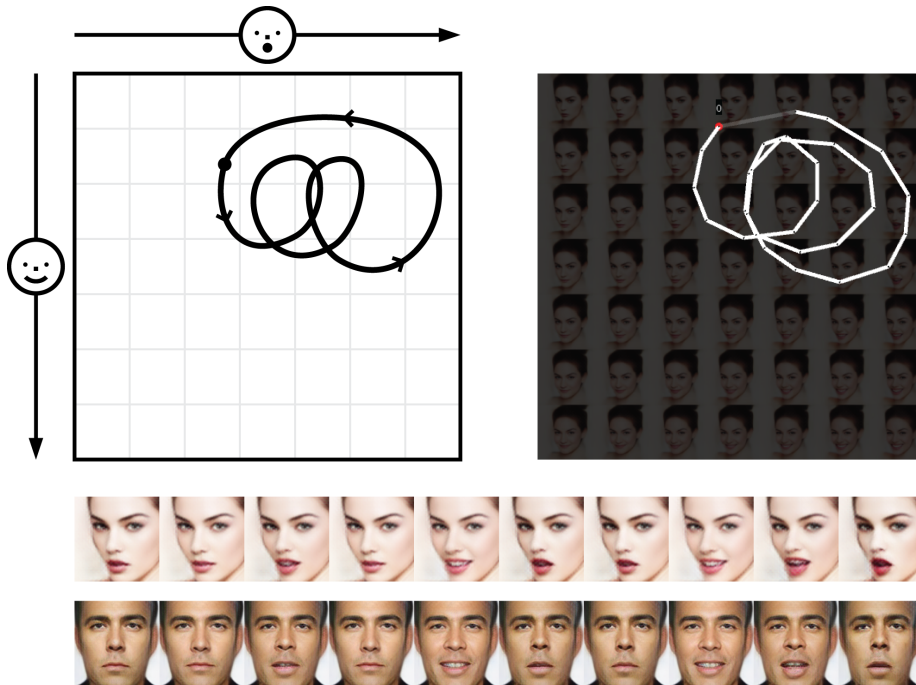
Appendix: Example Animations

Figure 3: Chewing Animation. The loops in the sketched sequence indicate repeated motions in the animation. The sketch gesture can also be transferred to a second input image without modifying the path.
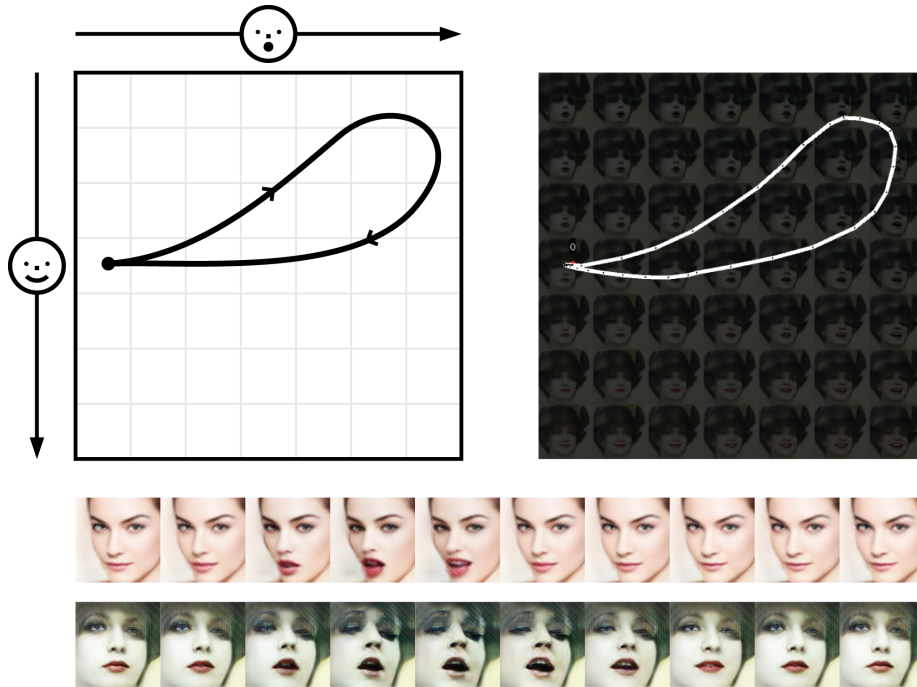
Figure 4: Kissing Animation shows how a sketched sequence is reused on a second input image. The captured sketch is a reusable component that can be applied to other inputs.
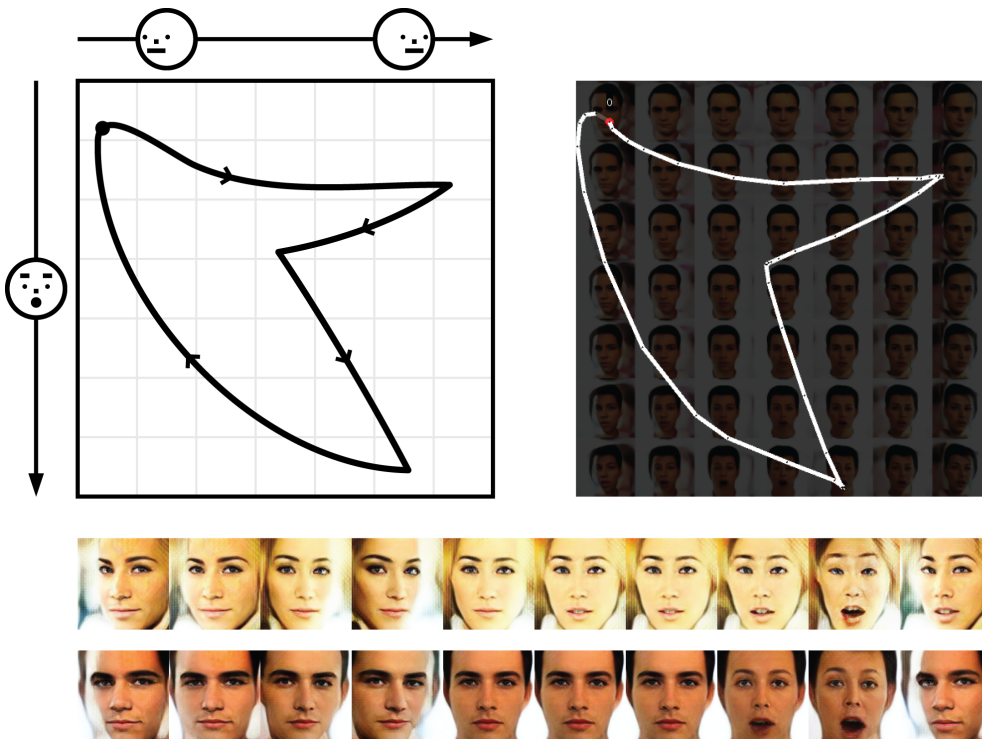

Figure 5: Parameters for animations can combine facial expressions with other changes such as orientation and lighting. Here a "double take" animation is constructed from face rotation and an expression of surprise.