Personalized Image Generation Through Swiping

Yuto Nakashima

The University of Tokyo
yuto-nakashima@g.ecc.u-tokyo.ac.jp

Abstract

Generating preferred images from GANs is a challenging task due to the high-dimensional nature of latent space. In this study, we propose a novel approach that uses simple user-swipe interactions to generate preferred images from users. To effectively explore the latent space with only swipe interactions, we apply principal component analysis to the latent space of StyleGAN, creating meaningful subspaces. Additionally, we use a multi-armed bandit algorithm to decide which dimensions to explore, focusing on the user's preferences. Our experiments show that our method is more efficient in generating preferred images than the baseline.

Introduction

GANs have been shown to generate more realistic images than conventional generative models. By manipulating the latent variables of GANs, users can generate their preferred images. However, GANs' latent space is often high-dimensional and can be challenging to explore. Existing research generates preferred images by asking users to adjust multiple sliders and edit images (Chong et al. 2021).

In this study, we propose a more lightweight method to generate preferred images by using a simple swipe operation. An overview of our approach is illustrated in Figure 1. When the system displays an image, users swipe to indicate their preference. Based on this feedback, the system explores latent variables in the latent space and generates new images that better match the user's preferences. With repeated interactions, the system becomes more adept at producing images aligned with user likes and dislikes.

To generate preferred images using swipe interactions, we need methods that operate on more limited feedback compared to previous approaches and that enhance the efficiency of exploring latent variables. To effectively explore the latent space with only swipe interactions, we apply principal component analysis to the latent space of StyleGAN, creating meaningful subspaces. Additionally, we use a multi-armed bandit algorithm to decide which dimensions to explore, focusing on the user's preferences.

To assess the efficiency of our proposed method in generating user-preferred images, we conducted user experiments. Our experiments show that our method is more efficient in generating preferred images than the baseline.

Swipe Interactions for Image Generation

We introduce a new user interface, illustrated in Figure 1, designed to generate user-preferred images using simple swipe actions. This interface captures user preferences by continuously processing swipe interactions to generate the preferred images.

Users are asked to perform pairwise comparison by comparing the currently displayed image with the previous one: swiping right indicates a preference for the current image, while swiping left favors the earlier one. Depending on this feedback, a subsequent image is generated and presented. This process continues until an image aligns with the user's preference.

Proposed Method for Efficient Latent Space Search

One issue with pairwise comparisons is that the number of inquiries to humans increases as a result of the small amount of user preference information obtained from each iteration. Therefore, it is necessary to devise more efficient ways to explore latent variables. Recently, with the advent of GANSpace (Härkönen et al. 2020), it has enabled us to acquire directions that significantly change the appearance of images. Taking inspiration from this, we apply PCA to the latent space of StyleGAN and performed exploration in a subspace composed of only principal components to streamline the search for latent variables.

Within this reduced subspace, it is unclear which dimensions should be prioritized for exploration from the user's perspective. Our study addresses this by determining which dimensions to explore within the subspace using a multi-armed bandit algorithm. This approach enables us to select the most informative dimension from among the dimensions with diverse characteristics of the subspace and explore its surroundings.

Experiment

To verify whether the proposed method can generate preference images more efficiently than the baseline when using

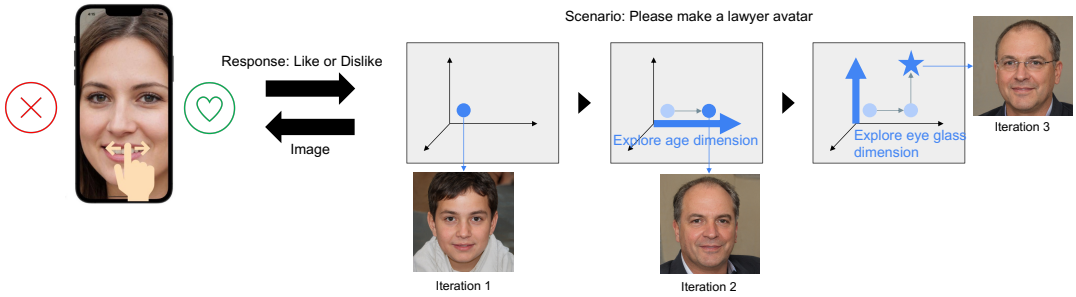

Figure 1: Overview of image generation through swipes: This example illustrates how our system generates a lawyer avatar based on user swipes. By interpreting continuous swipe feedback, the system dynamically adjusts the displayed image to match user preferences. Internally, it evaluates dimensions such as age or the presence of glasses. For instance, in the second iteration, the system updates the avatar's age based on estimated user preferences. By the third iteration, in response to feedback, glasses are added to the avatar.

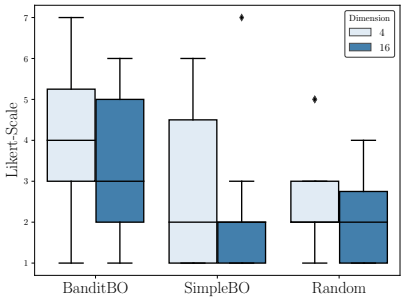

Figure 2: Distribution of response to "Were you able to reach the preferred image efficiently?"

actual user evaluations, a user study was conducted. Participants were asked to generate avatars according to specified scenarios. For each scenario, three methods were randomly assigned: the proposed method (BanditBO), a method that only performs Bayesian optimization in a subspace (SimpleBO), and a method that randomly samples latent variables from a uniform distribution across the search space (Random). The combinations were assigned with two subspace dimensions, $d' = 4, 16$. After completing each scenario, a survey was conducted on the image presentation and comparison process and satisfaction with the final avatar.

Figure 2 shows the results of the survey question, "Were you able to reach the preferred image efficiently?" It can be seen that more people responded that they were able to reach the preferred image more efficiently with the proposed method compared to the baseline. This indicates that the proposed method can generate preference images more efficiently than the baseline.

Social and Individual Well-being

The intuitive design of our system significantly decreases the cognitive load and frustration that are often linked to complex creative tools. By allowing users to easily explore and create personalized visual content with simple gestures, such as swipes, the system boosts personal satisfaction. Regarding social well-being, employing this system in applications like generating composite sketches for criminal investigations can have a beneficial impact on societal welfare.

Conclusion

In this study, we examined whether the proposed method can generate preference images more efficiently than the baseline, using only pairwise comparison results. By applying principal component analysis to the latent space of StyleGAN to reduce dimensions and focusing on exploring the user's interest dimensions, it was demonstrated that preference images can be generated more efficiently than the baseline. This study enabled the proposal of a new approach to generate preference images on smartphones using swipe interactions. In the future, we will work on developing efficient search methods that can generate preference images with fewer iterations, even when the search dimensions are large. Specifically, we aim to reduce the number of iterations required to generate preference images by utilizing information obtained from user behavior, such as the time elapsed from image presentation to swiping.

References

Chong, T.; Shen, I.-C.; Sato, I.; and Igarashi, T. 2021. Interactive Optimization of Generative Image Modelling using Sequential Subspace Search and Content-based Guidance. *Computer Graphics Forum*, 40(1): 279–292.

Härkönen, E.; Hertzmann, A.; Lehtinen, J.; and Paris, S. 2020. GANSpace: Discovering Interpretable GAN Controls. In *Proceedings of the Advances in Neural Information Processing Systems*, volume 33, 9841–9850.