Weaving ML with Human Aesthetic Assessments to Augment Design Space Exploration

Youngseung Jeon[†‡], **Matthew K. Hong**[‡], **Yan-Ying Chen**[‡], **Kalani Murakami**[‡]
Jonathan Li[‡], **Xiang Anthony Chen**[†], **Matthew Klenk**[‡]

[†]University of California, Los Angeles, Los Angeles, CA 90095
[‡]Toyota Research Institute, Los Altos, CA 94022

[†]{ysj,xac}@ucla.edu [‡]{matt.hong,yan-ying.chen,kalani.murakami,jonathan.li,matt.klenk}@tri.global

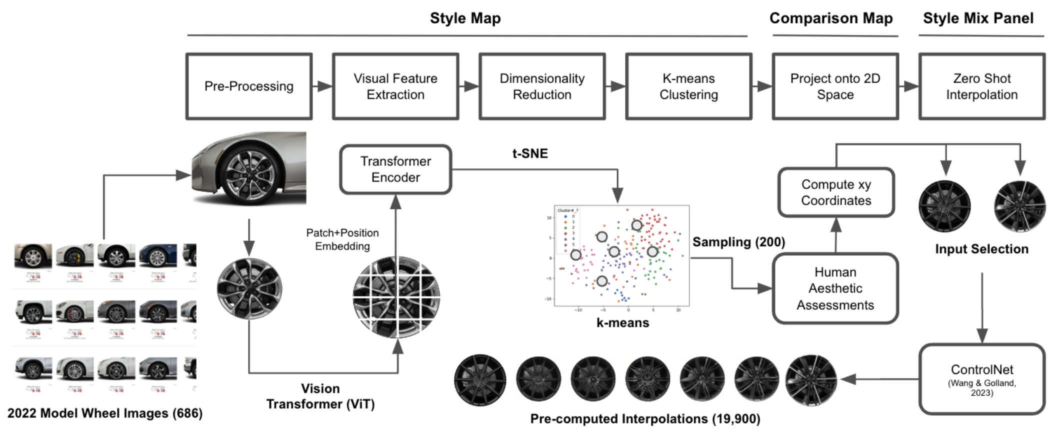

Figure 1: ML Pipeline for Design Space Organization System.

1 Introduction

People's semantic connection with product design is an important signal that drives purchase decisions and overall satisfaction. This relationship can be characterized by a set of semantic aesthetic attributes (e.g., rugged, sporty, luxury) assigned to perceptual features of a particular design. In the concept design phase, capturing and responding to this signal is an important part of a product designer's job. Yet in automotive design, where online experimentation is not a viable option, it is driven by speculation about consumers' aesthetic preferences, drawing from designers' intuition, prior experience, and domain knowledge. In this process, designers search and curate a set of potential design examples from experience and online repositories such as a product catalog or Pinterest in response to a design prompt (e.g., "*a rugged, adventure seeking vehicle*"), to establish a design direction and inspire generation of new design alternatives that best fit the design prompt.

We identify two problems that exist in this process to address through a novel framework that combines machine learning (ML), human assessments, and interface design. First, it does not consider a systematic exploration and investigation of design choices with respect to different semantic attributes that represent perspectives of existing and prospective consumers [3]. Additionally, consulting direct design feedback in later production phases is too costly[1], often resulting in additional design iteration cycles that cause potential delays in productization.

Prior work in ML has demonstrated feasibility in combining text-image joint embeddings and image generation models to achieve sufficient approximation of visual style based on a given concept keyword [4]. While enticing, current generative systems have limited applicability in professional

[1]In the automotive industry, targeted consumers are recruited to make aesthetic assessments in design clinic sessions based on descriptors such as sporty and luxurious. A single design clinic for a new vehicle design could cost over $100,000, with annual expenditures estimated at tens of millions of dollars for a single manufacturer[1].

37th Conference on Neural Information Processing Systems (NeurIPS 2023).

practice for two reasons: (1) designers lack means to anticipate outcomes based on text input, and (2) it is largely unknown whose aesthetic assessments are represented in the design outcomes.

In our work, we aim to reduce the psychological distance between designers and consumers and increase designers' capacity to make sense of a large design space. To this end, we introduce a creativity support tool (CST) that incorporates ML techniques to support designers in organizing a large space of automotive wheel designs informed by human aesthetic assessments, enabling flexible and creative modes of thinking about novel design spaces. In Appendix A, we report on three use cases that demonstrate how this CST can facilitate (1) rapid ideation and sketching, (2) finding happy surprises in form studies, and (3) communication in automotive design practice.

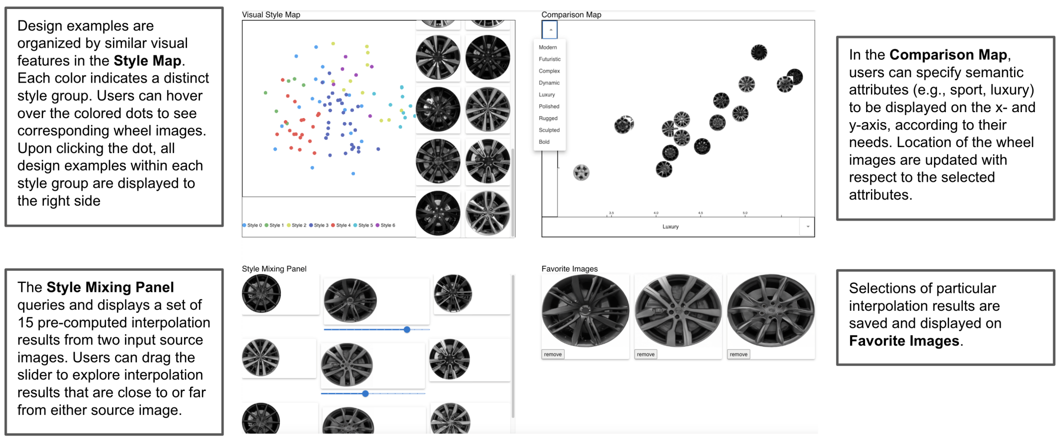

Figure 2: Design Space Organization User Interface.

2 System Design

Design Space Organization (DSO) is an AI-based CST that allows designers to explore a large number of design examples and alternatives. The system provides important information for designers during the wheel styling process, including visualization of style information organized by visual features and representation of human aesthetic assessments of design examples. Using latent diffusion models, designers can build on this visual information to experiment with novel visual concepts by interpolating between two wheel designs. In effect, the system aids designers in generating new alternative designs aligned with their intended design direction.

DSO consists of three main features (Fig. 2). **Style map** provides a large number of design samples (n=200) in a two-dimensional cluster map organized by visual features. These visual features are extracted via a vision transformer model (ViT) [2], projected onto a 2D space using t-SNE [5] dimensionality reduction, and subsequently clustered based on k-means algorithm. Each cluster group corresponds to a distinct style group within the Style map. Style selections are displayed on a 2D spatial layout, or **Comparison map**, based on quantitative human ratings (on a 7-point Likert scale) of images in association with 9 semantic attributes. Placement of images changes according to the chosen semantic attribute in x/y axis and corresponding values. Selections made on the comparison map are displayed on the **Style mixing panel**, where users can perform visual style mixing experiments using two selected input images. We use an implementation of ControlNet that enables zero-shot controllable interpolation [6] using latent diffusion models to support designers in generating and exploring multiple design alternatives in a narrow design space to inspire their work. Results of interpolation can be used as input to generate diverse combinations of images for further experimentation and exploration.

3 Conclusion

We developed a novel framework and system that combines machine learning, human aesthetic assessments, and interface design. We hope our demo can stimulate discussions around using this framework for professional product design practice.

A Case Study: Potential Use Cases of the DSO System

In an ongoing pilot evaluation with 10 automotive design professionals and students enrolled in automotive design schools, we uncovered insights that highlight potential use cases of this CST. We summarize three use cases below.

Rapid ideation and sketching. Many participants in our study expressed excitement over the tool's ability to support rapid ideation and sketching. For example, a junior transportation designer remarked,

> *I saw myself emotionally reacting to a wheel more quickly than sketching on my own—sometimes my hand is not putting down exactly what the brain is thinking.*

A 20 year veteran in automotive design added that the tool allows for quick benchmarking experiments with a diverse range of designs and even suggested a further use case to sketch on top of resulting inspirational images, using them as an underlay:

> *You can sketch over a Lexus, Cadillac.*

Using an inspirational source image as an underlay for sketching is a common technique that is used among industrial designers—to carry out visual experiments by adding lines that slightly deviate from the source design. A future iteration of this tool could include support for this extended use case by providing various levels of abstractions to the inspirational image, such as a outline sketches, within constraints specified by the user.

Figure 3: Comparison of interpolation results. Interpolating between completely different wheel styles resulted in poorer quality (Top), whereas compatible designs tend to yield higher quality results (Bottom).

Happy surprises in form studies. In industrial design, form studies consist of sequential, comprehensive manipulations of a geometric volume. Many participants' in our study saw the value of this tool in assisting their study of different forms in a wheel style. While designers saw the benefits of seeing unexpected AI artifacts, calling it a *happy surprise*, others also expressed disappointment. A junior professional automotive designer still appreciated and found meaning in the imperfect interpolation results:

> *However, there are also parts that can be utilized even if the overall lines are distorted. Even small changes that come from understandable lines can be very interesting and helpful to support my creativity.*

These reactions could be attributed to designers' unrealistic expectations of AI as well as the overall lower interpolation quality between designs that are incompatible (Fig. 3). Future work should explore techniques to improve the quality of interpolations across diverse styles. We speculate that removing noise (e.g., rotor, brake caliper) and controlling for varying image conditions (e.g., exposure, color contrast) in image pre-processing stage could help reduce unwanted distortion.

Visual communication. Designers also found immense value in using concrete design examples and aesthetic assessments as means to support design communication with management and colleagues. A second year MFA transportation design student commented on this topic, *It would also be very helpful in designer-to-designer communication. Most discussions about design are highly abstract, but if we can quickly visualize what we want, it could save us a lot of time.*

A professional automotive designer remarked on the value of aligning subjective interpretations to remove unnecessary inefficiencies in communication:

We often spend more time aligning our thoughts than we realize, and this feature could be applied to all design tasks [...] to support clear communication between design managers and designers.

References

[1] Alex Burnap, John R Hauser, and Artem Timoshenko. Product aesthetic design: A machine learning augmentation. *Marketing Science*, 2023.

[2] Alexey Dosovitskiy, Lucas Beyer, Alexander Kolesnikov, Dirk Weissenborn, Xiaohua Zhai, Thomas Unterthiner, Mostafa Dehghani, Matthias Minderer, Georg Heigold, Sylvain Gelly, et al. An image is worth 16x16 words: Transformers for image recognition at scale. *arXiv preprint arXiv:2010.11929*, 2020.

[3] Shang H Hsu, Ming C Chuang, and Chien C Chang. A semantic differential study of designers' and users' product form perception. *International journal of industrial ergonomics*, 25(4):375–391, 2000.

[4] Or Patashnik, Zongze Wu, Eli Shechtman, Daniel Cohen-Or, and Dani Lischinski. Styleclip: Text-driven manipulation of stylegan imagery. In *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pages 2085–2094, 2021.

[5] Laurens Van der Maaten and Geoffrey Hinton. Visualizing data using t-sne. *Journal of machine learning research*, 9(11), 2008.

[6] Clinton Wang and Polina Golland. Interpolating between images with diffusion models. 2023.