Kansei Engineering, Vol. 101, 2023, 23–33
https://doi.org/10.54941/ahfe1002988

AI Image Generation Boosts Kansei Engineering Design Process

Shigekazu Ishihara[1], Rueikai Kuo[2], and Keiko Ishihara[3]

[1]Department of Psychology, Hiroshima International University, Higashi-Hiroshima, 7392695, Japan
[2]Animen JP / VOLOR Inc., Taichung, Taiwan
[3]Department of Rehabilitation, Hiroshima International University, Higashi-Hiroshima, 7392695, Japan

ABSTRACT

This study delineates the integration of precisely tuned artificial intelligence (AI) image generators within the context of the Kansei engineering design process. Implementing Kansei engineering invariably encounters certain prevalent hurdles, including a limited diversity in evaluation samples, the exigency to incorporate non-designers within the design process, and the often overwhelming time commitments of designers that preclude their full engagement. To overcome these issues, the AI image generator known as StableDiffusion (Rombach et al. 2021) was introduced into the Kansei engineering design process. In an effort to more accurately depict the form and aesthetic of milk cartons, the StableDiffusion was adjusted using the "Hypernetworks" framework. By invoking Kansei design principles derived from a prior study of milk cartons (Ishihara et al., 1996), several attempts were undertaken to create innovative and aesthetically pleasing milk carton designs using AI-generated concepts. These designs involved the use of red and blue hues, abstract shapes, and cartoon-like illustration of meadow and cows. The implementation of AI technology demonstrates considerable potential as a powerful catalyst in fostering innovation within the Kansei engineering design process.

Keywords: Kansei engineering, Artificial intelligence, Generative image, Stable diffusion

INTRODUCTION

Common problems for implementing Kansei engineering to design process in companies.

Kansei engineering methodologies underpin the design process by scrutinizing the latent Kansei of existing and potential users, thereby infusing these insights into the development and continual refinement of products and services (Nagamachi, 1991, 2011). Through the implementation of Kansei Engineering in product development, a number of prevalent impediments have emerged.

A notable constraint is the dearth of diversity within evaluation samples. Upon orchestrating a Kansei evaluation experiment, the conspicuous lack of product diversity on the market becomes apparent. This hindrance is attributable to economic variables, the spectrum of product prices, and the inherent costs of materials and production. Consequently, innovation necessitates a departure from these real-world restrictions, and this departure is only rendered feasible by the wisdom and knowledge gleaned from various experts in design, production, materials, and procurement. Accordingly, 2D computer graphics techniques and 3D Computer Graphics (CG) have been systematically deployed to construct 'virtual products' for Kansei evaluation (Jindo et al., 1995, Matsubara et al., 2010). Virtual reality has also been introduced at an early stage (Nomura et al., 1998). Despite the significant strides made in CG techniques over the past two decades, object modeling and surface processing (including texturing and setting physical properties) continue to pose challenges for non-experts.

The ethos of Kansei engineering propounds the inclusive participation in the design process. Ishihara (2005) has empirically evidenced the catalytic effect of the Kansei Engineering system throughout the design process in international online design sessions spanning Sweden and Japan. Notably, Kansei engineering spurred the amplification of human creativity within the design and decision-making process.

Ishihara et al. (2010) presented a case study of rapid electric shaver design, deploying Kansei engineering and ergonomics at SANYO Electronics. Inc. In this instance, a diverse team, including the ergonomist, product manager, mechanical and electronic engineers, and the production planner, shared the latest findings from the Kansei ergonomic study. From the inception of an innovative design idea to its market introduction, less than six months elapsed. This brisk, agile, and novel design process necessitates the broad sharing of the most recent findings across almost all corporate departments. As such, it is imperative to articulate the product idea and the resultant analysis in near-comprehensive graphics.

Despite these advancements, there remains a shortage of designers, who are often overextended. In successful Kansei engineering applications, designers play a critical role from the nascent to the final stages of the design process. For example, significant effort is demanded in designing virtual products for Kansei evaluation experiments. However, this effort should be minimized to allocate more time for innovative ideation.

To attenuate these challenges, this study has integrated AI image generators, an emergent innovation in artificial intelligence, into the Kansei engineering design process. These tools could potentially facilitate the creation of evaluation samples and inspire design ideas among diverse stakeholders involved in product development.

Latent Diffusion Model

The image synthesis and generation AI system used in this research is called StableDiffusion, formally known as the Latent Diffusion Model (LDM) (Rombach et al. 2021). The structure is shown in Figure 1. The explanation of LDM in the following lines refers to Tokunaga (2022) and Shirai (2022). LDM operates with a combination of three distinct building blocks:

1. Diffusion model: U-net
2. Variational Autoencoder: VAE
3. Text Encoder: CLIP Coding

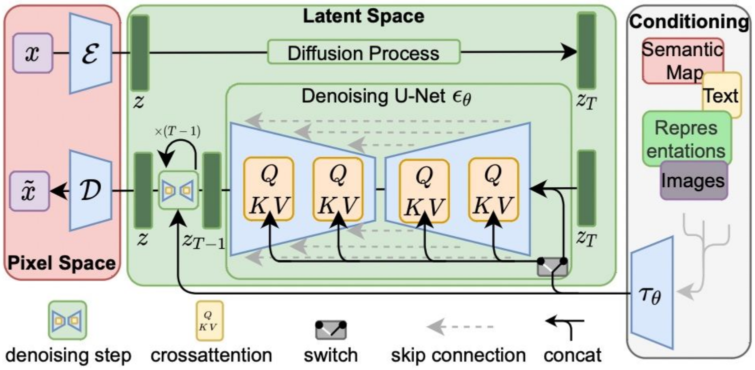

Figure 1: Structure of stable diffusion (referring from Rombach et al. 2021).

Stable Diffusion is a diffusion model trained on the latent space of a VAE for efficient high-resolution image generation. The text encoder is trained using CLIP coding (Radford, 2021), with text conditioning performed by cross-attention in U-Net.

The Stable Diffusion learning process takes an image and text pair as input. The image is embedded by the VAE encoder, and the text by the text encoder. Once embedded in the latent space, the image undergoes the diffusion and inverse diffusion processes of the diffusion model (Sohl-Dickstein et al., 2015). In the diffusion process, small amounts of Gaussian noise are continually added until the latent representation of the image becomes noise. The inverse diffusion process is the reverse; it gradually removes Gaussian noise from the noise. This removal process is trained on the U-net, with the final output of the U-net passed through the VAE decoder D to recover the image.

During generation, the only input is the text prompt. This text is fed into the U-Net during the inverse diffusion process, where the U-Net gradually removes noise from random Gaussian noise. As the U-Net is conditioned by the input text, the latent representation of the image will follow the text. The final output of the U-Net is passed through the VAE decoder to produce an image that aligns with the text. A trained Stable Diffusion model is available on Hugging Face and can be easily used with the U-Net implementation, "Diffusers."

A 64x64 Gaussian noise seed (Latent Seed) is provided to guide the noise reduction and resolution process so that user-provided prompts become latent variables in the text embedding (77x768 dimensions) generated by CLIP. CLIP (Radford, 2021) correlates the learning between embedded 768-dimensional vectors of words (sentence) and 77-dimensional vectors of the flattened 2D image vectors. An image is generated from the noise and super-resolved to a 512x512 image while aligning it to the vector.

StableDiffusion consists of an 860 MB Unet and a 123 MB TextEncoder.

Milk Carton Kansei Engineering Study

One of the authors conducted a Kansei engineering study on milk cartons (Ishihara et al., 1996). For the Kansei evaluation, seventy-one milk cartons were used. A 5-point semantic differential scale questionnaire was composed using sixty-nine Kansei words.

The Kansei evaluation data were analyzed with our self-organizing hierarchical clustering neural network, arboART (Ishihara, 1995). Figure 2 displays the clustering results. The nested enclosures illustrate the relationships between the inclusions of the clusters. These enclosures are

Clusters of all samples

1st cluster: Prototype vector: “simple”, “Proper” and “Monotonous”: Abstract shape and sans-self font

2nd cluster: Prototype vector: “Flesh”, “Lively”, “Natural” and “Healthy”: Photo or realistic illustrations

3rd cluster: Prototype vector: “Juvenile”, “Tender”, “Young” and “Bright” : Cartoon-like illustration of meadow and cows and Large white area

Figure 2: Milk carton Kansei analysis with arboART (Ishihara, 1996).

Figure 3: StableDiffusion 1.4 generated best pictures from "milkcarton" prompt.

intended to resemble contour lines on a map. Boxes encased by five lines form the first layer of arboART and exhibit very similar values. Clusters surrounded by fewer enclosures are grouped under a larger distance criterion. In Figure 3, there are three large clusters, two small clusters, and four independent samples are depicted on the right.

Figure 3 shows the clusters and the larger prototype values of the three main clusters. These results are utilized as cues to initiate the transfer of knowledge in the trained StableDiffusion.

METHODS

Initially, we tested StableDiffusion 1.4 and 2.0 using "milk carton" or "milkcarton" as the prompt. Figure 3 displays the best out of the 12 images generated by SD 1.4. While SD1.4 managed to capture some of the milk carton's features, such as the box and roof structure, it didn't accurately represent the milk carton's shape. Notably, the top roof structure differed from the actual carton. Figure 4 displays the best of the 5 images generated by SD 2.0. SD 2.0 seemed to struggle with the milk carton's shape, confusing it with other types of packaging like flour packets. The second image, despite featuring a red animal on the package, does not depict a cow but appears to be an Asian serow instead.

Based on these observations, we concluded that both versions of SD lacked accurate knowledge of the milk carton's shape. Given that SD 2.0 seemed to

Figure 4: StableDiffusion 2.0 generated best pictures from "milkcarton" prompt.

confuse the milk carton with other packages, we decided to train SD 1.4 to recognize the exact shape of a milk carton.

Fine-Tuning With Hypernetworks

The practice of fine-tuning through additive learning was explored in this study, where images of a white milk carton were captured from eight distinct orientations, each involving a rotation of 15 degrees. The creation of the pre-built model and its renderings were facilitated through Adobe Dimension version 3.4.6, as illustrated in Figure 5. Given the multi-angled presentation of the image to SD 1.4, there appears to be an establishment of implicit associations between the image and the sides.

The execution of the fine-tuning process on SD 1.4 was accomplished using stable-diffusion-webui, a tool authored by AUTOMATIC1111. The computations were performed on the Windows variant of this software, harnessing the computational power of an NVIDIA GTX 3060 GPU, which boasts 12GB of video memory. The system integrated the use of “Hypernetworks”, a fine-tuning framework proposed by NovelAI (NovelAI, 2022). The process of hypernetwork learning spanned across 4,000 iterations, where the shape learned demonstrated instability at the 500 and 1,000 cycle marks. Post 2,000 iterations, the learned shape exhibited a stable
pattern.

In contrast to traditional fine-tuning methodologies that frequently involve appending additional layers at the terminal end of a deep neural network, NovelAI (2022) posits that such a layer addition approach necessitates

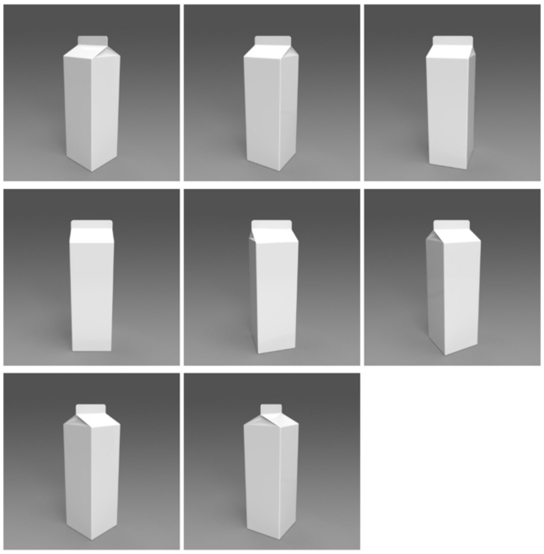

Figure 5: White milkcarton images rendered from 8 directions, used for fine tuning of StableDiffusion 1.4.

substantial computational resources. As a solution, they introduce the hypernetwork, a pioneering fine-tuning framework.

The hypernetwork strategy involves the incorporation of a singular compact layer into the internal layers of Key and Value vectors of the cross-attentional U-net within StableDiffusion. The remainder of the U-network components were left unaltered. This modification significantly mitigated the occurrence of overfitting, thereby promoting enhanced generalization at the culmination of the training phase.

RESULTS

It has been previously observed that design variations in available market products are somewhat restricted. Specifically, regarding the milk carton Kansei study it was noted that only two variants utilized a red color scheme (Ishihara et al., 1996). A commonly employed strategy for generating innovative ideas involves borrowing elements from related areas, in addition to adhering to traditional design conventions.

In a prior Beer Can Kansei study (Ishihara, 1998), we discerned a potent association of the color red with the following attributes: Premium, Gorgeous, Affected, and Showy. Given this observation, we endeavored to implement a non-conventional red hue to the milk carton design. As exemplified in Figure 6, this novel red color was effectively applied to the 3D model.

In an innovative maneuver, floral prints, infrequently seen on milk cartons yet common on candy packages, were incorporated. Figure 7 showcases the resultant image of the "red flower milk carton," striking a balance between elegance and tradition.

In previous studies, we found blue to be a favored color for milk cartons. Blue abstract shapes were recurrent, thereby leading to a Kansei perception of "simplicity," "propriety," and "monotony." In an effort to modernize this perception, we aimed to add a "refined" touch to this recurring blue and abstract design, as revealed in Figure 8.

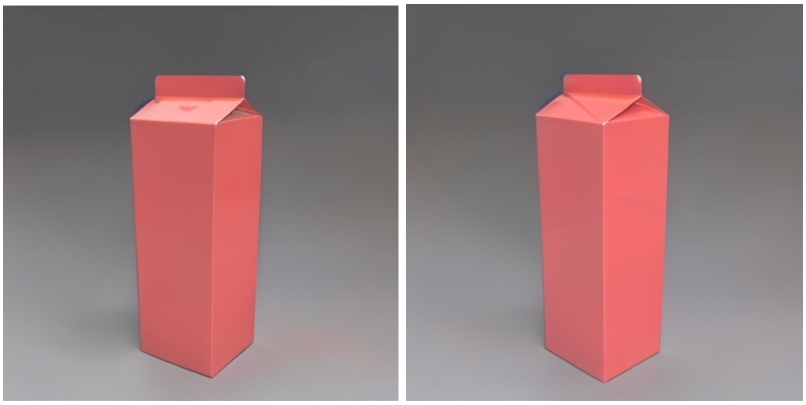

Figure 6: Generated from the prompt "red milkcarton", from the StableDiffusion learned milkcarton shape.

Figure 7: Generated from the prompt "red flower milkcarton", from learned StableDiffusion.

Figure 8: Generated from the prompt "blue abstract painting modern milkcarton", from learned StableDiffusion.

Figure 9 unveils an even more inventive concept, conceived from the prompt "colorful painting modern milk carton." The design of milk cartons with multiple colors was originally geared towards evoking a "Juvenile" and "Tender" Kansei. Our intention, however, was to infuse a touch of modernity.

Figure 10 shows the generated cartons corresponding to the third cluster of milk carton Kansei study (as shown in Fig. 2). Illustration or "Cartoon-like illustrations" of meadows and cows are used for the prompt.

Generated illustrations have another attractiveness than Japanese ones. The third illustration is an innovative abstraction of drawings.

Figure 9: Generated from the prompt "colorful painting modern milkcarton", from learned StableDiffusion.

Figure 10: Generated from the prompt "Cartoon-like illustration of meadow and cows on milkcarton", from learned StableDiffusion.

CONCLUSION

The KE methodology, through its dedicated pursuit of Kansei engineering, underscores its invigorating influence on innovative and problem-solving designs. This study highlights the potential of leveraging advanced AI technology in promoting more inventive design solutions. As AI technology incessantly advances, an increasing array of methods to encourage innovative design and facilitate human interaction will invariably arise.

ACKNOWLEDGMENT

The authors would like to thank Mr. "Milly" (https://github.com/Milly/) for technical assistance with stable-diffusion-webui and fine-tuning.

This work was supported by JSPS KAKENHI Grant Number JP 20K20500.

REFERENCES

Ishihara, S., Ishihara, K. & Nagamachi, M. (1996) Neural networks approach for Kansei analysis on milk carton design, Human Factors in Organizational Design and Management-V, 7–12, North-Holland.

Ishihara, S., Ishihara, K. & Nagamachi, M. (1998) Hierarchical Kansei analysis of beer can using neural networks, Human Factors in Organizational Design and Management-VI, 421–426, North-Holland.

Ishihara, S., Ishihara, K., Nagamachi, M & Matsubara, Y. (1995) arboART: ART based hierarchical clustering and its application to questionnaire data analysis, 1995 IEEE international conference on neural networks proceedings, 532–537.

Ishihara, S., Ishihara, K., Nakagawa, R., Nagamachi, M., Sako, H., Fujiwara, Y. and Naito, M. (2010), Development of a washer-dryer with Kansei Ergonomics, Engineering Letters, Vol. 18, No. 3, pp. 243–249.

Ishihara, S., Tsuchiya, T., Nagamachi, M., Nishino, T. and Ishihara, K. (2005) Catalytic effect of Kansei engineering system at collaborative design process, Human Factors in Organizational Design and Management-VIII, North-Holland, 401–406.

Jindo, T., Hirasago, K. Nagamachi, M. (1995) Development of a design support system for office chairs using 3-D graphics, Int. J. Industrial Ergonomics, 15(49–62).

Matsubara, T., Matsubara, Y., Ishihara, S. and Inokuchi, S. (2010), Virtual prototyping with real-time rendering for Kansei engineering of leather grain patterns on car dashboard panels, Journal of Japanese Kansei Engineering Society, 9(2), pp. 119–128

Nagamachi, M. (1991) An image technology expert system and its application to design consultation, International Journal of Human-Computer Interaction, Vol. 3, No. 3, pp. 267–279.

Nagamachi, M. (2011) Kansei/Affective engineering and history of Kansei/Affective engineering in the world, Nagamachi ed., Kansei/Affective Engineering, CRC press.

Nomura, J., Imamura, K., Enomoto, N. and Nagamachi, M (1998), "Virtual space decision support system using Kansei engineering", T. L. Kunii and A. Luchiani, Eds., Cyberworlds, Springer, Tokyo, Japan, pp. 273–288.

NovelAI (2022) NovelAI Improvements on Stable Diffusion, NovelAI Blog, https://blog.novelai.net/novelai-improvements-on-stable-diffusion-e10d38db82ac.

Radford, A. et al. (2021) Learning Transferable Visual Models from Natural Language Supervision, arXiv, https://doi.org/10.48550/arXiv.2103.00020.

Rombach, R. et al. (2021) High-Resolution Image Synthesis with Latent Diffusion Models, CVPR 2022, https://doi.org/10.48550/arXiv.2112.10752.

Shirai, A. (2022) Collaboration with AI makes you as a genius painter - Technical reading Stable Diffusion, Impress publishing (in Japanese).

Sohl-Dickstein, J., Weiss, E. A., Maheswaranathan, N. & Ganguli. S. (2015) Deep unsupervised learning using nonequilibrium thermodynamics. Proc. of the 32nd Int. Conf. on Machine Learning, Lille, France, 2015. JMLR: W&CP volume 37, abs/1503.03585.

Tokunaga, M. (2022) An in-depth look at Stable Diffusion, the image-generating AI that shocked the world, https://qiita.com/omiita/items/ecf8d60466c50ae8295b (in Japanese).