Aesthetic Design Based on the Analysis of Questionnaire Results Using Deep Learning Techniques

Masakazu Kobayashi[1], Suguru Fujita[2] and Takuya Wada[3]

[1]Toyota Technological Institute, kobayashi@toyota-ti.ac.jp
[2]Toyota Technological Institute, sd16070@toyota-ti.ac.jp
[3]Toyota Technological Institute, sd16092@toyota-ti.ac.jp

Corresponding author: Masakazu Kobayashi, kobayashi@toyota-ti.ac.jp

Abstract. In recent years, deep learning is attracting a lot of attention. Deep learning is a class of machine learning algorithms based on artificial neural networks. Deep learning uses multiple layers to extract higher-level features progressively and automatically from the raw input. Various types of deep learning models have been developed and applied to computer vision, speech recognition, natural language processing, audio recognition, self-driving car, board game, etc. In the field of sensitivity engineering, research that takes advantage of the features of deep learning has also been carried out. In this paper, a new aesthetic design method that creates new designs by analyzing customer's preferences using deep learning methods is proposed. More specifically, questionnaire investigations are carried out to collect customer's preferences for existing products and a CNN network that infer customer's preferences from product images is trained using questionnaire results. Aesthetic elements closely related to the customer's "like" evaluations are then identified and catalogued by analyzing the trained network using Grad-CAM and semantic segmentation. New designs are finally created by selecting and combining the favorite aesthetic elements from the catalog. In the case study, the proposed method was applied to a chair design. A lot of chair images were collected from the internet and subjects rated their preferences for them. Favorite aesthetic elements were then extracted and catalogued from the evaluation results and new chair designs were created by combining them. Created chair designs were finally rated by subjects and the effectiveness of the proposed method was confirmed.

Keywords: Kansei Engineering, Aesthetic Design, Deep Learning, Grad-CAM, Semantic Segmentation
DOI: https://doi.org/10.14733/cadaps.2022.602-611

1 INTRODUCTION

Due to the maturation of science and technology, it becomes increasingly difficult to differentiate products in terms of performance, functional features, or price. Therefore, companies are required to differentiate their products in terms of subjective and abstract qualities such as aesthetics and comfort that are evaluated by customers' feelings, which is called "Kansei" in Japanese. The quality evaluated by customer Kansei is called "Kansei quality" [18].

In the field of Kansei engineering (referred to as affective or emotional engineering), the methods for measuring customer Kansei or the impression of products have been developed and applied to many case studies. In these methods, semantic differential (SD) method [9] is widely used. Based on the measurement and analysis methods of customer Kansei, various aesthetic design methods have also been developed. These methods generate a new aesthetic design that a customer prefers best by revealing the relationships between the results of customers' Kansei evaluation of the same type of existing products as the design target and their aesthetic features. In these methods, various analysis methods such as artificial neural network [4] [5], fuzzy set theory [3], interactive reduct evolutionary computation [17], multidimensional scaling [5], rough set theory [6-8] [11] [16], self-organizing map [5], etc. are used.

In recent years, deep learning is attracting a lot of attention. Deep learning is a class of machine learning algorithms based on artificial neural networks. Deep learning uses multiple layers to extract higher-level features progressively and automatically from the raw input. In the case of image classification, lower layers may identify edges, while higher layers may identify the concepts relevant to a human such as digits, letters, or faces. Various types of deep learning models have been developed and applied to computer vision, speech recognition, natural language processing, audio recognition, self-driving car, board game, etc. Deep learning has also been used in the research on Kansei engineering. The most intuitive application of deep learning is to learn the relationships between product images and customers' preferences / impressions received from them using CNNs. In this application, when inputting unknown product images into the trained network, the customer's preferences / impressions received from those products can be inferred. More extensive applications of deep learning have also been developed in the following research. Ota et. al. proposed Kansei retrieval system to search for user's favorite clothing based on CNN and ANN [10]. Quan et. al. proposed the Kansei engineering-based neural style transfer for product innovation (KENPI) framework [12]. Dai et. al. proposed the approach for automatic design scheme generation based on generative adversarial networks (GAN) [1][2]. Schmitt and Weiss designed innovative chairs inspired by the chair images generated by GAN [15].

In this paper, a new aesthetic design method that generates new designs by analyzing customer's preferences using deep learning methods is proposed. More specifically, questionnaire investigations are carried out to collect customer's preferences for existing products and a CNN network that infer customer's preferences from product images is trained using questionnaire results. Aesthetic elements closely related to the customer's "like" evaluations are then identified by analyzing the trained network using Grad-CAM and semantic segmentation. Here, "Aesthetic element" is defined as a part of a product from the viewpoint of product aesthetics. For example, "backrest", "seat", "armrest", and "leg" are aesthetic elements of a chair. New designs are finally created by combining the collected favorite aesthetic elements. In general, product consists of various types of aesthetic elements. Customers judge whether they like or dislike the product by looking at the individual aesthetic elements, their combinations, or an overall design of the product. Therefore, by collecting aesthetic elements that customers prefer and combining them, new designs that are highly likely to be preferred by customers can be created. In addition, the proposed method confirms whether the created designs are preferred by customers or not by applying CNN and Grad-CAM to them. This avoids the case where individual aesthetic elements are desirable to the customer, but the entire product as a combination of them is not desirable.

2 PROPOSED METHOD

The proposed method consists of the following 4 steps. Their details are explained in the following sections.

Step1: Investigation of customer’s preferences using questionnaires
Step2: Leaning of customer’s preferences using CNN
Step3: Analysis of customer’s preferences using Grad-CAM
Step4: Generation of new designs

2.1 Step1: Investigation of Customer’s Preferences using Questionnaires

The first step is to collect customer’s preferences for existing products through questionnaire investigations. Since training a CNN requires a large amount of training data, e.g., thousands, the questionnaire program shown in Figure 1 is constructed in order to reduce the load on the subjects, in the case study of this paper. The program randomly presents prepared product images and allows the subject to enter the liking or disliking by two keys on the keyboard.

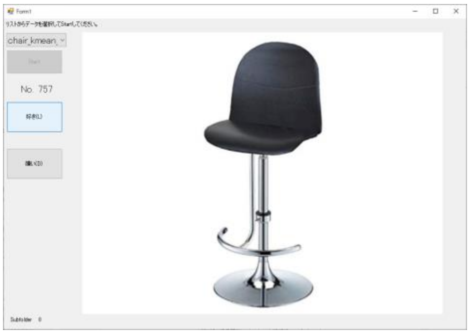

Figure 1: Questionnaire program.

Images of existing products of the same type as the design target are collected, their background is erased, and they are resized to the same size. A customer carries out a questionnaire investigation using the program. During questionnaire, a customer rates whether he / she likes or dislikes the displayed product images.

2.2 Step2: Leaning of Customer’s Preferences using CNN

The relationships between images of existing products and customer’s preferences for those products collected in step1 are learned by using a CNN. CNNs are often used for image classification tasks such as classifying an animal in an image as a dog or a cat. In the proposed method, a CNN infers customer’s preferences (like or dislike) for unknown products. Various types of CNN models have been proposed and any of them can be used in this step. In the case study of this paper, a relatively simple CNN model consisting of 4 convolution layers and 2 max pooling layers is used.

When a customer is asked to rate his / her preferences for many product images, as in this study, the accuracy and reliability of the questionnaire results may be insufficient. Therefore, the accuracy of the trained network is verified and if the accuracy is low, the customer is asked to re-evaluate the questionnaire investigations.

2.3 Step3: Analysis of Customer's Preferences using Grad-CAM

Using Grad-CAM and semantic segmentation, aesthetic elements that are closely related to the customer's "like" evaluations are identified and collected from the product images that a customer evaluates as favorable in the questionnaire investigation.

Grad-CAM is a technique for producing 'visual explanations' for decisions from a large class of CNN-based models, making them more transparent and explainable [14]. CNNs have been successfully applied to a wide variety of computer vision tasks, such as image classification, object detection, semantic segmentation (image captioning, visual question answering, visual dialog, and embodied question answering). The most critical problem for using CNNs is difficulty in understanding why the system did what it did. Therefore, various types of 'visual explanation' techniques that reveal what the CNN focused on in the input image to make its inference by analyzing the trained network have been proposed. In the proposed method, Grad-CAM is used.

Semantic segmentation is the task of clustering each pixel in an image together which belongs to the same object class. In the proposed method, semantic segmentation is used to identify aesthetic elements from product images. For example, it recognizes the backrest, seat, elbows, and legs from an image of a chair. Various types of semantic segmentation methods have been proposed. In the case study of this paper, U-Net [13] is used.

The detailed procedure of Step3 is as follows. (1) Using Grad-CAM, what part of images of the customer's favorite products is related to the customer's preference is analyzed. Since Grad-CAM calculates the effect of each pixel in an image on the inference result, the relationships between aesthetic elements and customer's preferences are not yet identified. (2) Using U-Net, aesthetic elements are identified from images of the customer's favorite products. (3) From the results of Grad-CAM and semantic segmentation, the impact of aesthetic elements that make up products on customer's "like" evaluations is evaluated. Specifically, the impact is calculated by dividing the sum of the importance weights of the pixels that make up the aesthetic element on the product image by the number of the pixels. The importance weight of each pixel is obtained by Grad-CAM while the design element to which each pixel belongs is identified by U-Net. The aesthetic elements with the impact higher than the threshold are closely related to customer's like evaluation and named "favorite aesthetic elements". The threshold value is determined by a designer by trial and error. The reason for dividing by the number of pixels that make up the aesthetic element is that the size of the aesthetic element can vary greatly depending on the design. In the example of chairs shown in Figure 1, the number of pixels that make up the legs of the left chair is much larger than the one on the right. (4) If more than one favorite aesthetic element is included in a single product, there may be a combination effect on those elements. The combination effect is a phenomenon in which these elements are perceived as preferable only when they are used in a single product simultaneously. If there is a combination effect between elements, those elements must always be used in combination. To confirm a combination effect between elements, one of the elements is deleted from the original product image and the image is analyzed by Grad-CAM. If the impact of the element left in the image is still high, a customer feels that the element is preferable in isolation. In this case, the element can be treated as favorite aesthetic elements by itself. If the impact of the elements left in the image becomes lower than the threshold value, the element is not preferable alone and must always be used in combination. Finally, favorite aesthetic elements are catalogued. In the catalog, aesthetic elements that can be used alone and ones that can only be used in combination with two or more other parts are stored separately.

2.4 Step4: Generation of new Designs

Finally, new product designs are generated by manually selecting favorite aesthetic elements from the catalogue developed in Step3 and combining them. Created designs are then evaluated by the CNN and Grad-CAM. This is because even if new designs are created by combining only the favorite aesthetic elements, the overall design may be unfavorable due to negative combination effects and other reasons. The CNN infers whether a customer prefers the overall design of the created products whereas the Grad-CAM analyzes whether a customer prefers each part of the created products. If CNN and Grad-CAM infer that a customer does not prefer the created design, it is discarded, and further design generation is attempted.

3 CASE STUDY

To show the flow of the proposed method, it was applied to the design of the chair. The design targets were a wide variety of chairs such as office chairs, dining chairs, and sofas. Since the development of the proposed method was carried out over a relatively long period of time, two case studies were carried out during that process. In the fist case study, the favorite aesthetic elements were manually extracted by the designer without calculating the impact of each design element, and new designs were created by combining them. In the second case study, new designs were created by following all the steps of the proposed method described in the previous section. To validate the effectiveness of the proposed method based on more results, both case studies are described here.

3.1 Collection of Product Images

Chair images for the questionnaire investigation were collected from the Internet. Only chairs with the same orientation were selected from them and the background of the chairs was changed to white using photo-retouching software. 4684 chair images were prepared. Figure 2 shows their examples. Collected chair images were used in both case studies.

Figure 2: Examples of chair images.

3.2 Case Study 1

In the first case study, 1 male undergraduate student was participated as a subject. A subject rated his preferences for 4684 chairs using a questionnaire program shown in Figure 1.

A CNN network was then trained using chair images and questionnaire results. As described above, the CNN network used in the case study consisted of 4 convolution layers and 2 max pooling layers. 75% of training data was used for training and 25% for test. The performance metrics of the trained network were shown in Table 1. This is worse than expected. This is probably because a large number of questionnaires, even with the questionnaire program, made it difficult for a subject to evaluate them consistently.

Accuracy	Precision	Recall	F-measure
0.813	0.901	0.805	0.85

Table 1: Performance metrics of the trained network.

The chairs that the customer evaluated as "like" were analyzed by using Grad-CAM. Figure 3 shows examples of the results of Grad-CAM. The redder the pixel in the image, the more relevant it is to the customer's "like" evaluation. As shown in Fig.3, there are also cases where the analysis cannot be done well.

Figure 3: Examples of the results of Grad-CAM.

Next, the U-Net network was trained to identify aesthetic elements from images of the customer's favorite products using semantic segmentation. For training the network, 1000 training data as shown in Figure 4 were manually prepared in advance. In the case study, a chair is divided into 3 elements namely "backrest & seat", "armrests", and "leg". Using the trained network, semantic segmentation was applied to the customer's favorite products and the results were compared with the results of Grad-CAM. Figure 5 shows examples of the results of semantic segmentation and Grad-CAM. By comparing these two images, it becomes clear which aesthetic elements of the chair are closely related to the customer's "like" evaluation. This series of operations was applied to multiple product images to collect candidates for aesthetic elements to be used in Step 4.

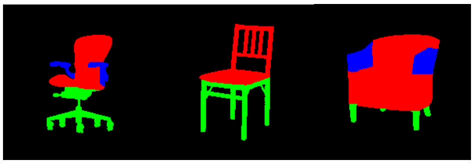

Figure 4: Examples of training data for semantic segmentation.

Figure 5: Comparison of the results of Grad-CAM and semantic segmentation.

Finally, 10 new designs were generated by combining the aesthetic parts and evaluated using CNN and Grad-CAM. Figure 6 shows examples of the generated design and evaluation results. The subject was asked to evaluate the generated designs and he evaluated 9 out of 10 products as "like".

Figure 6: Examples of the created chair designs.

3.3 Case Study 2

In the second case study, 3 male undergraduate students were participated as subjects. They were different from the student who participated in the first case study. Experiments with 3 subjects were conducted independently.

3 subjects rated his preferences for 4684 chairs and 3 CNN networks were trained independently by using their questionnaire results. Table 2 shows the performance metrics of 3 trained networks.

	Accuracy	Precision	Recall	F-measure
Subject1	0.806	0.858	0.833	0.845
Subject2	0.648	0.33	0.496	0.397
Subject3	0.823	0.949	0.853	0.898

Table 2: Performance metrics of 3 trained networks.

As shown in Table 2, performance of the network of Subject 2 was considerably worse than those of Subject 1 and 3. It would be desirable to conduct a questionnaire investigation for Subject 2 again, but due to time constraints, the case study was carried on without another questionnaire investigation. Favorite aesthetic elements were then catalogued by using Grad-CAM and U-Net. Figure 7 shows a part of the catalog of Subject 1. Here, the red, green, and blue dos marked on chairs indicate that the seats, armrests, and legs of those chairs are favorite aesthetic elements. A distinction is made between favorite aesthetic elements that can be used alone and ones that must be used in combination.

Finally, five new designs were generated for each subject by selecting and combining favorite aesthetic elements from their own catalogs. Figure 8 shows created chairs. 3 subjects were asked to evaluate the created chairs. For the evaluation, 10 chairs were prepared by mixing 5 newly created chairs and 5 chairs randomly selected from 4684 existing chairs. As a result, Subject 1 and 3 rated 4 out of 5 chairs as preferred, and Subject 2 rated 3 out of 5 chairs as preferred. The results of 2 case studies show that the proposed method can generate new designs from the results of questionnaires on existing products, and these new designs are preferred by customers.

4 CONCLUSION

To generate new designs that customers prefer from the results of questionnaires on existing products, a new aesthetic design method using on CNN and Grad-CAM. In the proposed method, Grad-CAM calculates the effect of each pixel in an image on a customer's "like" evaluation by analyzing the CNN network that learns customer preferences. Semantic segmentation then identifies aesthetic elements from product images.

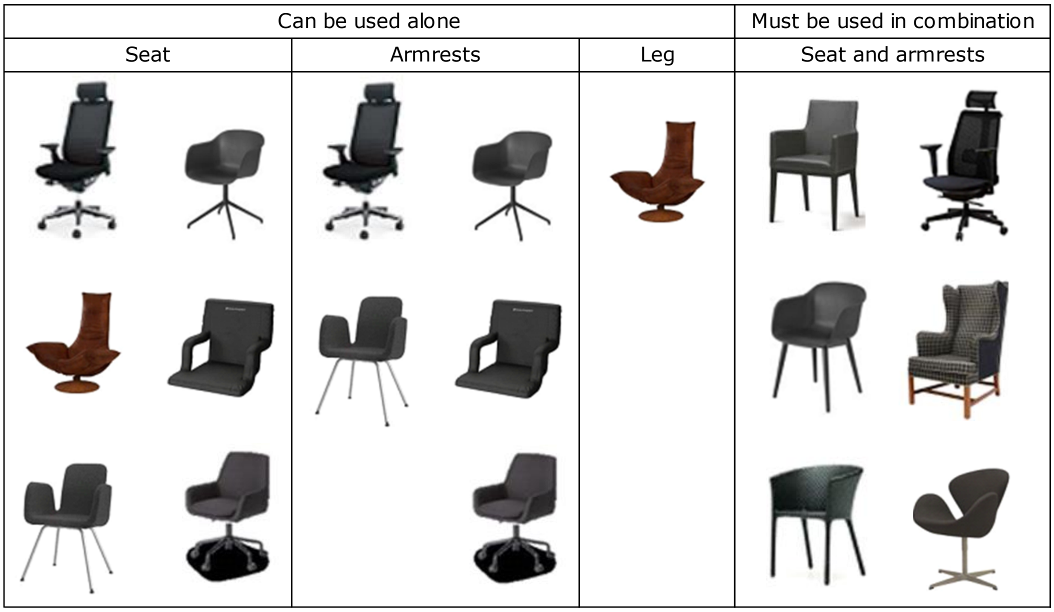

Figure 7: A part of the catalog of Subject 1.

Figure 8: Newly created designs.

By comparing two results, the aesthetic elements that affect customer's "like" evaluation are be clarified. Therefore, by collecting and combining their parts, new designs that customers prefer are generated. In the case study, the proposed method was applied to chair design, and its effectiveness was confirmed.

The limitation of the proposed method recognized as future issues are as follows. (1) The proposed method can analyze which aesthetic elements are favorable to customers, but cannot indicate which combination of elements is most favorable. (2) The process of integrating the favorite aesthetic elements into a single design needs to be done manually by the designer. (3) It is not a problem only for the proposed method, but it is difficult to collect a large amount of training data from customer questionnaires.

Masakazu Kobayashi, https://orcid.org/0000-0002-1212-2879
Suguru Fujita, https://orcid.org/0000-0003-1015-4727
Takuya Wada, https://orcid.org/0000-0001-7346-8104

REFERENCES

[1] Dai, Y., Li, Y., Liu, L.: New Product Design with Automatic Scheme Generation, Sensing and Imaging, 20(29), 2019. http://dx.doi.org/10.1007/s11220-019-0248-9
[2] Goodfellow, I J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., Bengio, Y.: Generative Adversarial Networks, NIPS'14: Proceedings of the 27th International Conference on Neural Information Processing Systems, 2, 2014, 2672-2680. http://dx.doi.org/10.1145/3422622
[3] Hsiao, S.W.; Huang, H.C.: Applying the semantic transformation method to product form design, Design Studies, 19(3), 1998, 309-330. http://dx.doi.org/10.1016/S0142-694X(98)00009-X
[4] Hsiao, S.W.; Huang, H.C.: A neural network based approach for product form design, Design Studies, 23(1), 2002, 67-84. http://dx.doi.org/10.1016/S0142-694X(01)00015-1
[5] Kobayashi, M.; Kinumura, T.; Higashi, M.: A Method for Supporting Aesthetic Design Based on the Analysis of the Relationships Between Customer Kansei and Aesthetic Element, Computer-Aided Design & Applications, 13(3), 2015, 281-288. http://dx.doi.org/10.1080/16864360.2015.1114385
[6] Kobayashi, M.; Niwa, K.: Method for Grouping of Customers and Aesthetic Design Based on Rough set theory, Computer-Aided Design & Applications, 15(4), 2018, 565-574. http://dx.doi.org/10.1080/16864360.2017.1419644
[7] Kuramaru, K.; Takanashi, R.; Mori, N.: The Method of Design Project Based on the Declaration of Checking Preference and Non-preference with the Product, Journal of Japan Society of Kansei Engineering, 1(1), 2001, 65-72. http://doi.org/10.5057/jjske2001.1.65
[8] Ohki, M.; Harada, T.; Inuiguchi, M.: Decision Rule Visualization System for Knowledge Discovery by Means of the Rough Set Approach, Journal of Japan Society for Fuzzy Theory and Intelligent Informatics, 24(2), 2012, 660-670. http://doi.org/10.3156/jsoft.24.660
[9] Osgood, C. E.; Suci, G. J.; Tannenbaum P.: The Measurement of Meaning, University of Illinois Press, 1967.
[10] Ota, S.; Takenouchi, H.; Tokumaru, M.: Kansei retrieval of clothing using features extracted by deep neural network, Transactions of Japan Society of Kansei Engineering, 16(3), 2017, 277-283. https://doi.org/10.5057/jjske.TJSKE-D-17-00003
[11] Pawlak, Z.: Raugh Sets, International journal of Information Computer Science, 11(5), 1982, 341-356. http://doi.org/10.1007/BF01001956
[12] Quan, H.; Li, S.; Hu, J.: Product Innovation Design Based on Deep Learning and Kansei Engineering. Applied Sciences, 8(12), 2018, 2397-2415. http://dx.doi.org/10.3390/app8122397
[13] Ronneberger, O.; Fischer, P.; Brox, T.: U-Net: Convolutional Networks for Biomedical Image Segmentation, MICCAI 2015: Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015, 2015, 234-241. http://dx.doi.org/10.1007/978-3-319-24574-4_28
[14] Selvaraju, R. R.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D.: Grad-CAM: Why did you say that? Visual Explanations from Deep Networks via Gradient-based Localization, International Journal of Computer Vision, 128, 2016, 336-359. http://dx.doi.org/10.1007/s11263-019-01228-7
[15] Schmitt, P.; Weiss, S.: The Chair Project, https://philippschmitt.com/work/chair

[16] Yamada, K.; Moroga, U.; Unehara, M.: Design Support for Generating Novelty with Rough Sets Theory and Conceptual Hierarchy, Transactions of Japan Society of Kansei Engineering, 11(1), 2012, 17-26. http://dx.doi.org/10.5057/jjske.11.17
[17] Yanagisawa, H.; Fukuda, S.: Kansei Design by Interactive Reduct Evolutionary Computation: With Attention Paid to Favored Feature of Design, Transactions of the Japan Society of Mechanical Engineers. C, 70(694), 2004, 1802-1809. http://dx.doi.org/10.1299/kikaic.70.1802
[18] Yanagisawa, H.: Kansei quality in product design, Emotional Engineering, 2011, 289-310. http://dx.doi.org/10.1007/978-1-84996-423-4_16