

An Article Retrieval Support System That Accepts Arbitrary Kansei Words

Yuichi Murakami, Shingo Nakamura, and Shuji Hashimoto

Waseda University, Tokyo, 169-8555, Japan
(Tel: 81-3-5286-3233, Fax: 81-3-3202-7523)

yuichi@shalab.phys.waseda.ac.jp, shingo@shalab.phys.waseda.ac.jp, shuji@waseda.jp

Abstract: In most article retrieval systems using Kansei words, there exists a gap between user's Kansei and system's Kansei model. Therefore, it is not always easy to retrieve the desirable articles. The purpose of this paper is to bridge this gap by introducing a novel function to accept arbitrary Kansei words. The parameters related to newly added Kansei words are automatically estimated from articles selected by the user during search. Thereby, the system is customized for individual users, to bridge the gap. In the evaluation experiments, we took actual paintings as the articles, and evaluated the user-satisfaction, accuracy, *novelty* and *serendipity* of the system through actual user operations. Most of users were satisfied with the system and could find unknown desirable articles. Moreover they could often find unexpectedly interested articles.

Keywords: User profiling, Kansei information processing, Cooperative system, Web retrieval

1 INTRODUCTION

Web services are increasing day by day and the demands for e-commerce have also become higher. It is, however, not easy to retrieve a proper article from a lot of articles in web pages. One way of retrieving articles is to use keywords. Especially using adjective keywords called "Kansei words" is effective to represent human feelings and impressions. However, the accuracy of retrieval results is not sufficient due to the gap between user's Kansei and system's Kansei model. Therefore, it takes a long time to retrieve the desirable articles. To bridge this gap, we consider that it is necessary to introduce user profiles, and implementation of the following two functions: 1) to add arbitrary Kansei words freely, 2) to fit system Kansei words to the user's Kansei. Hijikata pointed out that user profiling methods can be classified as either explicit or implicit methods [1]. In the explicit method, users directly input the information about individual taste. In the implicit method, the system estimates the individual taste based on the past user behavior. As examples of explicit method, individual tastes are estimated by the questionnaire in advance [2][3], or by asking users to grade the browsed internet pages on their interest and relevance [4]. Such explicit methods put a strain on users. On the other hand, implicit methods infer individual taste by usage history [5][6]. However, they need enough usage history. Most of studies about Kansei retrieval system took an explicit method. However, we take an implicit method in order not to put a strain on users together with above two functions.

We already proposed the system to fit system Kansei words to the user's Kansei in prepared default Kansei words decided by a system architect [7]. In this paper, we append the function of adding arbitrary Kansei words to the

previous system. The proposed system collects articles fitting or non-fitting to the added Kansei word while users search the desirable articles. Utilizing the collected articles, the system estimates the parameters related to the added Kansei words by using the support vector machine (SVM) method. Thereby, users can freely use arbitrary Kansei words, and customize to the own desirable system by editing retrieval criteria as they please. In the result, we expect to improve the satisfaction because users are not tired of the system and the flexibility of retrieval is improved. However, the parameters related to the added Kansei words are learned through user's most desirable article by neural networks because the added Kansei words are often low accuracy.

For the experiments, we developed an article retrieval support system implemented with the proposed method, and employed artistic painting images for article database. We evaluated the user-satisfaction and accuracy through actual operations. We also employed *novelty* and *serendipity* as the evaluation indices [8]. Users have potentially multiple preferences for diverseness of Kansei words [9]. Therefore, these two metrics are useful to measure how much ambiguous or unaware preferences are derived.

2 PROPOSED SYSTEM

Most users wish to search articles not only by given Kansei words but also by suitable their own Kansei words. However, the system cannot prepare all Kansei words in advance because all most of the words and expressions have the potential to be utilized as Kansei words. Therefore, it is important to add arbitrary Kansei words to the system freely. However, the system has to parameterize the arbitrary Kansei word in accordance with all articles, which

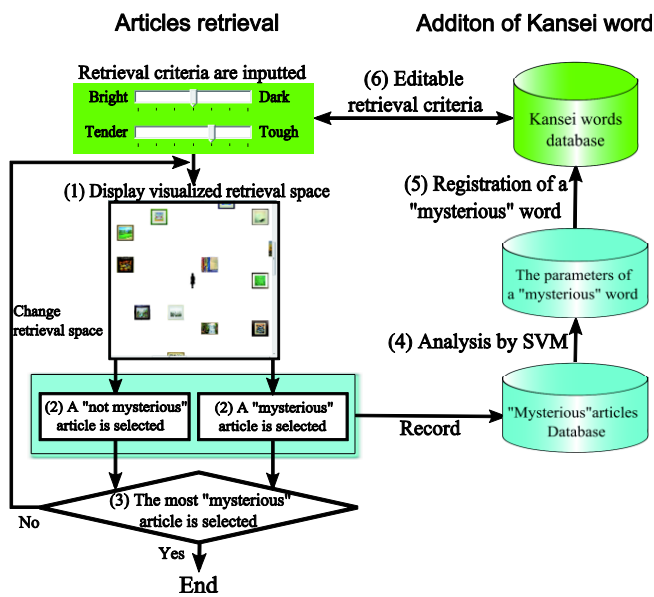


Fig.1. The schematic diagram of proposed system

is laborious. Meanwhile, Mukunoki et al reported that some typical and popular Kansei words can express other Kansei words [9], although it needs images selected by specialists in order to add arbitrary Kansei words. We suggest a new procedure to treat arbitrary Kansei words by utilizing the given default Kansei words.

2.1.1. Registration and usage of arbitrary Kansei words

Fig.1 shows the schematic diagram of the proposed system in case of adding an arbitrary word “mysterious”. First, the proposed system needs to find out which articles fit the added Kansei word for users as teacher data for supervised learning. The Kansei retrieval system we proposed before is employed to collect such data efficiently. The previous system can visualize the two-dimensional retrieval space converted from Kansei parameter space by Self Organizing Maps ((1) in Fig.1). Moreover, the user can point not only desirable articles but also undesirable ones. In the case of selecting undesirable articles, the retrieval space is changed to follow the user’s tendency. With the help of this system, users repeat the selection of articles fitting or non-fitting to the added Kansei word. This operation continues until users find the most fitting article for the added Kansei word, or select a sufficient number of articles. The user’s strain to collect teacher data is minimized because the system concomitantly collects the data while users repeat the selection to find the most desirable article. At this stage, the articles fitting the added Kansei word are scored by +1. On the other hand, the articles not fitting it are scored by -1. After the teacher data are collected enough, the system estimates the parameters

of the added Kansei word by SVM. The added Kansei word is registered as retrieval criteria in the same way as the default Kansei words after this procedure. Additionally, the system can also create the new retrieval space concerning the parameters of the added Kansei word. Therefore, the system can create an opportunity to discover unknown desirable articles as unexpectedly interested articles.

2.1.2. Editable retrieval criteria

Users can freely use arbitrary Kansei words added by themselves. Besides, Kansei words added by other users are also available as own retrieval criteria by accessing the database which stores the parameters of each added Kansei word. Users can also eliminate the added Kansei words from retrieval criteria ((6) in Fig.1). Therefore, users are not tired of the system because they can customize the system.

2.1.3. Learn parameters of added Kansei words

The previous system learns user’s Kansei based on user’s most desirable article with a three-layered perceptron. We extend the previous system to the range of added Kansei words. Learning method is the same as the previous system. The system learns the relationship between the initial input retrieval criteria’s parameters including added Kansei words and the finally selected article ones. In other words, the perceptron’s input data is the initial position on the visualized retrieval space, and the teacher data is the position of the article finally selected.

3 IMPLEMENTATION

We made some experiments to evaluate the usage of the proposed retrieval system for painting images to assist the customers who wish to buy paintings. In the experiments, 135 painting images were stored in the database. They were rated by 20 people with 7-point scale of 8 adjective pairs shown in Table 1. The adjective pairs were chosen according to the Inoue’s study on Kansei evaluation of paintings [10]. The average of human evaluation was set into the system database in advance as the initial system’s Kansei model. The system also allowed users to add arbitrary Kansei words based on this 8 adjective pairs.

We utilized the software library LibSVM for SVMs modeling. We ran a preliminary experiment to avoid over fitting. In the result, we chose C-SVM with RBF kernel. If the teacher data existed only in one side, we chose one-class SVM. Cost parameter and ν parameter were set to 2 and 0.1. Other tool parameters were set to the default values of the LibSVM. The other system parameters and software’s composition were conformed to the previous system.

Table 1. Pairs of Kansei word

Pair #	Kansei Words	
1	Bright	Dark
2	Tender	Tough
3	Cheerful	Melancholy
4	Fancy	Plain
5	Complicated	Simple
6	Sharp	Dull
7	Warm	Cold
8	Comfortable	Uneasy

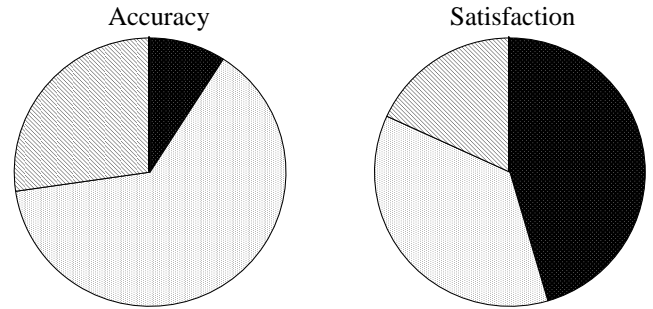
4 EVALUATION EXPERIMENTS AND RESULTS

In the evaluation experiments, we employed *novelty* and *serendipity* metrics as the evaluation indices other than the satisfaction or accuracy. They can measure the “non-obviousness” of the retrieval system as an important metrics for user-satisfaction. These metrics indicate whether the recommended item is both unknown and favorite. We evaluated that how many unknown desirable articles about a new Kansei word were recommended. A serendipitous recommendation helps users to find an unexpectedly interested item which might not be discovered in other way. For measuring serendipity, it is necessary to evaluate how the recommended items attract and surprise the user. However, no concrete calculation method has been proposed so far. Therefore, subjective assessment was adopted.

4.2.1. Satisfaction and accuracy

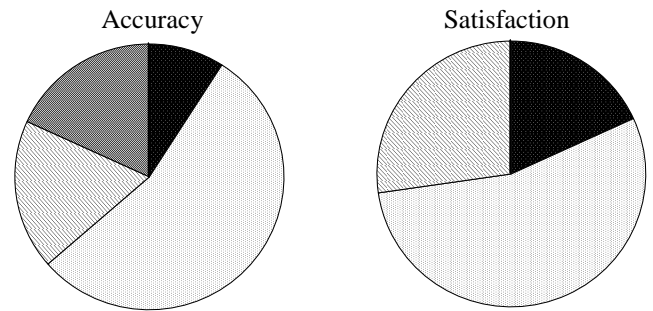
We asked 13 people the accuracy and satisfaction of an arbitrary Kansei word added by themselves. We also asked the accuracy and satisfaction when users utilized Kansei words added by other users. The result is shown in Fig.2. Here, we defined “Good” and “Excellent” evaluations mean “be satisfied”. We concluded that added Kansei words are useful because over 70% of the users were satisfied with it. Though accuracy of Kansei words added by other users are occasionally inadequate, satisfaction level is high. Interestingly, some users commented “it is fun to feel other user’s Kansei though the accuracy is low.” Next, we asked the satisfaction of editable retrieval criteria. The result is shown in Table 2. Most of the users were satisfied with the function of editable retrieval criteria. We consider that users were satisfied with not only searching user’s desirable articles but also customizing the system.

Moreover, we evaluated the function of neural networks. 8 users tried retrieving their preferred paintings five times. Afterward we compared the positions of Kansei learning with the initial positions. Table 3 shows the average result for 8 users. The first and second columns represent the



■ Excellent □ Good ▨ Fair ■ Poor ▩ Very poor

(a) Arbitrary Kansei words added by themselves



■ Excellent □ Good ▨ Fair ■ Poor ▩ Very poor

(b) Kansei words added by other users

Fig.2. The accuracy and satisfaction for added Kansei words

Table 2. Editable retrieval criteria

Evaluation	Excellent	Good	Fair	Poor	Very poor
Rate	23.1%	69.2%	7.7%	0.0%	0.0%

Table 3. Result of by Kansei learning

	Without Kansei learning	With Kansei learning	Difference
Average (5 times try)	1.00	0.64	-0.36

distance values between the initial position and the finally selected position with Kansei learning. The values are the normalized distance of visualized retrieval space. The third column indicates the difference. The initial position gets about 0.36 closer to the selected article. It can be concluded that the system fits to individual user’s preference in the course of using the system even if arbitrary Kansei words are included in retrieval criteria.

4.2.2. Novelty and serendipity

We evaluated the *novelty* and *serendipity*. We asked 10 people to show the articles in the database that fit to the added Kansei word. We then distilled teacher data used for creating a new Kansei word. We consider that these articles are regarded as novel articles because they are

Table 4. The novelty for adding Kansei words

Added Kansei words	Novelty		
	Precision	Recall	F value
<i>mysterious</i>	0.50	0.43	0.46
<i>boisterous</i>	0.33	0.50	0.40
<i>realistic</i>	0.38	0.26	0.31
<i>unusual</i>	0.43	0.20	0.27
<i>disorganized</i>	0.17	0.50	0.25
<i>expansive</i>	0.30	0.13	0.18
<i>enjoyable</i>	0.00	0.00	0.00
<i>likeable</i>	0.29	0.29	0.29
<i>good</i>	0.25	0.16	0.20
<i>beautiful</i>	0.17	0.06	0.09
average	0.28	0.25	0.26

Table 5. The serendipity for adding Kansei words

Evaluation	Yes	Rather yes	Yes and no	Rather No	No
Rate	38.5%	46.1%	0.0%	7.7%	7.7%

fit to the Kansei word. We evaluated the precision, recall and F value of novel articles. Here, if the values of the new Kansei word by SVM were above 0.85, they were regarded as fitting articles. The result is shown in **Table 4**. The F value of Hijikata's system specialized for novelty is about 0.3. However, the F value of typical system not specialized for novelty is under 0.1 [8]. Therefore, our proposed system showed better results than previous system in novelty. Moreover, the adjectives having high "evaluation factor" (e.g. *likeable*, *good* and *beautiful*) generally is difficult for utilizing as retrieval criteria because there are great differences between individuals. However, users can occasionally find unknown desirable articles.

We also evaluated serendipity by subjective assessment of the proposed system. 13 people tried and the result is shown in **Table 5**. Most of users answered "yes" or "rather yes" about adding arbitrary Kansei words. We consider that the system created a new retrieval space concerning the parameters of the added Kansei word, and an opportunity to find unexpectedly interested articles. In the result, we concluded that users could find unknown desirable articles, and could find unexpectedly interested articles by using the proposed system.

5 CONCLUSION

We have proposed an article retrieval support system that accepts arbitrary Kansei words not to put strain on users. Users can freely add arbitrary Kansei words, and edit retrieval criteria. Moreover, the parameters related to the added Kansei words are learned by neural networks. In the experiments, we confirmed that most of users were satisfied with adding arbitrary Kansei words, and customizing the system. Besides, the accuracy of the system was improved.

We also confirmed that users could find unknown desirable articles, and could find unexpectedly interested articles by using the proposed system.

We are trying to apply the developed system for other different article retrievals to examine the universality of the proposed method. In the future, we would like to link physical parameters to Kansei factors, and to implement the automatic indexing of articles.

ACKNOWLEDGMENT

This work was supported in part by the CREST project "Foundation of technology supporting the creation of digital media contents" of JST, and the "Global Robot Academia," Grant-in-Aid for Global COE Program by the Ministry of Education, Culture, Sports, Science and Technology, Japan.

REFERENCES

- [1] Hijikata Y. (2004), Implicit user profiling for on demand relevance feedback. Proceedings of the 9th International Conference on Intelligent User Interfaces, pp. 198-205
- [2] Sugihara T., Morimoto K. and Kurokawa T. (2004), An Improved Kansei-Based Music Retrieval System with a New Distance in a Kansei Space. Proceedings of the 2004 IEEE International Workshop on Robot and Human Interactive Communication, pp.141-146
- [3] Miyakawa A., Ishida T. and Shibata Y. (2008). Visual and Physical Feature Analysis for Kansei Retrieval Method in VR Japanese Crafting Presentation System. Workshops 22nd International Conference on Advanced Information Networking and Applications, pp.477-482
- [4] Pazzani M. and Billsus D. (1997), Learning and Revising User Profiles: the Identification of Interesting Web Sites. Machine Learning, 27(3), pp.313-331
- [5] Sugiyama K., Hatano K. and Yoshikawa M. (2004), Adaptive Web Search Based on User Profile Constructed without Any Effort from Users. Proc. of WWW'04, pp.675-684
- [6] Sakurai Y. and Tsuruta S. (2009), A Searching Method based on Mechanically Learned Subjective Image. Journal of Japan Society for Fuzzy Theory and Intelligent Informatics 21(2), pp. 214-221
- [7] Murakami Y., Nakamura S. and Hashimoto S. (2010), An article retrieval support system that learns user's Kansei. User Science and Engineering, 2010 International Conference on, pp.32-37
- [8] Hijikata Y., Shimizu T. and Nishida S. (2009), Discovery-oriented Collaborative Filtering for Improving User Satisfaction. 14th ACM International Conference on Intelligent User Interfaces, pp.67-76
- [9] Mukunoki M., Tanaka H. and Ikeda K. (2001), An Image Retrieval System Using the Feature Space Consisting of Antonymous Pairs of Adjectives. Transactions of Information Processing Society of Japan 42(7), pp.1914-1921
- [10] Inoue M., and Kobayashi T. (1985), The research domain and scale construction of adjective-pairs in a semantic differential method in Japan. The Japanese journal of educational psychology 33(3), pp. 253-260