DEIM Forum 2023 5a-8-5

Craft Visualization System based on Characteristic Extraction and Emotion Analysis

Zerong SONG† and Yuanyuan WANG†

† Graduate School of Sciences and Technology for Innovation, Yamaguchi University
2–16–1 Tokiwadai, Ube, Yamaguchi, 755–8611 Japan
E-mail: †{c109vgw,y.wang}@yamaguchi-u.ac.jp

Abstract Crafts express regional culture in a unique way and carry on the cultural continuity of a local region or even a country. Globalization of the economy has led to better dissemination of crafts worldwide, and the crafts have become bridges between different places and the world. Nowadays, users pay more attention to artifacts found on webpages, streaming media, etc. However, it is difficult for users to immediately obtain a thorough grasp of crafts from a region because of the large number and variety of crafts in the world. Therefore, in this paper, we develop a system for visualizing crafts based on those practiced in Japan's Kinki area. For visualizing crafts to represent their characteristics and emotional attributes, we utilize traditional crafts provided by the Ministry of Economy, Trade and Industry (METI) through characteristic extraction and sentiment analysis. Finally, we evaluated the usability of the proposed craft visualization system, and we also confirmed that our proposed system could help users understand regional crafts quickly and comprehensively.

Key words traditional crafts, characteristic extraction, emotion analysis, visualization

1 Introduction

According to the definition of the Ministry of Economy, Trade, and Industry (METI)(注1) of Japan, traditional crafts are items made by a small number of people in a given region for everyday use, using traditional raw materials as the primary raw material and having the main part done by the crafts and traditional skills. Japanese traditional crafts are very popular among users all over the world because of their practicality, aesthetics, and craftsmanship they convey. According to the data released by the Ministry of Economy, Trade, and Industry (METI), 240 traditional crafts are currently designated. Since 2016, when it fell below 100 billion yen, the volume of traditional craft production has been gradually decreasing. In 2020, there were approximately 54,000 related employees, a number that is also gradually declining. As a result, we can conclude that the number of traditional crafts in Japan is large and widely distributed, and it is difficult to preserve them because there are not enough successors. For this reason, the visualization and dissemination of traditional crafts are particularly important.

Therefore, the purpose of this paper is to develop a craft visualization system based on the traditional crafts of the Kinki region in Japan. We provide the details of our developed craft visualization system in our ongoing work [1]. Through characteristic extraction, emotion analysis, and visualization, we aim to make users around the world aware of the traditional crafts of the region, promote the crafts, and facilitate regional activation as well as international cross-cultural exchange.

(注1)：https://www.meti.go.jp

2 Related Work

In characteristic extraction, methods such as *TF-IDF*, LDA is often used. Kowsari et al. [2] applied LDA to topic modelings and other document processing tasks, such as document classification and collaborative filtering. Jha et al. [3] presented a new and simple approach towards text-to-emoticon conversion and vice-versa using NLTK and WordNet.

In text emotion analysis, Ahmed et al. [4] proposed an attention-based LSTM model to address aspect-level emotion analysis task based on the emotion score retrieved from the proposed dictionary to weight down the non-emotion parts among aspect-related information in a given sentence. BERT, a language representation model of the Google AI Language API, aims to quantify where linguistic information is captured in the network. In 2019, qualitative analysis experiments conducted by Tenney et al. [5] showed that the BERT model can dynamically adapt to the traditional NLP pipeline, modifying low-level decisions based on ambiguous information from higher-level representations.

In visual analysis,methods include wordcloud, treemap, themeriver, etc. Chen et al. [6] proposed Tagnet, a model based on node linkage graph and wordcloud, which color-codes nodes and linkages to show the emotion attributes of tags. Brooks et al. [7] used an emotion-themeriver to show the changes in the number of Twitter comments containing different emotion messages by representing different emotions through colors.

In this paper, we use NLTK to pre-process the data and LDA to extract feature words in the description text of traditional crafts. We use Google Natural Language API [注2] based on the BERT model to get the emotional scores of description texts for traditional crafts. By using Tagcloud and ThemeRiver [8] to represent the feature words of traditional crafts and their emotion attributes and the word frequency changes of the feature words of traditional crafts in the temporal dimension.

3 Characteristic Extraction and Emotion Analysis of Traditional Crafts

3. 1 Pre-processing

In this section, we pre-processed the description texts. We extract the description text of "Traditional Crafts" provided by the Association for the Promotion of Traditional Craft Industries "Japan Traditional Crafts Aoyama Square," and use NLTK to divide the description texts by sentence and remove stopwords in the sentence.

3. 2 Characteristic Extraction

In this section, we perform characteristic extraction of craft description texts from six dimensions. Following preliminary studies, we examined the results of characteristic extraction from pre-processed description texts using NLTK, *TF-IDF*, and LDA in five dimensions. We concluded that LDA offers superior outcomes.

After using GENSIM to vectorize words in the pre-processed description texts, we used LDA to extract the topic words (topic: 1, #words: 20) and artificially categorize them into five dimensions: region, category, material, atmosphere, and manufacturing.

We constructed a dictionary using Wikipedia's Timeline of Japanese History as a reference and extracted the temporal dimension from the description texts.

Using the above characteristic extraction, traditional crafts are analyzed in six dimensions: region, category, material, atmosphere, manufacturing, and era. We show the classification results of six dimensions for three traditional crafts in Table 1. Taking **KYO Uchiwa** as an example, the six dimensions of the classification enable us to quickly learn about the basic characteristics of this traditional craft: a **traditional fan** made of **bamboo**, which was produced by **painting** and **rib** during the **Edo** period.

(注2)：https://cloud.google.com/natural-language

Table 1 The Classification Results of Six Dimensions for Three Traditional Crafts.

Dimension	KYO Uchiwa	SHIGARAKI Yaki	KYO Yuzen
Region	Kyoto	Shigaraki	Kyoto
Category	fan	ware, yaki	textile
Material	bamboo	pottery, kiln	cloth
Atmosphere	traditional	sabi, unique	pictorial
Manufacturing	rib, painting	firing	dyeing, silk
Era	Edo	Kamakura	Edo

3. 3 Emotion Analysis

In this section, we applied the Google Natural Language API to perform the emotion analysis on the historical description texts of crafts based on the BERT model. In the pre-processing section, we divided the historical description text of crafts based on sentences. So we calculate the emotion score of each sentence and then use the average score of all sentences to represent the emotion score of the traditional craft. The range of emotion scores was -1.0 to 1.0.

We defined sentences with absolute values of emotion scores below 0.1 as unemotional sentences, and because unemotional sentences affect the calculation of the total score, we removed the emotion scores of unemotional sentences when calculating the emotion scores of crafts. Meanwhile, the emotion attribute of traditional crafts is defined as positive and is shown as red when the emotion score is between 0.3 and 1.0; it is defined as neutral and shown as green when the emotion score is between -0.3 and 0.3; and it is defined as negative and shown as blue when the emotion score is between -1.0 and -0.3.

Figure 1 shows the color distribution of emotion attributes of traditional crafts.

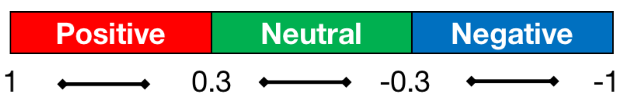

Figure 1 Colors and emotions.

4 Visualization of Traditional Crafts

4. 1 Visualization with Tagcloud

In this section, we generate Tagcloud utilizing the six dimensions of feature words that were derived in Section 3. 1, and we display the size and color of the feature words according to word frequency and emotion score. Furthermore, because the feature words themselves do not have emotion attributes, we calculate the emotion scores of the feature words by averaging the emotion scores of multiple sentences in which the feature words appear in the description texts. The formula for calculating the emotion score of a feature word is given in Eq. (1) as follows:

$$V_i = \frac{\sum_{i=1}^{n} V_u \ of \ S_u \ containing \ W_i}{\sum_{u=1}^{n} \ N \ of \ S_u} \tag{1}$$

where W_i stands for the feature word, S_u for the sentence in the description text that contains the feature word, N for the number of sentences that contains the feature word, and V_u for the emotion score of the sentences in the description text that contains the feature word W_i. The final emotion score for the feature word W_i is represented by the variable V_i. Additionally, the absolute value of V_u needs to be greater than 0.1.

4.2 Visualization with ThemeRiver

In this section, we created a ThemeRiver chart with a chronological horizontal axis and a word frequency vertical axis using ECharts(注3), an open-sourced, web-based, cross-platform framework that facilitates the quick building of interactive visualization. By showing the changes in word frequency of the characteristic words in different eras, the ThemeRiver chart reflects the temporal changes in traditional crafts in the region.

4.3 Visualization for Traditional Crafts

In this section, we use the Kinki region as an example to visualize traditional crafts. The interface of the visualization system is shown in Figures 2 and 3. The system consists of five parts: map, Tagcloud, images of traditional crafts, description texts of traditional crafts, and ThemeRiver.

(1) Using Leaflet(注4) and Google Maps API to present the traditional crafts on the map.

(2) Using WordCloud of the python library to create a tagcloud to show the six dimensions of feature words of traditional crafts.

(3) In the part of images of traditional crafts, three colors of red, green, and blue are used to represent emotional attributes in the edge of the images of traditional crafts.

(4) Extracting the history of the traditional crafts and presenting the description texts.

(5) Using ThemeRiver to show the change in word frequency of the feature words in different eras.

The images and description texts of traditional crafts of Hyogo Prefecture in Figures 2 and 3 are quoted from the Association for the Promotion of Traditional Craft Industries "Japan Traditional Crafts Aoyama Square"(注5). As shown in Figure 2, by entering the region, users can view all of the traditional crafts in Hyogo Prefecture. They can also browse the common words of the traditional crafts' feature words to gain a general idea of the traditional crafts and can comprehend the emotional attributes of regional traditional crafts and the development of related feature words in different eras. As shown in Figure 3, when the user chooses one of the traditional crafts, such as the Tamba Tachibana of Hyogo Prefecture, the image on the map is magnified, and the user read the Introduction and Tagcloud to find out more information about the traditional craft.

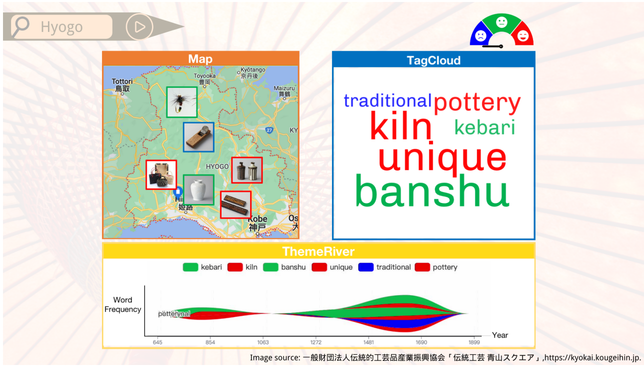

Figure 2 The visualization interface of Hyogo Prefecture.

Figure 3 The visualization interface of Tamba Tachikui.

5 Evaluation

5.1 Data Collection and Processing

We used the publicly available data of traditional crafts provided by the METI, and used the description text data of "Traditional Crafts" provided by the Association for the Promotion of Traditional Craft Industries "Japan Traditional Crafts Aoyama Square." From the data, we extracted 40 different types of traditional crafts in the Kinki region. Then, through characteristic extraction, 800 feature words were extracted and 339 feature words were classified manually. Through emotion analysis, we calculated the emotion scores of traditional crafts and the emotion scores of feature words. Finally, we developed a system for visualizing crafts.

(注3): https://echarts.apache.org/zh/index.html
(注4): https://leafletjs.com
(注5): https://kougeihin.jp/

Table 2 The results of the usability questionnaire by *SUS* scores

	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	Ave
SUS Score	70.8	79.2	83.3	83.3	68.8	79.2	83.3	79.2	83.3	66.7	77.7

5.2 Usability of Visualization System

In this section, we evaluated the usability of the Visualization system using the System Usability Scale (*SUS*) created by Bangor et al. [9]. Seven males and five females in their 20s served as subjects. Following a visualization demonstration of the system, subjects responded to the following questions.

Q1. I think that I would like to use this system frequently.

Q2. I found the system unnecessarily complex.

Q3. I thought the system was easy to use.

Q4. I think that I would need the support of a technical person to be able to use this system.

Q5. I found the various functions in this system were well integrated.

Q6. I thought there was too much inconsistency in this system.

Q7. I would imagine that most people would learn to use this system very quickly.

Q8. I found the system very cumbersome to use.

Q9. I felt very confident using the system.

Q10. I needed to learn a lot of things before I could get going with this system.

The System Usability Scale (*SUS*) was rated on a 5-point scale (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, and 5: strongly agree). We calculate the *SUS* score from the above questions. Firstly, each subject's rating of every positive question (Q1, Q3, Q5, Q7, Q9) minus 1 (rating-1), and each subject's rating of every negative question (Q2, Q4, Q6, Q8, Q10) is subtracted from 5 (5-rating) to adjust the score for each subject's question to make sure it is between 0 and 4. Secondly, we took the average subject's rating of each question and multiplied it by 25 to convert it to a *SUS* score scale from 0 to 100.

As shown in Table 2, the average score of all question items is 77.7. According to general guidelines on the *SUS* score, A) Excellent: >80.3, B) Good: >68.0-80.3, C) Okay: >68.0, D) Poor: >51.0-68.0, E) Awful: >51.0, given that the *SUS* score was "Good;" we could confirm that our proposed visualization system is very useful.

6 Conclusion

In this work, we developed a craft visualization system by characteristic extraction and emotion analysis of traditional crafts. Through the questionnaire, we received a "Good" rating from *SUS* scores, which indicates the high usability of the visualization system. In addition, we conducted a user satisfaction survey, and based on the results, we concluded that the visualization system received a high rating in terms of overall performance and functionality. However, it received a moderate rating in terms of demonstrating the emotional attributes of traditional crafts. We also requested feedback from the subjects. As a result, we will make improvements in the future.

In the future, we intend to expand the experimental dataset to include data from other regions as well as crafts from other countries to improve the applicability of the system and to improve the presentation of emotion attributes in the craft visualization system. Furthermore, we plan to explore the relationships between traditional crafts of Japan and present them in the visualization system.

Acknowledgment

This work was partially supported by JSPS KAKENHI Grant Numbers JP19H04118 and JP21K17862.

References

[1] Song, Z., and Wang. Y. (2023, March). "Visualization system for traditional crafts based on text mining with sentiment analysis." Proceedings of the 2023 11th International Conference on Information and Education Technology (to appear).

[2] Kowsari, K., Heidarysafa, M., Brown, D. E., Meimandi, K. J., and Barnes, L. E. (2018, April). "Rmdl: Random multimodel deep learning for classification." Proceedings of the 2nd International Conference on Information System and Data Mining (pp. 19-28).

[3] Jha, N. K. (2018, September). "An approach towards text to emoticon conversion and vice-versa using NLTK and WordNet." Proceedings of the 2018 2nd International Conference on Data Science and Business Analytics (pp. 161-166).

[4] Ahmed, M., Chen, Q., and Li, Z. (2020). "Constructing domain-dependent sentiment dictionary for sentiment analysis." Neural Computing and Applications, 32(18), 14719-14732.

[5] Tenney, I., Das, D., and Pavlick, E. (2019). "BERT rediscovers the classical NLP pipeline." arXiv preprint arXiv:1905.05950.

[6] Chen, Y. (2018, April). "TagNet: Toward tag-based sentiment analysis of large social media data." Proceedings of the 2018 IEEE Pacific Visualization Symposium (pp. 190-194).

[7] Brooks, M., Robinson, J. J., Torkildson, M. K., Hong, S. R., and Aragon, C. R. (2014, September). "Collaborative visual analysis of sentiment in Twitter events." Proceedings of the International Conference on Cooperative Design, Visualization and Engineering (pp. 1-8).

[8] Li, D., Mei, H., Shen, Y., Su, S., Zhang, W., Wang, J., Zu, M., and Chen, W. (2018). "ECharts: A declarative framework for rapid construction of web-based visualization." Visual Informatics, 2(2), 136-146.

[9] Bangor, A., Kortum, P. T., and Miller, J. T. (2008). "An empirical evaluation of the system usability scale." International Journal of Human–Computer Interaction, 24(6), 574-594.