DEIM2020 P1-12(day1 p77)

User Interests Extraction by Analyzing User Operations on Multimedia Museum Devices

Honoka KAKIMOTO†, Yuanyuan WANG††, Yukiko KAWAI†††, and Kazutoshi SUMIYA†

† Kwansei Gakuin University 2-1 Gakuen, Sanda-shi, Hyogo, 669-1337 Japan
†† Yamaguchi University 2-16-1 Tokiwadai, Ube-shi, Yamaguchi, 775-8611 Japan
††† Kyoto Sangyo University Kamigamo-motoyama, Kita-ku, Kyoto, 603-8555 Japan
††† Osaka University 5-1 Mihogaoka, Ibaraki-shi, Osaka, 567-0047 Japan
E-mail: †{dmi91695,sumiya}@kwansei.ac.jp, ††y.wang@yamaguchi-u.ac.jp, †††kawai@cc.kyoto-su.ac.jp

Abstract Nowadays, information on museum collections have been stored as digital archives, and virtual museums using digitized information have been provided. With this rapid digitization, the information that visitors can acquire about the museum exhibits has also diversified. For this reason, various learnings starting from the museum, such as the use of mobile devices in museums and applications for pre-learning. Also, in the recent education, interactive on-site learning is performed. Therefore, in the field of information engineering, the development of interactive learning systems in museums is very active. However, the existing learning support systems mainly focused on support for pre-learning and on-site learning, and there is not enough support to deepen the interests gained in on-site learning and lead to more advanced learning. Therefore, it is necessary to connect the interests gained at the museum to the next interests and more effective, scalable learnings. In this paper, we aim to extract learners' interests by analyzing learners' interactions for exhibits on museum devices. For this, we propose the scoring method based on four features of interaction data: keyword appearance frequency, keyword transition, media type, and media transition. Finally, we evaluate our proposed scoring method through a user study.
Key words Museum, museum education, user interaction

1. Introduction

Nowadays, information on museum collections have been stored as digital archives, and virtual museums using digitized information have been provided. With this rapid digitization, the information that visitors can acquire about the museum exhibits has also diversified. For this reason, various learning is starting from the museum, such as the use of mobile devices in museums and applications for pre-learning. Also, in the recent education, interactive on-site learning is performed. Information on exhibits is commonly used in education as supporting information of learning contents. Such information is provided on the online learning platform. Wu [1] studied the use of "iPalace Channel" (注1), the online learning platform of the National Palace Museum of Taiwan. The result of the study shows the effectiveness of the online content of exhibits. In particular, it is showed elementary and middle school teachers generally think that the video content of exhibits in the online learning platform is quite effective in teaching. Also, recent studies have shown the effectiveness of interactive learning using multimedia devices. Li et al. [2] proposed a multimedia interactive system for interactive learning of preschoolers. Their research showed that adoption of the multimedia interactive system and method based on preschool education make learners creative.

In our research, we define it as media in which learners can get more information about exhibits than images and text in section (2. 2). While learners explore the information of exhibits in class, there are three types of learning styles at the museum: pre-learning, on-site learning, and post-learning. The existing learning support systems mainly focused on support for pre-learning and on-site learning. There are some researches on approaches that connect learner's interests and features of exhibits on-site [3], while many existing learning support systems are based on the approach of estimating learner interests from the behavior in the museum. As one type of post-learning, Spence et.al [4] proposed mobile application that visitors exchange their experiences at the museum. However, there is not enough support to deepen the interests gained in on-site learning and lead to more advanced

(注1) : iPalace Channel, National Palace Museum, `https://ipalace.npm.edu.tw/\#`

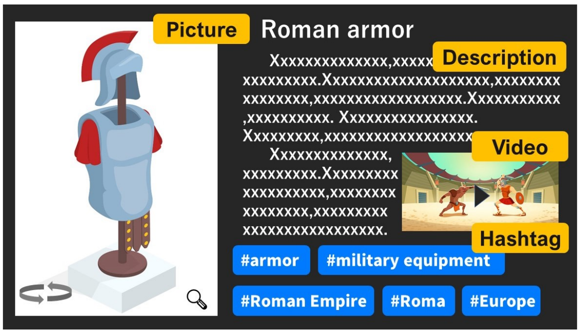

Figure 1 Example of the multimedia museum device.

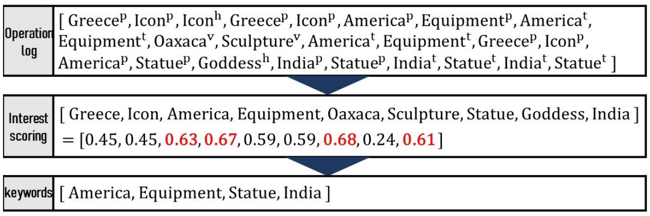

Figure 2 Example of interests extraction.

learning. Therefore, it is necessary to connect the interests gained at the museum to the next interests and more effective, scalable learnings.

In our research, we aim to extract user interests gained at the museum to provide post-learning content. Some researches treat learners' interests in a museum as dynamic. Hatala et.al [5] proposed an augmented audio reality system "ec(h)o", the knowledge-based recommender system for museums. It is linking the environment, interaction, objects and users at an abstract semantic level instead of at the content level. Our research also treats user interests as dynamic, and they are extracted from user interaction log data. In this paper, in order to extract learners' interests, we first analyze learners' interactions for exhibits on museum devices. Then, we propose the scoring method based on four features of interaction data: keyword appearance frequency, keyword transition, media type, and media transition. Also, we provide an interface to visualize learner's interests.

2. Extraction of Learners' Interests

2.1 Keywords Extraction from Operation Data

In this research, we extract learner's interactions on a multimedia device of a museum as their interests: tablet-type devices with direct access to videos, pictures, related links, etc as in Figure 1. The top part of Figure 2 presents an example of the extraction of keywords from operation log data. All nouns are extracted from words in hashtags, text, caption, title in operation log data of the museum device based on the morphological analysis. Then the keyword string is generated with nouns and media types. In this research, we extract four media types: hashtag, picture, text, and video.

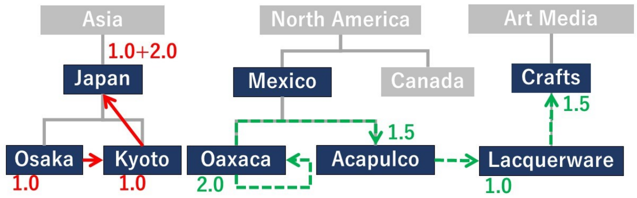

Figure 3 Example of scoring.

2.2 Interests Extraction by Scoring Methods

To determine keywords as learner's interests, we score on nouns in operation log data based on the following four scoring methods: keyword appearance frequency (f), keyword transitions on Wikipedia category structure (g), media type (h), and media type transitions (i). These scoring methods were set based on the following three conditions: nouns selected repeatedly by the learner, nouns selected consecutively by the learner, and nouns selected in multiple media by the learner.

a) f: keyword appearance frequency

The left part of Figure 3 shows an example of scoring based on keyword appearance frequency. Keyword appearance presents learner's interaction on keyword, and interaction on a keyword may indicate an interest in keywords that belong to a superordinate category. Therefore, the score is given to keywords in the superordinate category of the Wikipedia category structure when it exists in operation data.

b) g: keyword transition on Wikipedia structure

The scoring by keyword transitions is based on the number of transitions on the Wikipedia category structure, and the score of the first keyword is 0. The middle part of Figure 3 presents an example of scoring based on keyword transitions on the Wikipedia category structure. Since interaction on a keyword may indicate an interest in keywords that belong to a superordinate category, there are three types of score: transitions to keywords without superordinate category (1.0), keywords within the same superordinate category (1.5), the same keyword (2.0).

c) h : media type

The scoring by the media type is conducted based on a specific score of each type of media interacted by learners. This score was set as below based on the assumption of the amount of information that each media can provide to the learner and the viewing time. The scores of the hashtag, picture, text, and video are represented as follows: $h(1.0) < p(2.0) < t(3.0) < v(4.0)$.

d) i : media type transition

In this scoring method, switching to different media is scored in binary. When learners interact with the same media continuously, the score will be 0.

Interest score of a keyword $Interest_k$ is calculated by fol-

Table 1 Results of extraction

Respondents		Extracted keywords	Recall	Precision	F
1	A	America, bangle, Zuni, pincushion, Navajo, necklace, decoration, Venezuela, ornament, Solomon Islands, statue	**0.67**	0.75	0.71
	B	America, bangles, decoration	0.33	1.00	0.50
2	A	Morocco, candlestick, board, bow, Papua New Guinea, basket, Oceania, ritual, statue, Hawaiian Islands, jar, hook, living in the sea, Caroline Islands, Yap, Samoa, fishing basket	0.67	0.20	0.31
	B	Morocco, Oceania, Hawaiian Islands	0.67	0.67	0.67
3	A	Oaxaca, statue, Peru, calendar, North Korea, mask, South Korea, ritual, China, dynasty, Mali, clothes	**0.63**	0.42	0.50
	B	America, necklace, Zuni, Peru, calendar, North Korea, mask	0.38	0.43	0.40
4	A	America, tool, Oaxaca, statue, sculpture, goddess, India	0.67	0.33	0.44
	B	Greece, icon, America, tool, Oaxaca, statue, sculpture, goddess, India	1.00	0.50	0.67
5	A	America, Colombia, Pipe, Peru, collapsing tool, Bolivia, Mexico, Mask, Europe, USSR, dress, Model, Kazakh	**0.40**	0.17	0.24
	B	America, Mexico, Mask, Europe, USSR	0.20	0.20	0.20

Table 2 Results of Q1

Respondents		Answered keywords
1	C	America, bangles, Zuni, Navajo, necklace, decoration, Venezuela, Solomon Islands, statues
2		Papua New Guinea, Oceania, Hawaiian Islands
3		necklace, statue, Peru, mask, ritual, Nigeria, sculpture, Mali
4		Greece, America, sculpture
5		Guitar, Eating, Bolivia, Europe, dress

lowing equations. Firstly, the score a the keyword (k) is normalized as F_k. Then score of g, h and i of a keyword (k) at a point of operation log are normalized as $G_{k,p}$, $H_{k,p}$, and $H_{k,p}$. $Y_{k,p}$ is the normalized value of $X_{k,p}$, and $Interest_k$ is the Average value of the all values of an identical keyword (k) at all points of operation log. Then, keywords with a threshold of 0.50 or more are extracted as objects of interest for the learner.

$$X_{k,p} = F_k + G_{k,p} + H_{k,p} + I_{k,p} \tag{1}$$

$$Y_{k,p} = \frac{X_{k,p} - X_{min}}{X_{max} - X_{min}} \tag{2}$$

$$Interest_k = \frac{\sum_{i=1}^{n} Y_{k,i}}{f_k} \tag{3}$$

3. Evaluation

To evaluate the effectiveness of the proposed method, we compared the evaluation result of recall, precision and F value of keywords extracted by the proposed method (A) with the evaluation results based only on the keyword appearance frequency f (B). Evaluation data is the operation log data in which five respondents operated a multimedia device for three minutes and keywords extracted from the log data based on the proposed method. To evaluate the relationship between the intentions of the respondents and the tendency of the operations, the respondents answered the intentions of operations while they are operating the device. In the evaluation, the respondents watched the video of their operation, and then answered the questions below. In Q1, nouns extracted from the respondent's operation log were presented as a keyword list. In Q2, the respondent freely described the intention of the operation.

- Q1: Select keywords of your interest from the list.
- Q2: For what purpose did you interact with the app?

The results of Q1 shown in table 1 and table 2. The recall of keywords (A) extracted by the proposed method is high (bold part). This result indicates that it is possible to extract keywords indicating the learner's interest. In addition, the highlighted part in the table indicates that the keywords (A) have a higher F value than the keywords (B). In Q2, respondents 1, 3, and 5 answered that "I operated to learn more about the topic I was interested in.", "I operated to see the information of related exhibits based on pictures that I was interested in.". According to these results, it became clear that the interest extraction method in our research is effective for learners who want to know in detail about the exhibits and topics of interest. This result also shows the validity of the setting conditions of f, g, h, and i in the scoring method described in section 2.

On the other hand, respondent 2 answered that "First I operated to browse the objects of interest at random. After that, I operated to learn more about the exhibits I was interested in." in Q2. The F value of (A) was lower than the value of (B) in the result. The reason that the keywords of interest could not be accurately extracted from the log data based on the proposed method is probably due to the random interaction of respondent 2 whose purpose of the operation was not specific. Therefore, It was found that it was difficult to extract the interest of learners who interact with the information of the exhibits randomly based on the proposed method. Besides, respondent 4 also replied, "I operated to see the information on the exhibit I was interested in." However, the F value of (A) was low. This is probably because respondent 4 spent a long time browsing the media containing each word, and the number of keywords extracted from the log was extremely small. Therefore, It is considered that a more effective extraction of learner's interest can be achieved by extending the log acquisition time or incorporating the browsing time for each medium in the log as an additional scoring method.

Respondents who browsed in detail on a particular topic imply their interests in various keywords, and respondents who browsed randomly in various topics mainly imply their interests in location names or area names. therefore, when the target is limited to contents having characteristics of time and place, such as exhibits data of museums and art museums, post-learning based on the area and location is may be suitable for learners whose interests are not clear.

Regarding the media selection tendency, there was no difference in the type of media selected by respondents. They tend to select mainly images regardless of intention.

As a result, this evaluation revealed that the proposed method can extract interests that reflect the learner's intention to obtain detailed information on specific topics contained in the exhibits. Therefore, it is possible to extract a keyword of interest suitable for the post-learning for the learner to deepen the interest obtained in the on-site learning. Also, it was found that more effective interest extraction can be expected by incorporating the scoring method media browsing time in the future.

Our proposed scoring method focused on the hierarchical relationship of keywords in the category structure of Wikipedia. Interest extraction method based on Wikipedia category information has already been proposed in a research by Yang et al [3]. They extracted more conceptual and various interests of visitor of museum based on the analysis of the links of the Wikipedia pages that corresponded to the keywords. The problem with their proposed interest extraction method is that they extracted some keywords that are too general and conceptual. On the other hand, the keywords of interest extracted based on our method are limited to nouns included in the multimedia device in the museum,

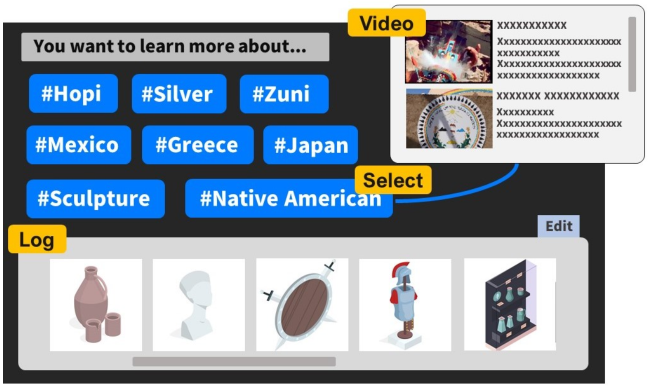

Figure 4 User interface.

and may not sufficiently express the conceptual interests of learners. Therefore, similar to the method of Yang et al. , it is necessary to extract not only keywords included in multimedia devices but also keywords of interest from Wikipedia pages. In future work, it may be possible to extract more conceptual interests from the upper categories of the pages corresponding to the keywords; based on the hierarchical relationship in the category structure of Wikipedia we focused on this research.

4. User Interface

This interface in Figure 4 is designed to visualize a learner's interests in exhibits and recommend video scenes for post-learning. The learner's operation log is displayed at the bottom of the interface, and keywords extracted by interest extraction from the log are displayed as tags. Then, the scene(s) related to the tag is recommended by selecting the tag(s).

To make the post-learning effective, it is necessary to extract a scene suitable for the post-learning, not the whole video. Also, the learner should be able to dynamically select the extracted scenes. Huh et al. [6] focused on hyperlinks in a system that recommends videos based on learner interest. They proposed a method of extracting the target of the video viewer's interest based on the ER (entity-relationship) model and creating a smart video in the e-learning domain. This smart video contains hyperlinks to other resources such as books, dictionaries, location information, people, and other videos, and these hyperlinks will be effective in deepening the interest and knowledge that the learner has gained in the museum's on-site education. Therefore, our proposed method extracts interest from learner's interaction with multimedia devices in the museum. Similar to the work of Huh et al. it is possible to generate hyperlinks based on the extracted interests. Since post-learning requires more advanced and advanced learning, it is necessary to link learners' interests with the topics they need to learn. However, creating links based only on learners' interests is not enough to support post-learning. The topics that learners should learn could be based on their learning history, their tendency to answer test questions, important topics in textbooks, etc. These methods will correspond to the Sequential-global continuum (how learners prefer to organize and progress toward understanding information) among the four dimensions in the learning style model of Felder and Silverman [7]. Sheeba et al. [8] proposed and evaluate the algorithm that determines the learning style of learners based on the Felder and Silverman model in the e-learning domain. Their evaluation method will be very effective in extracting and evaluating post-learning content.

Based on the evaluation results, our proposed method can classify learners into two types: learners of type A are learners who want to learn more about particular topics; they tend to repeatedly select the same topics, continuously select the same topics, and being interest in various topics. By combining the keywords extracted by the proposed method with the items to be learned, it will be able to help them learn about the topics they want to know in-depth. Type B are learners who are interested in topics at random. They change topics frequently and are less likely to return to the same topics. Also, they tend to be interested mainly in the location names and area names. Items to be learned and location names among keywords extracted by the proposed method should be used to extract post-learning content for them. However, the development of these methods will be excluded from our research topics as future works.

5. Conclusion

In this paper, we proposed a method to extract and visualize learners' interests from operation logs of multimedia devices in museums. As future work, we plan to propose a video content extraction method for post-learning support by analyzing the learner's learning styles based on classifications of important topics in textbooks and the tendency of the learner's operations. Bakkay et al. [9] developed a support protocol for the creation of educational video content for teachers without technical background. The goal of our research is to propose a dynamic scene extraction method using hyperlinks between video scenes based on an interest extraction method to recommend video content for post-education. If this goal is achieved, it is expected that teaching using educational video content will be easier for teachers and learners will be able to learn more scalable.

References

[1] Shaochun Wu. 2019. *Online Learning and Opinions of Educator: A Quantitative Study of Museum Educational Video Platform's User.* In Proceedings of the 8th International Conference on Informatics, Environment, Energy and Applications (IEEA '19). ACM, New York, NY, USA. pp 258-262. DOI:https://doi.org/10.1145/3323716.3323756

[2] Yan, Li, Ma Yan, and Feng Xu. 2017. *Multimedia Interactive System and Method Based on Preschool Education.*

[3] ZeHua Yang, Yusuke Yamamoto, Takehiro Yamamoto, Noriko Kando and Hiroaki Ohshima. 2019. *Discovering Method of Information Linking Museum Exhibits to Visitor Interests.* In Proceedings of the 11th Forum on Data Engineering and Information Management (DEIM'19). Nagasaki, Japan. E4-3.

[4] Jocelyn Spence, Benjamin Bedwell, Michelle Coleman, Steve Benford, Boriana N. Koleva, Matt Adams, Ju Row Farr, Nick Tandavanitj, and Anders Sundnes Løvlie. 2019. *Seeing with New Eyes: Designing for In-the-Wild Museum Gifting.* In Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19). ACM, New York, NY, USA. Paper 5, 13. DOI: https://doi.org/10.1145/3290605.3300235

[5] Marek Hatala and Ron Wakkary. 2005. *Ontology-Based User Modeling in an Augmented Audio Reality System for Museums.* User Modeling, Adaptation, and Personalization (UMAP'05). Volume 15. Issue 3–4. pp 339-380. Springer. DOI:https://doi.org/10.1007/s11257-005-2304-5.

[6] Seyoung Huh, Yoomi Park, onghyun Jang and Wan Choi. 2015. *Making a video smart for smart e-learning.* In Proceedings of the 2015 International Conference on Information and Communication Technology Convergence (ICTC'15). Jeju, Korea. pp. 858-863. DOI:https://doi.org/10.1109/ICTC.2015.7354686

[7] Richard M. Felder and Linda K. Silverman. 1988. *Learning and teaching styles in engineering education.* Engineering Education 78, 7. pp 674-681.

[8] T. Sheeba and Reshmy Krishnan. 2018. *Prediction of student learning style using modified decision tree algorithm in e-learning system.* In Proceedings of the 2018 International Conference on Data Science and Information Technology (DSIT ' 18). Association for Computing Machinery. New York, NY, USA. pp 85–90. DOI:https://doi.org/10.1145/3239283.3239319

[9] Mohamed Chafik Bakkay, Matthieu Pizenberg, Axel Carlier, E. Balavoine, Géraldine Morin, and Vincent Charvillat. 2019. *Protocols and Software for Simplified Educational Video Capture and Editing.* Journal of Computers in Education 6 (2) pp 257–76. DOI:https://doi.org/10.1007/s40692-019-00136-6