From textual to visual image searching:

User experience of advanced image search tool

Elina Late[1][0000-0002-3232-1365], Hille Ruotsalainen[1][0000-0002-2513-9223], Mert Seker[1] [0000-0002-3837-8688], Jenni Raitoharju[2] [0000-0003-4631-9298], Anssi Männistö[1][0009-0008-3264-6726] and Sanna Kumpulainen[1][0000-0002-7016-257X]

[1] Tampere University, Kalevantie 4, 33100 Tampere, Finland

[2] University of Jyväskylä, Seminaarinkatu 15, 40014 Jyväskylä, Finland

Abstract. This paper reports findings from a study focusing on user experience of image search tool utilizing content-based image retrieval methods. Previous studies have indicated challenges in textual image search especially in the historical domain. As a part of the project, a prototype tool was created for searching digitized historical images based on their visual contents to provide support for user needs identified in earlier studies. The tool was tested by 15 participants who evaluated their user experience with User Experience Scale and by verbal feedback. Our results indicate that participants derived benefits from the search capabilities provided by the tool, which went beyond relying on textual image descriptions. However, problems occurred, for example, in evaluating the search results and in user skills. Results also emphasize the value of intellectually produced metadata for image searching and use. Therefore, future developments should focus on creating hybrid systems supporting both textual and visual image searching.

Keywords: Content-based image retrieval, image archive, user experience

1 Introduction

Historical photographs form an important part of our cultural heritage capturing how the world looked like in the past. During the recent decades efforts have been put in digitizing photograph archives to make the contents available for various users. Indeed, digital image archives have become popular sources of historical information, for example, for scholars, information specialists, amateurs, and for the general public. For example, images are important primary sources for academic historians, and they are used for verification, documentation, or corroboration [1]. Although many digitized collections are openly available, access is often difficult because of the lack or incompleteness of image metadata [2,3]. However, textual metadata is vital since images are mostly searched using textual queries [4,2]. Yet, creating metadata manually is resource-consuming and challenging as the same image may have varying interpretations

depending on the user's viewpoint. Also, previous experiences have demonstrated that information needs in humanities research can be highly diverse, making it difficult to create a single unified metadata scheme. Therefore, flexible systems are needed [5].

Content-based image retrieval methods (CBIR) have been proposed as a solution to the problem. These methods enable the recognition of people, objects, events, and landscapes within images, all without relying on textual metadata. Another valuable application of CBIR is reverse image search, which allows users to find images by uploading a sample image as a query [6]. Novel methods are already widely available in commercial image search engines, but cultural heritage collections often lack such functionalities because of limited resources in their maintenance and development. As some studies have shown, users are longing for new image search possibilities [7] others have argued that users have conflicting attitudes and needs for automatic methods [8]. In general, users value possibilities for searching conceptual attributes by querying and browsing [9]. However, image use varies according to the user's task and profession [10,2]. Nevertheless, there is a gap in research in this respect and we do not yet know how the users of historical photograph archives benefit from the recent developments in the automatic query.

This paper aims to fill this gap in knowledge by evaluating the user experience of an image search tool based on CBIR. As a part of our research project, a prototype tool was created for advanced image searching utilizing computer vision methods and machine learning models to identify searchable contents from the images. Our test collection included historical photographs from the Second World War many of which lack original metadata. The prototype was tested by 15 users and user experience was measured using User Engagement Scale [11]. Additionally, user experiences were collected from verbal feedback during and after users tested the prototype.

Our research questions are:

RQ1. How satisfied are users with the advanced image search tool?

RQ2. What benefits and barriers do users see in content-based image retrieval?

Next, we will introduce our research setting with the description of the prototype tool and the data collection and analysis. Finally, we present the findings followed by discussion and conclusions.

2   Research setting

2.1   Advanced Image Search Tool

Advanced Image Search Tool (AIST) [13] was developed for improved access to digitized photographs. We tested AIST on photographs captured during the Second World War in Finland. The original collection (FWPA, sa-kuva.fi) provided by the Finnish Defense Forces contains in total almost 160.000 photographs captured by photographers who served in Information Company troops in 1939-1945. The search is based on textual metadata of the images that were mostly created during the wartime by the photographers. However, metadata is partly missing because of the chaotic times

during the photography. For our sample collection, we selected 23 800 images including 3800 images without any kind of original metadata or captions.

Based on the information collected during our previous studies [2,7,9], AIST was designed to provide an easy-to-use implementation for many aspects of Automatic Image Contact Extraction [6] by applying different computer vision methods and machine learning models trained on large publicly available datasets. AIST enables conducting search tasks by a graphical user interface and the tool is publicly available at GitHub [13]. AIST allows various automatic content-based search types ranging from low-level features, such as color distribution, to higher-level semantic information, such as environment or objects, using search options. As image archive users have emphasized the importance of analyzing people and objects from the images [7], several AIST search features are also related to people: their amount, age, gender, facial expression, and gaze direction. It is possible also to use images and text for querying. Different combinations of search features can be freely used.

2.2   Data collection and analysis

We invited in total 15 participants to test the prototype in May-June 2022. The participants were recruited partially from the previous interviews and partially through the contacts of the research group. The participants were either experienced users (N=8) of the original collection (researchers, museum curators, journalists, war history enthusiasts) or novices (history students N=7). The tests were audio and video recorded and the participants' consents were collected. One test session took on average 45 minutes.

The tests were conducted remotely via Zoom connection. The prototype was installed on the computer of the researcher and the participants used it via Zoom with "Ask for Remote Control" option. The users were asked to conduct five predefined tasks with AIST. The search tasks were formulated based on the actual searches that emerged in the previously collected interviews. This procedure followed the guidelines by Borlund [14]. The predefined tasks were used to ensure that all the participants were exposed to the different functionalities of the system. After completing the search tasks the respondents were asked to answer a short post-test questionnaire, which was based on the UES short form [11] to measure the user engagement in four factors. The scale consists of 12 statements evaluated with a five-point Likert scale; Strongly disagree, disagree, neither agree nor disagree, agree, strongly agree. We translated the UES into Finnish. We also added one question from the UES long form [11] to measure utilitarian achievement (UA) by asking to evaluate the success of the search task with the system. After completing the survey, the respondents were asked informally how they felt about using AIST and whether it would be useful for themselves.

We analyzed the data using SPSS and Atlas.ti. First, we created five computed variables to evaluate the user experience (FA, PU, AE, RW and UES total, see Fig. 1). Because some of the questions were negative and some positive, the scores were reversed if needed. UA was analyzed separately. We studied the correlations between UA and UES variables using Pearson bivariate correlation. Second, we uploaded the discussion transcripts into Atlas.ti where verbal expressions of user experiences were identified and coded. Quotes were further coded according to the categories used in UES

scale (FA, PU, AE and RW). Analyses were done by one researcher, but the codings were discussed in detail with another researcher in several rounds during the analyses process to reach a consensus.

3   Results

The image search tool gained an overall good evaluation by the test users with the user experience scale resulting an average 3.8 with 5 being the highest value. The scores of the four subscales varied (Fig. 1).
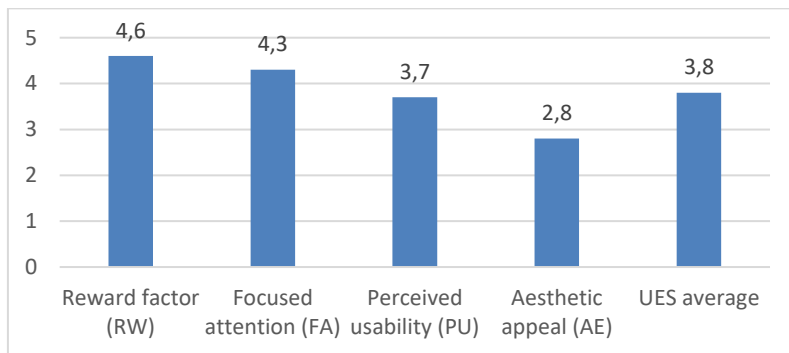

Figure 1. Mean scores from the 4 UES items and the UES score

Out of the five measures, the **Reward Factor** was scored the highest (mean 4.6). RW consists of three items measuring the experiences of success and reward when using the system. The scores show the users found the experience interesting. In their verbal feedback, the participants discussed the future possibilities of the tool and visioned the tool being even more rewarding for bigger collections. The participants described the tool as supportive, enabling them to overcome the shortcomings of the image metadata and access the images beyond the textual descriptions. They found the tool showing the full potential of the collection providing also more opportunities for research use such as data analysis.

Focused attention (FA) was measured with three items focusing on users' experiences absorbed in the interaction and losing track of time. The mean score for FA was 4.3. In their verbal comments, the users expressed feelings of happiness, excitement, and fun. These feelings were raised by discovering new photographs from the collection and realizing the potential of new methods for retrieving the images.

Perceived usability measured the negative affections experienced as a result of the interaction with the system and the degree of control. The mean score received for PU was 3.7. In the verbal feedback, various problems were brought up, many of them relating to unsuccessful searches and the lack of possibilities to evaluate the search results. Some users talked about the "black box" effect as they did not understand how the system produced the results. When collecting images used as research data, scholars had a need to understand what the search was based on. Searching images by visual contents demanded a new approach also from the users and users hoped for more

support and guidance from the system for making the searches. They did not know, for example, what words they should use for querying. Many participants agreed that the old and the new systems should be integrated to allow users to utilize the best features from both approaches (original metadata and content-based searching). Furthermore, users reminded that providing access to the images does not necessarily remove all the problems in using them. For example, using an image for illustrating a book requires trustworthy contextual information about the image. The tool cannot derive this information solely based on the image analyses alone.

Aesthetic appeal factor measures the attractiveness and visual appeal of the interface with three items. The mean score for AE was the lowest compared with other factors totaling 2.8. Indeed, in their verbal comments, users agreed that the visual appearance of the prototype was not aesthetically pleasing but at the same time adding that their expectations for not-for-profit services were not similar as for systems by big corporations. However, participants brought up that the visual design should support the user better, for example by selecting colors guiding the use.

Additionally, we asked if the users were able to find the images they were searching for with the system (UA). More than one fourth (26.7%) agreed and 60% partly agreed with the statement. The UA factor correlated significantly only with the PU factor (r= .577, p= .024). The users successful with searching had fewer negative experiences compared with those experiencing a lower rate of success.

4 Discussion and conclusions

The aim of our study was to analyze user experiences on a CBIR tool. As CBIR methods have been seen as a solution for problems of lacking metadata for textual searching of historical images, there is a lack of recent studies of the usefulness of such systems for the actual users [7,8]. As a part of our project, a prototype tool was created for searching images from historical image collections to provide support for user needs identified in earlier studies [2,7,9]. The tool was tested by 15 participants who evaluated their user experience. Our results indicate participants having high expectations for the tool but experiencing some difficulties when using it.

Our first research question was: How satisfied are users with the advanced image search tool? Overall, the study participants were very satisfied with the image search tool when evaluated by the UES. The aesthetic appeal of the tool was scored the lowest, although the users did not have high expectations for the prototype looks and the aesthetic appeal was not prioritized in the development. However, more studies with larger samples are needed to cover the variety of CBIR based tools to provide more reliable results of the user experience. Also, comparative studies on different user groups are needed as Beaudoin [8] observed differences in user needs. Nevertheless, this study provides a good starting point for future research.

Secondly, we asked: What benefits and barriers do users see in the content-based image retrieval? Our results show that CBIR has much to offer for searching the contents from historical image collections with limited metadata. Most participants were

excited about the possibilities of the novel methods and described such tools as being the "future". With the prototype tool the participants could already find images they had not found before from the collection. Indeed, earlier studies have showed that users desire CBIR methods and experience the lack and incompleteness of metadata as a major barrier to accessing the images [7]. CBIR systems may be helpful also for searching known items as before this has been frustrating for users lacking information of the specific image [2]. Another benefit of CBIR is overcoming the limitations caused by the language of the captions [12].

However, for professional use AIST should be further developed, evaluated and documented. Users value and expect transparency in use, ability to evaluate search results and clear guidelines for use. CBIR based tools demand new approaches also from the users. Before users have tried to imagine what words, the original photographer may have used for describing the image [7], but with the CBIR they need to learn to think about the contents of the image and how the tool might interpret them. Thus, future research should focus on search behaviors in real-life activities to find ways to support the information seekers with AI tools. Additionally, user training is needed.

Although new functionalities were appreciated, users want to also keep the features of the original search tool. Because, for example, the location, time and name of the photographer are among the most important access points for images [7], automatic metadata creation cannot totally replace the original metadata. Original captions also have their own value for image use in addition to accessing them [1,2]. Historians place significant importance on the trustworthiness associated with reputational institutions, such as archives, and the provenance of photographs when utilizing them for their research. They value original descriptive information, including captions, keywords, subject headings, the original medium of the photographs, and even details like the image size [1]. Our participants also reminded us that providing access to the images does not solve all the problems in using them. Many images lack metadata that is crucial for interpreting the contents. When gathering research data, scholars need information, for example, about the aboutness of data, characteristics of data, metadata, and secondary information about data [15] that CBIR is unable to produce. More metadata could be produced intellectually by crowdsourcing, i.e., allowing users to annotate contents directly and integrate knowledge from different sources into the collection.

Therefore, new features and search possibilities should be built on top of existing systems or earlier functionalities should be integrated with the new ones to create hybrid systems [8]. Different metadata types could be provided as layers on top of the original metadata and let the users decide which to use. Developing cultural heritage collections requires both financial and intellectual resources to ensure the continuation of the digital curation [16]. Collecting real-life user experiences and use practices of digital tools is crucial in future research to ensure their evidence-based development.

Acknowledgements

References

1. Chassanoff, A. M.: Historians' experiences using digitized archival photographs as evidence. The American Archivist 81(1): 135-164 (2018). https://doi.org/10.17723/0360-9081-81.1.135
2. Late, E., Ruotsalainen, H., Kumpulainen, S.: Searching images from open photograph archive. Search tactics and faced barriers in historical research. To be published in the International Journal on Digital Libraries, (2023).
3. Roberts, H. E.: A picture is worth a thousand words: Art indexing in electronic databases. Journal of the American Society for Information Science and Technology, 52(11): 911-916 (2001). https://doi.org/10.1002/asi.1145
4. Cho, H., Pham, M., Leonard, K.N., Urban A.C.: A systematic literature review on image information needs and behaviors. Journal of Documentation, 78(2), 207-227, (2022). https://doi.org/10.1108/JD-10-2020-0172
5. Lund, H., Bogers, T., Larsen, B., Lykke, M.: CHAOS: User-driven Development of a Metadata Scheme for Radio Broadcast Archives. In Proceedings of the iConference 2013, pp. 990-994 (2013). https://doi.org/10.9776/13510
6. Männistö, A., Seker, M., Iosifidis, A., Raitoharju, J.: Automatic Image Content Extraction: Operationalizing Machine Learning in Humanistic Photographic Studies of Large Visual Archives. arXiv:2204.02149, 2022.
7. Late, E., Ruotsalainen, H., Kumpulainen, S.: In a Perfect World. Exploring the Desires and Realities for Digitized Historical Image Archives. Accepted for publication in the proceedings of the ASIST conference, Wiley, London (2023).
8. Beaudoin, J.: Content-based image retrieval methods and professional image users. Journal of the Association for information science and technology, 67(2), 350-365, 2016. https://doi.org/10.1002/asi.23387
9. Kumpulainen, S., Ruotsalainen, H.: Searching Wartime Photograph Archive for Serious Leisure Purposes. In International Conference on Theory and Practice of Digital Libraries, pp. 81-92. Springer, Italy (2022). https://doi.org/10.1007/978-3-031-16802-4_7
10. Beaudoin, J.: A framework of image use among archaeologists, architects, art historians and artists. Journal of Documentation, 70(1): 119-147, 2014. https://doi.org/10.1108/JD-12-2012-0157
11. O'Brien, H. L, Cairns, P., Hall, M.: A practical approach to measuring user engagement with the refined user engagement scale (UES) and new UES

short form. International Journal of Human-Computer Studies, 112, 28-39 2018. https://doi.org/10.1016/j.ijhcs.2018.01.004

12. Menard, E., Khashman, N.: Image retrieval behaviours: users are leading the way to a new bilingual search interface. Library Hi Tech, 32(1), 50-68, 2014. https://doi.org/10.1108/LHT-06-2013-0067

13. GITHUB https://github.com/mertseker-dev/advanced_image_search_tool

14. Borlund, P: A study of the use of simulated work task situations in interactive information retrieval evaluations. Journal of Documentation, 72(3), 394–413, 2016. https://doi.org/10.1108/JD-06-2015-0068

15. Korkeamäki, L., Keskustalo, H., Kumpulainen, S.: Task information types related to data gathering in media studies. Journal of Documentation, 78(7): 528-545, 2022. https://doi.org/10.1108/JD-04-2022-0082

16. Barbuti, N.: From digital cultural heritage to digital culture: Evolution in digital humanities. In Proceedings of the 1st International Conference on Digital Tools & Uses Congress (DTUC'18). Association for Computing Machinery, New York, Article 21, pp. 1-3 (2018).