Understanding IIIF Image Usage Based on Server Log Analysis

Chifumi Nishioka (Kyoto University Library, Japan)

Kiyonori Nagasaki (The University of Tokyo, Japan)

Abstract. Numerous libraries and museums have adopted the International Image Interoperability Framework (IIIF) that promotes mutual use of images among different institutions. In a IIIF-compatible digital collection, images are retrieved via IIIF Image API, wherein regions of images can be specified. Thus, it is possible to investigate image usage in detail by examining the image regions that have been requested. In this article, we propose a log analysis method to measure the image usage and to visualize the analysis outcomes. Specifically, we employ the number of accesses to each pixel and visualize the results using heat maps. Given that a pixel is the smallest image unit, we apply herein a fine-grained analysis. The analysis method can be used in different applications including research tools in which researchers examine images and heat maps that show which regions of the images that have already been examined by collaborators.

1. Introduction

Staff in libraries and museums should understand how digital collections and their contents have been used for many reasons, including evaluation of digital collections and accountability of stakeholders. Analyzed results can be used to improve digital collections (Hughes, 2011). In general, quantitative usage analysis is conducted based on two steps: (1) selection of a measurement and (2) visualization of the result.

Numerous libraries and museums have adopted the International Image Interoperability Framework, pronounced as “Triple-Eye-Eff” (IIIF) (Snydman et al., 2015) that promotes mutual use of digital images. IIIF defines a couple of APIs to enable interoperable use of images. In IIIF-compatible digital collections, images are fetched via the IIIF Image API whose syntax is defined as:
{scheme}://{server}{/prefix}/{identifier}/{region}/{size}/{rotation}/{quality}.{format}.

Every time an image is zoomed and panned on an image viewer, different regions of the image are requested via IIIF Image APIs. Thus, detailed image usage is possible to investigate by examining the regions that have been requested.

In this paper, we present a method to analyze image usage on IIIF-compatible digital collections and to visualize the analyzed results. Specifically, we employ the number of accesses to each pixel and visualize them by heat maps. Given that the pixel is the smallest unit of an image, we enable a fine-grained analysis that is different from those used in previous studies (Warwick et al., 2008; Jones et al., 2000). As described in the syntax, digital collections and users can retrieve images and can specify different parameters, such as region, size, rotation, quality (e.g. color, gray, bitonal) and format (e.g. JPEG, TIFF, and PNG). In addition, IIIF Image API allows users to fetch the information of images, such as sizes and available formats. In this paper, we focus on regions among different parameters.

There are different applications associated with usage analysis. First, the visualization of the analysis result can facilitate collaborative research. Collaboration has become a hallmark of digital humanities (DH) research (Nowviskie 2012). The heat maps can depict the image regions that have already been examined by collaborators. Thus, researchers can identify regions that have not been investigated and work on these. Second, the usage analysis may facilitate transcription of cultural resources. To date, numerous transcription projects and platforms have been launched (Carletti et al., 2013). If a platform and target images are compatible with IIIF, it is possible to explore patterns, such as whether there is a difference in transcription performance (e.g. accuracy) between regions being zoomed and those not being zoomed. The patterns can facilitate the verification process for transcriptions.

The remainder of this paper is organized as follows. Section 2 describes related works. In Section 3, we propose a method to analyze the IIIF image usage and to visualize the analyzed result. Section 4 presents a couple of examples of analyses using server logs obtained from the Kyoto University Rare Materials Digital Archive and discuss improvements of the analysis method. In Section 5, we discuss possible use cases of the analyzed method. Section 6 discusses validity of server log analysis as well as the possible risks and concerns that must be considered when we employ the proposed method as a service. Finally, Section 7 concludes this paper.

2. Related Work

This section presents related studies regarding methods used for the investigation of the usage of digital collections and images. In general, the quantitative usage analysis is conducted according to the following two steps:

(1) Selection of a measurement: A measurement is chosen for usage analysis. The measurement is obtained from data, such as server logs. To date, measurements, such as the number of accesses to materials (e.g. books and manuscripts) and images, have been employed extensively (Jones et al., 2000)
(2) Visualization of results: Usage analysis outcomes are visualized to facilitate users to understand. Charts (e.g. line and bar charts) have been extensively employed

In the field of digital libraries, usage analysis had been conducted to evaluate collections and usability. For instance, Jones et al. (2000) calculated and reported the number of search queries per user session, number of visits to the digital library within a certain period, and number of browsed materials per search query to understand the search behavior on digital libraries. Hashemi et al. (2016) used the user's duration time on each exhibition to investigate to which extent the location of the exhibitions influenced user behaviors. Pääkkönen et al. (2015) used the number of clips (i.e. the number of annotations) and the number of keywords added to annotations to evaluate the usage of annotation and curation functions on digital collections. Luo et al. (2017) used page dwell time, mouse clicks, page re-visits, and the number of slider movements to understand user interests based on web server logs. Schultheiß et al. (2020) analyzed transaction logs to identify frequencies of known item searches and search tactics in library search systems. They used session durations and the number of search engine results page (SERP) sessions as measurements.

Nowadays, many studies are being conducted on usage analysis using an eye-tracker. An eye-tracker is a device that captures eye positions and eye movements. According to Holmqvist et al. (2011), there are 120 measures captured by eye-trackers that can be classified into four categories: movement, position, numerosity, and latency measures. Mokatren et al. (2016) developed a concept in which mobile eye-trackers are used to identify visitors' position and points-of-interest in museums. Jung et al. (2018) used a mobile eye-tracker to investigate children's interactions in a museum environment. They concluded that mobile eye-tracking allows a better understanding on the specific subpart a learner engaged visually among the multiple subparts of educational exhibits that would be very difficult to capture with other types of data. Dunst et al. (2017) reported that data captured by eye-trackers enabled the identification of candidate areas for objects to-be-annotated and identify types of objects with fairly satisfactory accuracy. For instance, areas with many fixations interspersed with short saccades are likely to contain text (Dunst et al., 2017).

In the works on digital libraries, tables, pie charts, and histograms have been adopted as methods to report and visualize the results of measurements. Administrators of digital libraries and digital collections have also used tables and charts to provide the results of measurements, such as the number of accesses. Studies using eye-trackers have adopted different visualization methods, such as heat maps and scanpaths (Holmqvist et al., 2011). A heat map visualizes the magnitude of a phenomenon in the form of color in two dimensions. The heart of the heat map is a color-shaded matrix display that has been used for more than a century (Wilkinson and Friendly 2009). A heat map is calculated based on different measures recorded by eye-trackers and shows where users tend to look. In contrast, a scanpath focuses on the temporal properties of eye positions and eye movements (Menges et al., 2020). It represents each fixation (i.e. gaze maintained on a single location) as a circle. The transition between two fixations is plotted as a connecting line.

In this study, we selected "the number of accesses to each pixel of an image" as a measurement and used the heat map as a visualization method for the measurement. We used "the number of accesses to each pixel of an image" as the pixel is the smallest image unit. There are similarities to other studies that used eye trackers in terms for the clarification of the usage of all image regions. However, in this study, we aimed to reveal the usage of each image region by only analyzing the server logs that have been recorded. Thus, in this study, we do not use any devices, such as eye-trackers. The purpose of the measurement and the visualization method is to understand detailed usage of images on IIIF-compatible digital collections. However, the results given by the analysis method are not limited to the purpose of understanding the detailed usage of images. As discussed in Section 5, the analysis method can be applied to different applications, including collaborative research tools and transcription platforms.

3. Analysis Method

This section describes methods to measure and visualize the detailed usage of images on IIIF-compatible digital collections and ways to display the results. Section 3.1 presents a method used to measure the number of accesses for each pixel of an image and for generating a heat map that visualizes the detailed usage of the image. In Section 3.2, we introduce a method to display the generated heat map along with the image to be analyzed, using Mirador, a IIIF-compatible image viewer.

3.1. Measurement and visualization

The method is comprised of the following two steps. Each process corresponds to the steps described in Section 2.

(1) Measuring the number of accesses of each pixel
For each image, an $H \times W$ matrix is generated wherein all the elements are zero. H and W are the height and width of the image in pixels, respectively. Thus, each element of the matrix corresponds to each pixel of the image. The height and width of images are retrieved by info.json[1] provided by the IIIF Image API. Subsequently, the requested images and regions are acquired by parsing the logs of the IIIF Image API. Based on the requested regions, the number of accesses to each pixel of each image is counted and recorded to the matrices.
In the IIIF Image API, "full" is used as the value of the region to request the entire region of an image. The value is used frequently, e.g. when requesting an image as a thumbnail. If the region is set to "full" in a server log, the server log has no influence on the generated heat map in (2) given that the value of one is added to all elements of the matrix. Hence, it is reasonable to filter out logs in which "full" is set as a requested region, and we filter out server logs where "full" is set as a requested region. However, if we display the number of accesses along with the colors of a heat map, filtering is inappropriate. Thus, it is necessary to decide whether we should filter out server logs with the requested region "full" according to the purpose of the analysis.

(2) Generating heat maps
After counting the number of accesses to each pixel, the result is output as a heat map. The RGB value of each pixel is calculated by considering the minimum and maximum values of the number of accesses to the pixel of an image. Regarding visualization methods, other methods, such as a bivariate histogram, can be applied. However, given that the heat maps have been used in numerous domains and contexts, and given that it is suitable to overlay them with the target image (refer to Section 3.2), we decided to employ heat maps.

The IIIF Presentation API allows the overlay of multiple images on one canvas that is an object that corresponds to a page. In this study, we counted the number of accesses and generated the heat map per image, rather than per canvas for simplicity.

3.2. Display of heat maps

This section introduces ways on how to display heat maps to facilitate users to understand the detailed usage of an image on a IIIF-compatible digital collection.

Users[2] can understand the detailed usage of an image by displaying the heat map along with the target image. The specification of IIIF Presentation API allows the overlay of multiple images. In practice, we edit IIIF manifests to overlay two images (i.e. the target image and heat map) in each canvas (i.e. page). A IIIF manifest is a JSON file in which metadata and material structures (e.g. books and manuscripts) are specified. We describe the image data to be displayed on each canvas in IIIF manifests. If a heat map is stored in a size smaller than the target image, it is necessary to specify that the heat map should be displayed at the same size as the target image. In addition, heat maps should be stored in a IIIF-compatible image server. Mirador, a popular viewer among the IIIF community, implements a function that allows the display of overlaid images, as shown in Fig. 1. One can manipulate the visibility and opacity for each image in the left-side panel.

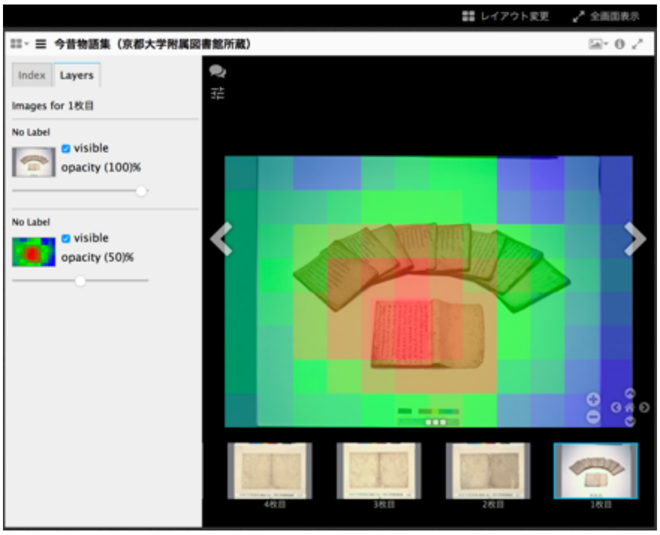

Figure 1: Overlay display of a heat map and its target image using Mirador. Photograph is courtesy of the Main Library of Kyoto University-*Konjaku monogatarishuu*.

4. Example and Improvement

This section illustrates examples of the analyzed results using server logs recorded on the Kyoto University Rare Materials Digital Archive and improvements of the log analysis method.

4.1. Analysis considering probabilities to be accessed

Fig. 2 illustrates a typical heat map that represents the image usage. The number of accesses close to the center is higher than that of other regions. This tendency has been observed in many other images. IIIF is oriented toward the visualization of high-resolution images. In IIIF, images of various resolutions were prepared for each image file, and they were divided into tiles.[3] Each time a user zooms in or out on an image in a viewer, the image is requested through the IIIF Image API depending on the size and region of the image. At this time, the viewer calculates the resolution and tiles requested according to the display size and regions of the image. Fig. 3 shows a heat map in which the regions to be accessed are randomly simulated. Fig. 3 is generated as follows:

(1) The size of a heat map is set as $H \times W$ pixels. An $H \times W$ matrix is prepared along with Section 3.1 (1). In the case of Fig. 3, we set H = 300 and Y = 400

(2) A pixel (x, y) is randomly selected from the range [1, X] and [1, Y], respectively

(3) Starting from the pixel selected in (2), the rectangle to be accessed is determined randomly. Specifically, a and b are randomly set from the range [$-1 \times x$, $(X - x)$] and [$-1 \times y$, $(Y - y)$]. The region surrounded by [x, y, a, b] is regarded as randomly accessed. Elements in the matrix corresponding to the region are incremented by unit

(4) Steps (2) and (3) are repeated 100,000 times

(5) The heat map is generated based on the matrix described in Section 3.1 (2)

As can be observed in Fig. 3, when browsing various regions with the viewer, the region close to the center of the image is likely to be included in the access area. Therefore, access is biased toward the center. When an image is browsed with an image viewer that allows zooming and panning of different image regions, the pixels close to the center are likely to be included in the requested regions. To treat each pixel equitably, it is necessary to adjust the number of accesses according to the access probability.

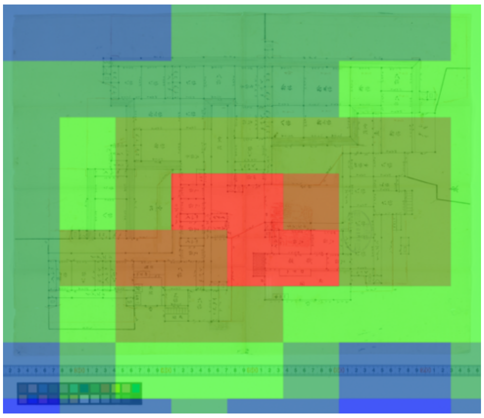

Figure 2: Typical heat map. Photograph courtesy of the Main Library of Kyoto University-*Yashiki-zu (design drawing of a mansion) from Nakai Collection*.

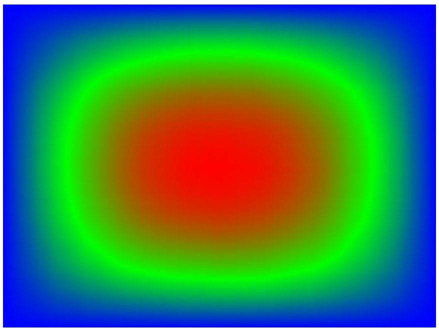

Figure 3: Heat map for an image wherein regions to be accessed are randomly chosen.

We compute a pixel's access probability that is located *a* and *b* pixels from the midpoint of each side of the image as

$$p(w,h,a,b) = \frac{w^2+2w-4a^2-1}{2w^2} \cdot \frac{h^2+2h-4b^2-1}{2h^2}.$$

Let $c(w, h, a, b)$ be the number of accesses of a pixel that is located *a* and *b* pixels from the midpoint of each side of the image. Taking the difference in probabilities to be accessed into account, the number of accesses can be adjusted by the following equation:

$$c_{rem}(w,h,a,b) = c(w,h,a,b) \cdot \log\left(\frac{w^2+2w-1}{w^2+2w-4a^2-1} \cdot \frac{h^2+2h-1}{h^2+2h-4b^2-1} + \alpha\right),$$

where $\frac{w^2+2w-1}{w^2+2w-4a^2-1} \cdot \frac{h^2+2h-1}{h^2+2h-4b^2-1}$ is the probability of a pixel access at the center of the image divided by the probability of a pixel access at a point that is located *a* and *b* pixels from the midpoint of each side. We use the logarithm to mitigate the influence from the ratio of probabilities. The parameter α controls the extent to which probability is taken into account. It is necessary as the adjusted number of accesses at the center becomes 0 because $\frac{w^2+2w-1}{w^2+2w-4a^2-1} \cdot \frac{h^2+2h-1}{h^2+2h-4b^2-1} = 0$ at the center. As a result, the number of accesses to pixels for Fig. 2 is adjusted as shown in Fig. 4.

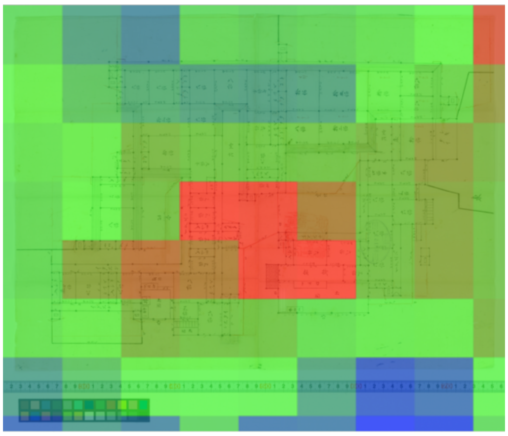

Figure 4: Heat map in which probabilities to be accessed are considered for Fig. 2. Photograph is courtesy of the Main Library of Kyoto University-*Yashiki-zu (design drawing of a mansion) from Nakai Collection*.

4.2. Analysis considering duration time

In this article, the analysis was conducted based on the assumption that the duration times of each access trial were equal. In practice, however, the duration time at each access trial was not equal. For example, the duration time for an access that is generated in the process of expanding the region on an image viewer will be extremely short. Conversely, the duration time will be longer when zoomed regions are looked. In general, the duration time of a certain page is calculated by the difference between the time at which a user accesses the page, and the time of the subsequent access. However, it is difficult to apply this general method to IIIF-compatible digital collections. When a large image on a IIIF-compatible digital collection is requested, the image is divided into small tiles (i.e. regions) that are accessed simultaneously. Therefore, if we apply the general method to calculate duration times, we can calculate a correct duration time only for the region that was associated with the last server log recording during the access. The duration time for other regions that were accessed simultaneously will be extremely short. It is difficult to determine this time accurately from the server logs irrespective of whether different accesses are simultaneous or not. For the determination, we should arbitrarily set some threshold regarding the time difference between consecutive server logs. In addition, some digital collections prefetch images and allow caches to enable immediate access to them. This decouples the request and duration times. Thus, it is necessary to consider how digital collections work to prefetch images and make caches when the duration time is taken into account.

We can obtain an accurate duration time by embedding a plug-in in an image viewer that measures the duration time of each region of each image. Of course, we need to obtain permission from users to use the plug-in and send the duration time. The validity of the duration time as an evaluation measurement has been shown in web-search systems (Kelly and Belkin, 2004; Fox et al., 2005) and recommender systems (Orad and Kim, 1998; Yi et al., 2014). For this reason, we would like to take the duration time into consideration for future work.

4.3. Referrer of images

As exemplified in Fig. 5, we observe images in which accesses are concentrated in specific regions. In general, when zooming into specific regions with an image viewer, regions around them are also accessed. Therefore, the regions are represented as a gradation in the heat map. Furthermore, image viewers usually have a limit on the enlargement ratio. Thus, it is not usual that only a small specific region is displayed on image viewer as shown in Fig. 5.

Consideration of referrers of access logs has indicated that these regions are referenced from the IIIF Curation Platform[4]. Given that IIIF enables mutual use of images, regions and images have more opportunities to be referenced from other organizations and platforms. By indicating the referrer, it is possible to show the motivation and background behind accesses. Furthermore, if the website that the referrer indicates is completely disclosed, it is possible to present a link to the website on a viewer in the form of an annotation. In this way, users can discover regions and images that are highly relevant.

Figure 5: Example wherein specific regions are extensively accessed. The photograph is courtesy of the Main Library of Kyoto University-*The story of Benkei, a tragic warrior.*

5. User Interactions and Possible Applications

This section describes user interactions with IIIF images and heat maps, and it lists possible applications of the results of the usage analysis.

5.1. User interaction with IIIF images and heat maps

Before the description of possible applications, we show how researchers interact with IIIF images and heat maps. We assume that heat maps are generated by (among others) an administrator of digital collections who has access to the server logs. The administrator is recommended to prepare a IIIF manifest that enables the display of a heat map over a target image, as described in Section 3.2.

Researchers can then interact with heat maps and images as follows,

(1) A researcher finds the IIIF manifest icon or URI of a record that includes an image of his/her interest. Many IIIF-compatible digital collections provide a IIIF manifest icon along with each record as shown in Fig. 6. The IIIF manifest icon provides a link to the IIIF manifest. The IIIF manifest URI can be obtained from the link destination of the IIIF manifest icon.

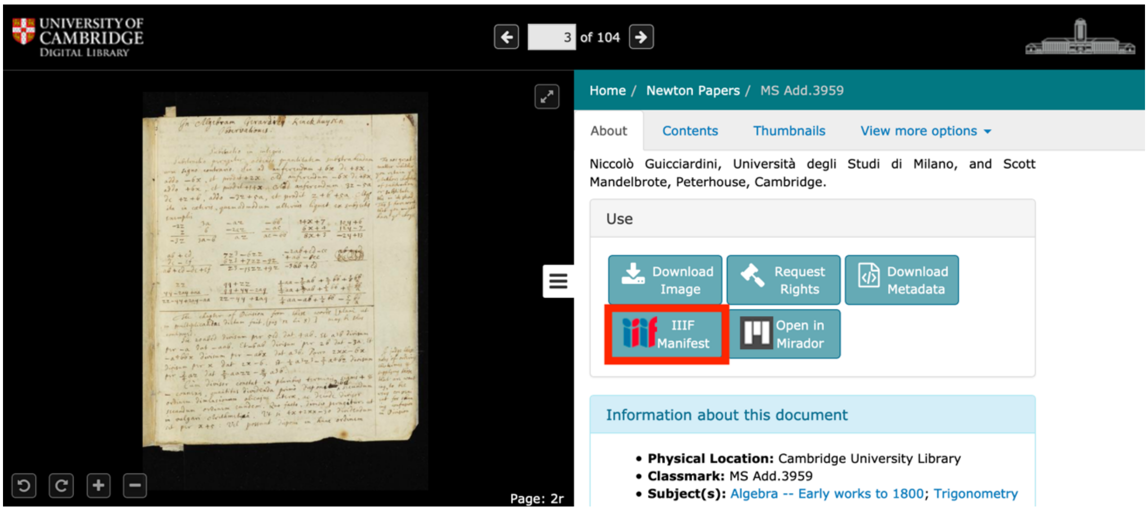

Figure 6: Example of a IIIF manifest icon in a digital collection. The photograph is from the Digital Library of the University of Cambridge–*Newton Papers, MS Add.3959.*

(2) A researcher opens Mirador (e.g. https://projectmirador.org/demo/), a IIIF image viewer, on his/her web browser. The researcher then imports the record to Mirador using the IIIF manifest icon or URI. As described by Winget (2016), a researcher can import the record by dragging and dropping the IIIF manifest icon directly into a workspace area of Mirador. Alternatively, he/she can import the record by navigating to the "replace options" at the upper left and pasting the URI into the box at the top right labeled "add new object from URL," as shown in the video provided by IIIF (2020).

(3) A researcher selects an image of his/her interest from the thumbnails displayed at the bottom of the viewer. He/she can observe the "layers" tab in which he/she can select the visibility state of the images (i.e. target image and heat map), as shown in Fig. 1. In addition, he/she can adjust the opacity of images with a bar.

5.2. Possible applications

This section describes possible applications of the outcomes of the usage analysis.

Collaborative research tool: Collaboration has become a hallmark of digital humanities (DH) research (Nowviskie 2012), with researchers either actively engaging with different parties or becoming increasingly aware that this is what they should be, or are at least expected to be doing (Griffin and Hayler, 2018). According to the European survey on scholarly practices and digital needs in the arts and humanities conducted by Costis et al. (2017), 71% of DH researchers answered that they "often or very often" collaborate with others on a research project. Until now, different research tools and infrastructures (Hinrichs and Krauwer 2014; Nagasaki et al., 2017; Wloka et al., 2013; Allori and Paltrinieri, 2020) have been developed to facilitate researchers to share resources and work for them collaboratively. In a collaborative research tool, researchers can share resources, such as images and texts. They can then explore the resources, and use them to conduct various work tasks, such as the addition of annotations and transcriptions to achieve a common goal.

Interoperability is an issue for sharing resources and collaboration. Green and Courtney (2015) pointed out interoperability as a critical need for digital collections. The data model used in IIIF follows the Web Annotation Data Model[5] that is recommended by W3C.

Therefore, IIIF facilitates the sharing of images and information accompanying images (e.g. annotations such as transcripts). For this reason, IIIF-compatible collaborative research tools have been developed (Sato and Ota, 2017; Nagasaki et al., 2017; Allori and Paltrinieri, 2020).

The analysis method shown in this article is expected to help understand which regions of images each researcher has looked at. By presenting heat maps, researchers can understand which regions of images have already been examined by collaborators. Thus, collaborators can see regions that have not been investigated to work on them. In addition, a researcher can see contexts of, for example, annotations made by a collaborator as heat maps show regions of images investigated by him/her.

Transcription Platform: Numerous transcription projects and platforms have been launched in the past (Carletti et al., 2013). Transcribers zoom and pan images during the generation of transcriptions. If a platform is compatible with IIIF, it is possible to verify a pattern, that is, ascertain whether there is a difference in transcription performance (e.g. accuracy) among regions that are zoomed and those that are not zoomed. If we find a pattern, we can facilitate the verification process for transcriptions.

In addition, as the heat maps show popular regions in images, transcribers can find regions that should be transcribed or annotated as priority. This contributes to meeting the demands of users.

Selection of thumbnails: In many cases, images displayed on the first page of materials are used as thumbnails. However, the first image does not necessarily represent the material. We may select the most-viewed regions of images in the material as a thumbnail. The most-viewed regions can be revealed by the analysis method presented herein.

Understanding the research process: It is important for researchers to reflect and review their research process to make their research efficient. Researchers can reflect and review their research process by looking into heat maps that show how they have investigated images. In addition, it is possible for students and young researchers to understand and learn how experienced researchers conduct research by exemplifying the details of their image usage using heat maps.

6. Discussion

This section discusses validity of server log analysis and the possible risks and concerns that should be carefully considered when employing the proposed method as a service.

6.1. Validity

In this article, we assume that heat maps generated based on server logs of a IIIF image server reflect user attention. In this section, we look at works that investigate how cultural resources (e.g. paintings) are looked at using eye-trackers and discuss validity of our assumption by confirming whether heat maps generated based on server logs correlate with the findings of these works.

Bailey-Ross et al. (2019) explored the viewing behaviors of different participants to examine whether the accompanying written context influences how digital reproductions are experienced, using eye-trackers. They revealed that the majority of first fixations are on the face. According to Bailey-Ross et al. (2019), previous studies have also showed that the viewer's gaze is focused predominantly on the human figure in particular on human faces, independently of contextual elements also depicted in the image (Ro et al., 2007; Massaro et al., 2012; Villaniet al., 2015). Figure 7 shows an image usage heat map of a photograph that includes six human figures. We see that upper regions including human faces received more accesses compared to other parts such as feet and shoes, which correlates with observations gained from the previous studies. However, it cannot reveal focused regions as finely as eye-trackers.

Figure 7: Example of an image heat map wherein human figures, especially regions around faces, are more frequently accessed. The photograph is courtesy of the Main Library of Kyoto University-*Six warriors.*

According to Teeuwes (2010), eye movements are the result of the visual (bottom-up) input reaching the eye and the expectations and intentions of the observer (top-down goals) (Walker et al., 2017). Bottom-up attentional processes are stimulus driven and automatic, directly determined by the physical properties of the environment (Walker et al., 2017). On the other hand, top-down attentional processes are determined by goals, intentions, and interpretation of the observer (Walker et al., 2017). Walker et al. (2017) investigated the influence of bottom-up and top-down attentional processes on participants' gaze behavior, using eye-trackers in the Vincent Van Gogh Museum. They observed that both children and adults spent more time at regions mentioned in the background information. In addition, Bailey-Ross et al. (2019) also observed written interpretation of paintings redirect the gaze toward areas of conceptual significance and away from faces. Thus, viewers focus on different regions depending on contexts such as background information and motivation.

Although server logs of IIIF image servers do not record users' contexts of accesses of images very well, we may infer them by looking into referrers of server logs. For example, if we look into server logs whose referrer is "Minna de Honkoku" (Hashimoto et al., 2018), a transcription platform for Japanese cultural resources, we can assume that regions that include characters are accessed more frequently compared to other regions. Figure 8 is a heat map of an image that was generated from server logs whose referrer is "Minna de Honkoku" (i.e., http://honkoku.org/), which supports the assumption.

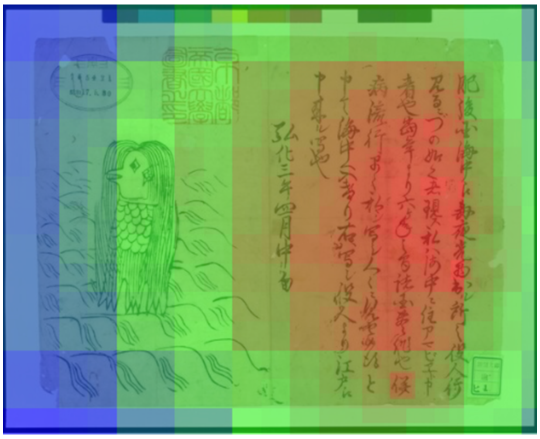

Figure 8: Example of heat map based on server logs whose referrer is a transcription platform, wherein characters are more frequently accessed. The photograph is courtesy of the Main Library of Kyoto University-*Amabie.*

From the above, it can be concluded that the analysis method presented in this paper cannot clarify regions that users pay attention to as finely as eye-trackers; however, it is advantageous in that it does not require installation of any equipment.

6.2. Challenges

Visualization of access logs is not a problem if anonymization is conducted appropriately. However, anonymization does not make sense in some cases (e.g. images with a small number of accesses). For instance, for images that are only accessed by researchers in a specific field, colleagues can easily guess who accessed images and regions. Even if the anonymization is complete, a series of activities of a researcher on images may reveal his/her viewpoint that would be a key issue of his/her academic outcome. In this case, his/her priority of the research discovery may be infringed. Therefore, careful consideration will be necessary to exploit the analysis result as a service.

These challenges sometimes occur in the context of text databases and web-search platforms. Therefore, we would like to tackle the challenge, by referring to guidelines pertaining to the handling of user logs on these websites.

As noted in Section 6.1, we assume that heat maps generated based on server logs of a IIIF image server accurately reflect user behavior and research process. In the future, we would like to verify this assumption by comparing the heat maps based on the server logs and data recorded by eye-trackers as well as conducting interviews with researchers. Therefore, when using heat maps in an application, it is necessary to reflect feedback from a researcher himself/herself about browsing regions.

7. Conclusion

This article presents a method to analyze image usage and to visualize the analyzed results. In particular, we count the number of accesses of each pixel and visualize the result in the form of heat maps generated based on the number of these accesses. Heat maps enable users to understand the detailed usage of images, such as the regions of images that users are more likely to access. Although the analysis method cannot capture focused regions as finely as eye-trackers, it is advantageous in that it does not require installation of any equipment. As possible applications, we list a collaborative research tool and a transcription platform. In the future, we would like to consider in detail the various challenges related to the anonymity of accesses and priority rights of the research discovery.

Funding

This research was supported, in part, by the JSPS KAKENHI [JP18K13235 and JP20K20132].

References

Allori, L. and Paltrinieri, C. (2020). A Collaborative Workspace for Archival Research: MIA and the EURONEWS Project. Proceedings of Digital Humanities 2020 (DH2020), (online) available from https://dh2020.adho.org/wp-content/uploads/2020/07/548_ACollaborativeWorkspaceforArchivalResearchMIAandtheEURONEWSProject.html

Awh, E., Belopolsky, A. V., and Theeuwes, J. (2012). Top-down versus Bottom-up Attentional Control: A Failed Theoretical Dichotomy. Trends in Cognitive Sciences, 16(8), 437-443. doi: 10.1016/j.tics.2012.06.010

Bailey-Ross, C., Beresford, A. M., Smith, D. T., and Warwick, C. (2019). Aesthetic Appreciation and Spanish Art: Insights from Eye-tracking. Digital Scholarship in the Humanities, 34(1): i17–i35. doi:10.1093/llc/fqz027

Carletti, L., McAuley, D., Price, D., Giannachi, G., and Benford, S. (2013). Digital Humanities and Crowdsourcing: An Exploration. Proceedings of Museums and the Web, (online), available from https://mw2013.museumsandtheweb.com/proposals/digital-humanities-and-crowdsourcing-an-exploration/.

Dunst, A., Hartel, R., and Laubrock, J. (2017). The Graphic Narrative Corpus (GNC): Design, Annotation, and Analysis for the Digital Humanities. Proceedings of 14th IAPR International Conference on Document Analysis and Recognition (ICDAR). IEEE, 3: 15–20. doi: 10.1109/ICDAR.2017.286.

Fox, S., Karnawat, K., Mydland, M., Dumais, S., and White, T. (2005). Evaluating Implicit Measures to Improve Web Search. ACM Transactions on Information Systems (TOIS), 23(2): 147–168.

Green, H. E. and Courtney, A. (2015). Beyond the Scanned Image: A Needs Assessment of Scholarly Users of Digital Collections. College & Research Libraries, 76(5): 690–707. doi: 10.5860/crl.76.5.690.

Griffin, G. and Hayler, M. S. (2018). Collaboration in Digital Humanities Research – Persisting Silences. Digital Humanities Quarterly, 12(1). http://www.digitalhumanities.org/dhq/vol/12/1/000351/000351.html.

Hashemi, S. H., Hupperetz, W., Kamps, J., and van der Vaart, M. (2016). Effects of Position and Time Bias on Understanding Onsite Users' Behavior. Proceedings of the 2016 ACM on Conference on Human Information Interaction and Retrieval, ACM, pp. 277–280.

Hashimoto, Y., Kano, Y., Nakasnishi, I., Ohmura, J., Odagi Y., Hattori, K., Amano, T., Kuba, T., and Sakai, H. (2018). Minna de Honkoku: Learning-driven Crowdsourced Transcription of Pre-modern Japanese Earthquake Records. Proceedings of Digital Humanities

2018 (DH2018), (online) available from https://dh2018.adho.org/en/minna-de-honkoku-learning-driven-crowdsourced-transcription-of-%E2%80%A8pre-modern-japanese-earthquake-records/

Hinrichs, E. and Krauwer, S. (2014) The CLARIN Research Infrastructure: Resources and Tools for e-Humanities Scholars. Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC-2014). European Language Resources Association (ELRA), pp. 1525–1531.

Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Halszka. J., and van de Weijer J. (2011). Eye Tracking: A Comprehensive Guide to Methods and Measures. Oxford University Press, United Kingdom.

IIIF (2020). Import Manifest to Mirador. YouTube, (online), available from https://www.youtube.com/watch?v=R33IJtil8tg

Jones, S., Cunningham, S. J., McNab, R., and Boddie, S. (2000). A Transaction Log Analysis of a Digital Library. International Journal on Digital Libraries, 3(2): 152–69. doi: 10.1007/s007999900022.

Jung, Y. J., Zimmerman, H. T., and Pérez-Edgar, K. (2018) A Methodological Case Study with Mobile Eye-Tracking of Child Interaction in a Science Museum. TechTrends, 62: 509–517. doi: 10.1007/s11528-018-0310-9.

Kelly, D. and Belkin, N. J. (2004). Display Time as Implicit Feedback: Understanding Task Effects. Proceedings of the 27th annual international ACM SIGIR Conference on Research and Development in Information Retrieval, ACM, pp. 377–384.

Luo, X., Wang, J., Shen, Q., Wang, J., and Qi, Q. (2017). User Behavior Analysis Based on User Interest by Web Log Mining. Proceedings of the 2017 27th International Telecommunication Networks and Applications Conference (ITNAC). doi: 10.1109/ATNAC.2017.8215435.

Massaro, D., Savazzi, F., Di Dio, C., Freedberg, D., Gallese, V., Gilli, G., and Marchetti, A. (2012). When Art Moves the Eyes: A Behavioral and Eye-Tracking Study. PLOS One, 7(5), e37285. doi: 10.1371/journal.pone.0037285

Menges, R., Kramer, S., Hill, S., Nisslmueller, M., Kumar, C., and Staab, S. (2020). A Visualization Tool for Eye Tracking Data Analysis in the Web. Symposium on Eye Tracking Research and Applications. Association for Computing Machinery, pp. 1–5. doi: /10.1145/3379156.3391831.

Mokatren, M., Kuflik, T., and Shimshoni, I. (2016). Using Eye-Tracking for Enhancing the Museum Visit Experience. Proceedings of the International Working Conference on Advanced Visual Interfaces (AVI '16). Association for Computing Machinery, pp. 330–331. doi: 10.1145/2909132.2926060.

Nagasaki, K., Tsuda, T., Yang, X. J., Kitazaki, Y., Muller, A. C., and Shimoda, M. (2017) A Collaborative Approach between Art History and Literature via IIIF. Proceedings of Digital Humanities 2017 (DH2017), (online) available from https://dh2017.adho.org/abstracts/185/185.pdf.

Nowviskie, B. (2012). Evaluating Collaborative Digital Scholarship (or, Where Credit is Due). Journal of Digital Humanities 1/4 (Winter). http://journalofdigitalhumanities.org/1-4/evaluating-collaborative-digital-scholarship-by-bethany-nowviskie/.

Oard, D. W. and Kim, J. (1998). Implicit Feedback for Recommender Systems. Proceedings of the AAAI Workshop on Recommender Systems, AAAI, pp. 81–83.

Pääkkönen, T. (2015). Crowdsourcing Metrics of Digital Collections, LIBER Quarterly, 25(2): 41-55.

Ro, T., Friggel, A., and Lavie, N. (2007). Attentional Biases for Faces and Body Parts. Visual Cognition, 15(3), 322–348. doi: 10.1080/13506280600590434

Sato, M. and Ota, I. (2017). Collaboration System based on Crowdsourcing with Mirador - Proposal of a System to Support Analysis and Theory in Collaborative Research of Humanities. SIG Technical Reports Computer and Humanities, 2017-CH-114(7): 1–6.

Schultheiß, S., Linhart, A., Behnert, C., Rulik, I., and Lewandowski, D. (2020). Known-item Searches and Search Tactics in Library Search Systems: Results from Four Transaction Log Analysis Studies. The Journal of Academic Librarianship, 46(5): 102202. doi: doi.org/10.1016/j.acalib.2020.102202.

Snydman, S., Sanderson, R., and Cramer, T. (2015). The International Image Interoperability Framework (IIIF): A Community and Technology Approach for Web-based Images. Proceedings of the Archiving Conference. Society for Imaging Science and Technology, pp. 16–21.

Theeuwes, J. (2010). Top–down and Bottom–up Control of Visual Selection. Acta Psychologica, 135(2), 77–99. doi: 10.1016/j.actpsy.2010.02.006

Villani, D., Morganti, F., Cipresso, P., Ruggi, S., Riva, G., and Gilli, G. (2015). Visual Exploration Patterns Of Human Figures in Action: An Eye Tracker Study with Art Paintings. Frontiers in Psychology, 6, 1636. doi: 10.3389/fpsyg.2015.01636

Walker, F., Bucker, B., Anderson, N. C., Schreij, D., and Theeuwes, J. (2017). Looking at paintings in the Vincent Van Gogh Museum: Eye movement patterns of children and adults. PLOS One, 12(6), e0178912. doi: 10.1371/journal.pone.0178912.

Warwick, C., Terras, M., Huntington, P., and Pappa, N. (2008). If You Build It Will They Come? The LAIRAH Study: Quantifying the Use of Online Resources in the Arts and Humanities through Statistical Analysis of User Log Data. Literary and Linguistic Computing, 23(1): 85–102. doi:10.1093/llc/fqm045.

Wilkinson, L. and Friendly, M. (2009). The History of the Cluster Heat Map. The American Statistician, 63(2): 179–184. doi: 10.1198/tas.2009.0033.

Winget, D. (2016). Create and Share IIIF Items Quickly and Easily with Drag and Drop over Email, (online) available from https://medium.com/@aeschylus/create-and-share-iiif-items-quickly-and-easily-with-drag-and-drop-over-email-879f13c9caba.

Wloka, B., Winiwarter, W., and Budin, G. (2013). DASISH: An Initiative for a European Data Humanities Infrastructure. Proceedings of International Conference on Information Integration and Web-based Applications & Services (IIWAS '13). Association for Computing Machinery, pp. 433–437. doi: 10.1145/2539150.2539237.

Yi, X., Hong, L., Zhong, E., Liu, N. N., and Rajan, S. (2014). Beyond Clicks: Dwell Time for Personalization. Proceedings of the 8th ACM Conference on Recommender Systems, ACM, pp. 113–120.

---

[1] https://iiif.io/api/image/2.1/#image-information (accessed on 2019-11-03).

[2] In this case, “user” refers to a person who uses the analysis result (e.g. an administrator of a digital collection).

[3] https://iipimage.sourceforge.io/documentation/images/ (accessed on 2020-10-27).

[4] http://codh.rois.ac.jp/icp/index.html.en (accessed on 2019-11-03).
[5] https://www.w3.org/TR/annotation-model/ (accessed on 2018-11-03).