Virtual Reality Lifelog Explorer

Lifelog Search Challenge at ACM ICMR 2018

Aaron Duane
Insight Centre for Data Analytics
Dublin, Ireland
aaron.duane@insight-centre.org

Cathal Gurrin
Insight Centre for Data Analytics
Dublin, Ireland
cathal.gurrin@insight-centre.org

Wolfgang Huerst
Utrecht University
Netherlands, Netherlands
huerst@uu.nl

ABSTRACT

The Lifelog Search Challenge (LSC) invites researchers to share their prototypes for lifelog exploration and retrieval and encourages competition to evaluate effective methodologies for this. In this paper. we present a novel approach to visual lifelog exploration using a virtual reality (VR) platform. Findings from our initial experiments with known-item search from lifelog data have motivated us to build a retrieval engine for virtual reality that uses visual concepts automatically extracted from the lifelog visual data as the basis for it's filtering mechanism.

ACM Reference Format:
Aaron Duane, Cathal Gurrin, and Wolfgang Huerst. 2018. Virtual Reality Lifelog Explorer: Lifelog Search Challenge at ACM ICMR 2018. In *LSC'18: The Lifelog Search Challenge, June 11, 2018, Yokohama, Japan.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3210539.3210544

1 INTRODUCTION

Lifelogging interfaces and applications have appeared on a wide range of prevailing platforms such as laptops, tablets and phones[13] and continuing research is constantly enhancing the user experience of these platforms. Virtual reality, however, has yet to be properly considered as a legitimate platform for lifelog exploration or retrieval.

Our research to date has focused on examining the potential of virtual reality to support intuitive interaction with lifelog archives and we have developed initial prototypes to evaluate the feasibility of exploring lifelogs in a virtual reality environment[4]. This work is motivated by our belief that virtual reality platforms will become more lightweight and ubiquitous. In this paper we describe the most recent iteration of this prototype which is designed to support the interactive known-item search tasks of the Lifelog Search Challenge at ACM ICMR 2018. For this research, we used a HTC Vive virtual reality headset and accompanying wireless controllers, which is a market-leading device in 2018.

LSC'18, June 11, 2018, Yokohama, Japan

ACM ISBN 978-1-4503-5796-8/18/06.
https://doi.org/10.1145/3210539.3210544

2 BACKGROUND

Effectively addressing the retrieval of lifelog documents (images, tweets, emails, etc.) has seen much attention in the lifelogging community in recent history. This can be clearly observed in conferences such as NTCIR[6] where researchers from a host of countries develop tools to effectively retrieve lifelog documents in datasets released by the conference organisers. Some of this work has focused on enhancing the performance of the visual concept detectors to be used for retrieval[10] whereas other work has focused on a purely textual approach[9]. Other notable research has utilised long, descriptive paragraphs of text to annotate the lifelog content, as opposed to the conventional automatic tag-based approach[11] and some even created an interactive systems which employed a semantic content-tagging mechanism[1]. Event segmentation[2] has also been a popular method of enhancing the retrieval process. However, to date no submissions to an NTCIR conference has ever considered virtual reality as a platform for these methods of lifelog retrieval.

Though there has been almost no research explicitly targeting the exploration of lifelogs in virtual reality, there has been many applications developed for the platform that facilitate elements of exploring and examining life experiences. One obvious example is the playback of 360-degree video which is considerably more immersive when viewed in virtual reality and is especially so when the footage is recorded from a more familiar first-person perspective. This evolution of immersion within virtual reality extends to many interaction methodologies that would better facilitate lifelog exploration. This is not to suggest that explicit examples of lifelog interaction in virtual reality do not already exist. For example, an art installation by Alan Kwan titled 'Bad Trip'[8] was developed in 2012 which enables users to explore a manifestation of the creator's mind and life experience within virtual reality.

3 DATASET

The LSC dataset that was released to accompany the search challenge consists of 27 days of multi-modal lifelog data from one active lifelogger. The dataset is based on the NTCIR-13 Lifelog dataset[6], but it is enhanced and compacted to represent one month of detailed user activity. Although the dataset contains many different sources of time-aligned sensor data, our focus in this prototype was on the visual lifelog data from wearable cameras (about 1,500 per day).

In the dataset, each image was accompanied by the output of a state-of-the-art computer vision concept detector, thereby providing a listing of real-world concepts (e.g. computer, car, coffee, etc.) for each image. We used these concepts as the primary source of data for supporting interactive search. Additionally, we included the date metadata to allow temporal filtering of the search space. We note that additional metadata (such as activities, locations, etc.) were also available, but they were not integrated into this prototype.

4 USER INTERACTION

The virtual reality lifelog explorer developed for the LSC has two components, each of which needed to be optimised for the VR environment. The querying component was a virtual interface designed to provide a quick and efficient means for a user to generate a filter query within the VR system. While there are many approaches that one could take to input queries, a decision was made to focus on gesture-based interaction, as opposed to other forms of interaction such as voice-controlled. To the best of our knowledge, this is the first lifelog interaction mechanism that has been developed for a VR environment.

The gesture-based querying interface consists of two sub-menus, one for selecting lifelog concepts of interest and the second for selecting the temporal aspect of the query (e.g. hours of the day or days of the week). A typical query to the system, such as 'using the computer on a Saturday afternoon' would require the user to use the concept sub-menu to select the appropriate visual descriptors (e.g. computer or laptop) and the temporal sub-menu to select the time range (afternoon) and the day of the week (Saturday). The user then hits the submit button and the query is executed and the result is displayed for the user to browse. The concept sub-menu is shown in Figure 1 and the temporal sub-menu is shown in Figure 2.

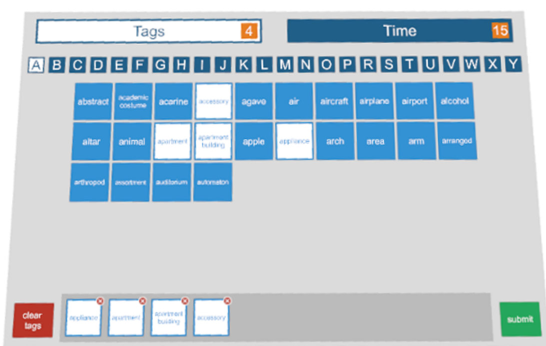

Figure 1: Concept Filter

This querying interface is available for the user to bring up at any time by pressing a dedicated button on either of the two VR handsets available with the HTC Vive. When the user submits their query, the interface disappears and the user is free to explore/browse the results inside the virtual environment.

The lifelog concepts that populate the concept sub-menu represent the original lifelog search challenge concepts that accompanied the dataset release; no additional computer vision algorithm was implemented. The concepts were divided into pages corresponding to their first letter and organised alphabetically on each page from left to right. The user can select no concepts or anywhere up to a maximum of 10 concepts per filter query. In our experimentation, no user has ever selected ten concepts, so this is a reasonable upper-bound for the current work. The temporal sub-menu presents the user with the 7 days of the week and the 24 hours of the day. These days and hours can be selected in any combination to generate a temporal filter on the search results.

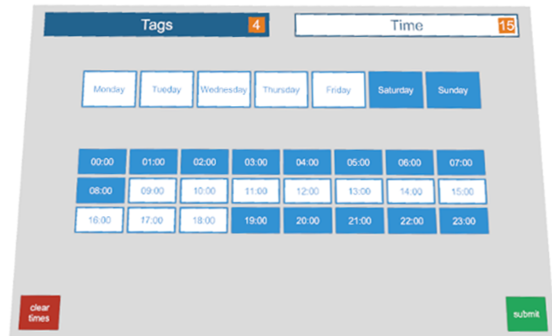

Figure 2: Day/Hour Filter

Part of the research into developing a prototype for visual lifelog exploration in a virtual environment is to identify the most efficient and preferred methods of interacting with that environment's user interface. The relative infancy of virtual reality as a modern platform means that there is not a clear answer for how to best interact with a user interface in this context. There are no well defined and understood interaction best practices to integrate (e.g. point-and-click in the desktop environment, or sweep-a-finger in a touchscreen environment). Without such norms, we developed two approaches to interacting with the our querying interface.

4.1 Distance-Based Interaction

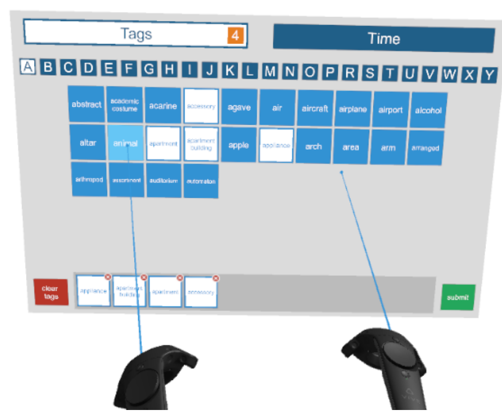

Figure 3: Distance-Based User Interaction

The distance-based approach utilises interactive beams which originate at the the top of the user's wireless controllers. These beams are projected when the controllers are pointed at any relevant interface in the virtual environment and directly interact with that interface's elements (see Figure 3). This method of interaction is comparable to a lean-back style of lifelog browsing[7] and is functionally similar to using a television remote or other such device. Pressing a button on the controllers selects the concept or time-range that is being pointed at. Naturally, it is possible to use both hands to select concepts in parallel, should a sufficiently dexterous user be generating queries.

4.2 Contact-Based Interaction

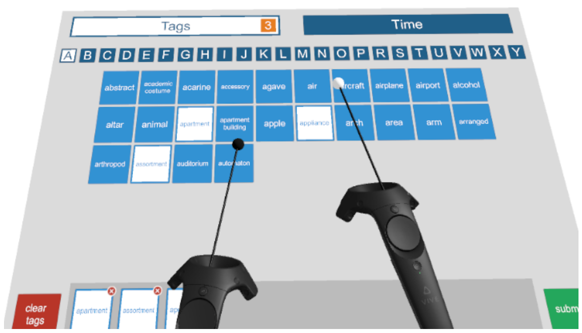

Figure 4: Contact-Based User Interaction

The contact-based approach utilises a much more direct form of interaction where the user must physically touch the interface elements with their controllers. To facilitate this process, the controllers are outfitted with a drumstick-like object protruding from the head of each controller (see Figure 4). This object was added to enhance precision and fidelity when contacting interface elements. This method of interaction is reminiscent of a more conventional style of lifelog browsing where the controller drumsticks mimic how our fingers interact with a keyboard or touchscreen. Tactile feedback is provided through the hand-controllers to signify hitting the keys.

These virtual reality user interface interaction methodologies are based on real-world analogues (television, keyboard, touchscreen, etc.) and can be observed in various forms in industry standard virtual reality applications such as the HTC Vive's main menu[12] or Google's popular Tilt Brush interface[5]. To date we have not observed a clear user preference for one interaction methodology over another. There has also been no notable drawback in terms of user querying efficiency across multiple interaction types[3].

5 LIFELOG DATA BROWSING

As previously stated, after a filter query is submitted to the system, the querying interface disappears and the user is presented with the highly-ranked filtered images (see Figure 5) in decreasing rank order. These images are ranked using a combination of concept relevance and the time of capture (maintaining the temporal organisation of the data), where concept relevance takes precedence over the temporal arrangement. For example, if the user creates a filter query containing 3 different concepts, then images containing all 3 concepts will be ranked first in the list, followed by images containing 2, and then 1. When multiple images contain the same amount of relevant concepts, then those images are ranked temporally according to the image capture time.

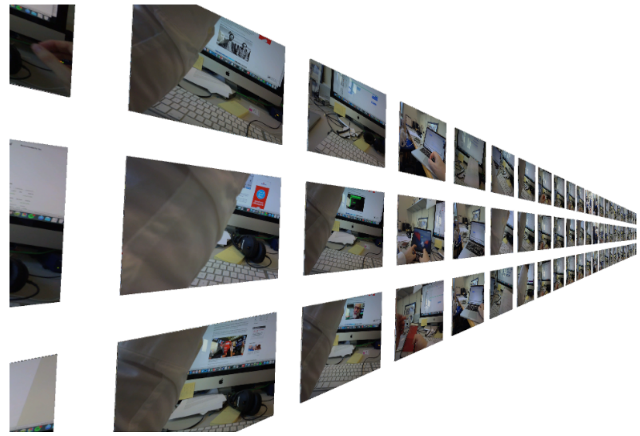

Figure 5: A VR Ranked List

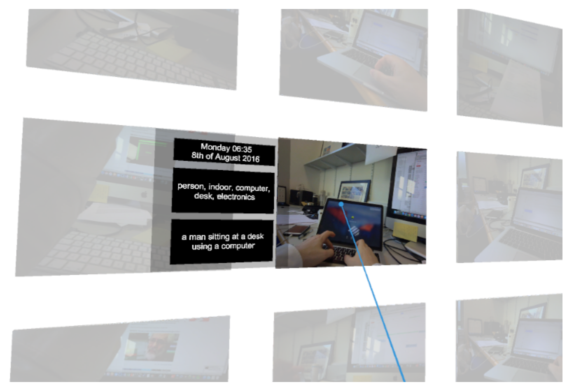

Figure 6: Image Metadata

Any image displayed on the VR ranked list can be selected for further exploration by pointing the user's controller at it and pressing a button. This displays additional metadata about the image such as the specific capture date and time and what concepts have been detected (see Figure 6). Additional filtering options are also made available along with this metadata. For example, the user can choose to see other images contained in the manually annotated event this image was labelled under. Finally, the user has the option of simply viewing all the images captured before and after the target image within a specific timespan. A summary of the process is shown in Figure 7.

6 INITIAL FINDINGS

Preliminary investigations into the LSC dataset using the virtual reality lifelog explorer have been promising. There are obvious drawbacks where a topic relies heavily on computer usage or biometric data as our system does not utilise these parts of the dataset. Yet oftentimes there is sufficient information exposed later in the topic outline to effectively locate a relevant section of time in the dataset, described accurately by the concept description.

We have yet to observe any noteworthy increase in querying efficiency upon moving from a conventional platform to a virtual

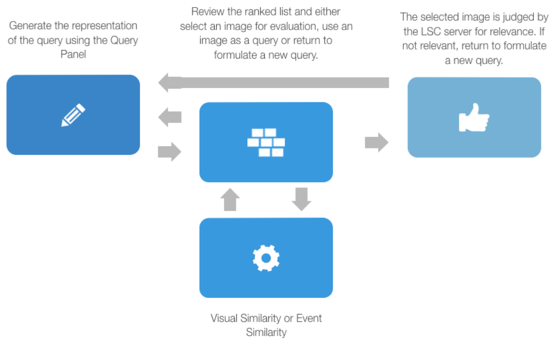

Figure 7: Query Process Flow

reality platform, but we have also noted no significant drawback either. Considering the relative infancy of virtual reality in this context, these observations have encouraged us to continue our research into developing the platform further.

7 CONCLUSION

In this paper we present our research into virtual reality as a potential candidate for lifelog exploration and retrieval and describe the most recent prototype we have developed for this purpose. The dataset released for the Lifelog Search Challenge 2018 contained a diverse range of data topics such as captured images (via wearable camera), concepts (via computer vision), physical activity (via phone gyroscope), computer usage (via keylogger) and biometrics (heart rate, steps taken, etc). Our prototype was developed to primarily target lifelog concepts as we felt it was sufficient to accomplish adequate lifelog retrieval and compete with other systems in this context.

REFERENCES

[1] Gabriel DeOliveira-Barra, X Giró-i Nieto, Alejandro Cartas-Ayala, and Petia Radeva. 2016. LEMoRe: A Lifelog Engine for Moments Retrieval at the NTCIR-Lifelog LSAT Task. *The 12th NTCIR Conference, Evaluation of Information Access Technologies* (2016), 366–371. https://imatge.upc.edu/web/sites/default/files/pub/cdeOliveira-Barra.pdf

[2] Aiden R. Doherty, Chris J. a. Moulin, and Alan F. Smeaton. 2011. Automatically assisting human memory: A SenseCam browser. *Memory* 19, 7 (2011), 785–795. https://doi.org/10.1080/09658211.2010.509732

[3] Aaron Duane and Cathal Gurrin. 2017. Pilot Study to Investigate Feasibility of Visual Lifelog Exploration in Virtual Reality. *LTA '17 Proceedings of the 2nd Workshop on Lifelogging Tools and Applications* (2017).

[4] Aaron Duane and Cathal Gurrin. 2018. *Lifelog Exploration Prototype in Virtual Reality.* 24th International Conference on Multimedia Modeling.

[5] Google. 2018. Tilt Brush. (Jan. 2018). Retrieved January 1, 2018 from https://www.tiltbrush.com

[6] Cathal Gurrin, Hideo Joho, Frank Hopfgartner, Liting Zhou, Rashmi Gupta, Rami Albatal, and Duc-Tien Dang-Nguyen. 2017. Overview of NTCIR-13 Lifelog-2 Task. *13th NTCIR Conference on Evaluation of Information Access Technologies* (2017).

[7] Cathal Gurrin, Hyowon Lee, Niamh Caprani, Zhenxing Zhang, Noel O 'connor, and Denise Carthy. 2010. Browsing Large Personal Multimedia Archives in a Lean- back Environment. *International Conference on Multimedia Modeling* (2010).

[8] Alan Kwan. 2012. Bad Trip. (2012). https://www.kwanalan.com/blank

[9] Hsiang-Lun Lin and Liang-Pu Chen. 2016. Image Searching by Events with Deep Learning for NTCIR-12. *The 12th NTCIR Conference, Evaluation of Information Access Technologies* (2016), 372–374. http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings12/pdf/ntcir/LIFELOG/04-NTCIR12-LIFELOG-LinH.pdf

[10] Bahjat Safadi, Philippe Mulhem, Georges Qu, F Grenoble, and F Grenoble. 2016. LIG-MRIM at NTCIR-12 Lifelog Semantic Access Task. *The 12th NTCIR Conference, Evaluation of Information Access Technologies* (2016), 361–365. https://pdfs.semanticscholar.org/fbf4/416b8e4846c5069e4f4c6f6d16d783605bd2.pdf

[11] Harrisen Scells, Guido Zuccon, and Kirsty Kitto. 2016. QUT at the NTCIR Lifelog Semantic Access Task. *The 12th NTCIR Conference, Evaluation of Information Access Technologies* (2016), 375–377. http://research.nii.ac.jp/ntcir/workshop/OnlineProceedings12/pdf/ntcir/LIFELOG/05-NTCIR12-LIFELOG-ScellsH.pdf

[12] Valve. 2018. HTC Vive SteamVR. (Jan. 2018). Retrieved January 1, 2018 from http://store.steampowered.com/steamvr

[13] Yang Yang, Hyowon Lee, and Cathal Gurrin. 2013. Visualizing lifelog data for different interaction platforms. *CHI '13 Extended Abstracts on Human Factors in Computing Systems on - CHI EA '13* (2013), 1785. https://doi.org/10.1145/2468356.2468676