Pilot Study to Investigate Feasibility of Visual Lifelog Exploration in Virtual Reality

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

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

ABSTRACT

The prevalence of modern technology has enabled people to record a digital trove of life experiences and these datasets continue to grow exponentially day by day. Exploring these huge datasets effectively has been the subject of much research in the lifelogging community. In this paper we describe a pilot study performed to investigate the feasibility of using virtual reality as a platform for exploring visual lifelog data. The dataset used in this experiment consisted of image data captured by wearable cameras and keywords describing the visual content of each image. The results of this experiment suggested there was no notable reduction in user performance when using the virtual reality platform, compared to a conventional desktop environment. This research was performed as part of a larger study to investigate the overall potential of virtual reality as a platform for visual lifelog exploration.

KEYWORDS

Lifelog, Lifelogging, Virtual Reality

ACM Reference format:
Aaron Duane and Cathal Gurrin. 2017. Pilot Study to Investigate Feasibility of Visual Lifelog Exploration in Virtual Reality. In *Proceedings of LTA'17, Mountain View, CA, USA, October 23, 2017,* 4 pages.
https://doi.org/10.1145/3133202.3133208

1 INTRODUCTION

Dodge and Kitchin[1] refer to lifelogging as "a form of pervasive computing, consisting of a unified digital record of the totality of an individual's experiences, captured multimodally through digital sensors and stored permanently as a personal multimedia archive" and the prevalence of modern technology and sensors has enabled people to capture this digital trove of life experiences automatically and continuously with newfound ease and efficiency. However, this automated and passive collection of data has produced very large datasets that are often difficult to disseminate or explore. Attempts to derive these semantics typically manifest in the form of graphic user interfaces that rely on additional technologies such as machine learning image classifications. There has been significant development in this domain in recent history as the importance of intuitive visual access to lifelogs becomes increasingly apparent. Lifelogging interfaces are appearing across all prevailing platforms such as laptops, tablets and phones[7] and continuing research is constantly optimising the user experience on these systems.

However, virtual reality, which has seen a recent resurgence in popularity due to advancements in technology, has gone largely unexplored. When we consider the complex and multifaceted nature of lifelog datasets in the context of virtual reality's multidimensional axes of exploration, there are numerous applications we can consider. For example, the SAS Institute[6] states that we are limited to processing less than 1 kilobit of information per second when reading from a screen yet the human optic nerve has an estimated bandwidth of about 8 megabits per second. In virtual reality this limit is significantly increased as we harness multiple axes and depths of information and are able to fully utilise our periphery vision. An argument against virtual reality often cites the expense of the platform as well as how long it takes to setup the hardware versus a conventional platform. However, we consider these issues as problems stemming from the current generation of virtual reality platforms rather than the medium itself. As technology naturally advances and hardware prices decrease, it is feasible to assume virtual reality platforms will become as lightweight and easy to use as a laptop or tablet. In this paper we outline a pilot study to identify the feasibility of using virtual reality in this context. This experiment is part of a larger study to investigate the overall potential of virtual reality as tool for visual lifelog exploration.

2 DATA

Suitable lifelogging datasets are often difficult to source in the research community due to data protection and other privacy concerns. This is because photographs are a common part of visual lifelogs and, if not correctly blurred, can contain sensitive information such as people's faces. Fortunately, the research challenges posed by lifelogging has been addressed by both NTCIR[4] and ImageCLEF[3] in recent years and we now note the availability of appropriate datasets. The datasets utilised in the pilot study discussed in this paper were all sourced from the NTCIR-12 Lifelog retrieval challenge [5] in 2016 where they were released as part of a series of research tasks aimed at exposing new and innovative methods for scrutinising lifelog data. We chose these datasets to standardise our research methodology and enable us to compare our work with other lifelog analysts.

The primary dataset[2] we used consisted of 3 months of continuously captured image data generated by 3 lifeloggers using wearable cameras (approximately 1 month per lifelogger). The dataset was further enriched via an automatic image classification algorithm

LTA'17, October 23, 2017, Mountain View, CA, USA

ACM ISBN 978-1-4503-5503-2/17/10…$15.00
https://doi.org/10.1145/3133202.3133208

which described the content of each image using keywords (e.g. laptop, car, phone, etc.). Alongside this dataset a number of topics were released to facilitate the lifelog tasks outlined by NTCIR. One of these tasks was referred to as the Lifelog Semantic Access Task (LSAT) which was comparable to a typical 'known-item search' task. It aimed to have participants develop a system to retrieve a number of specific moments in a lifelogger's life which were defined as semantic events or activities that happen throughout a typical day (e.g. hailing a taxi, eating breakfast, etc.). There was a total of 48 LSAT topics released for this dataset.

Using these topics, we constructed an experiment wherein a user is described a semantic event and then they must search through the entire dataset and locate at least one image which was representative of that event. To expedite this process, we decided to focus on only one month of data generated by one lifelogger, which reduced the 48 LSAT topics down to 24. This number was further reduced when it was determined that not all of the LSAT topics were compatible with the nature of our experiment. This is because several of the topics contained images flagged as positive results for a semantic event even though an indication of that event was not strictly present. For example, an event where the lifelogger is known to be eating a salad, but the salad is not visible in every image of the event. Instead of searching through every topic and culling images not compatible with our experiment, we decided to target 15 of the 24 topics where this issue did not arise. This had no impact on our research as 15 topics was sufficient to perform our intended pilot study.

3 EXPERIMENT

Three lifelog exploration prototypes were developed for this pilot study, one designed on a conventional platform (PC) and two designed on a virtual reality platform (HTC Vive). The conventional platform was created as a control to compare typical lifelog interactions and how effectively they translated to a virtual reality environment. We created two prototypes in virtual reality, based on two prevailing mainstream user interaction paradigms, to compare their effectiveness and ease of use. Each prototype targeted the same dataset of lifelog images and the same LSAT topics. As stated previously in this paper, the task on each platform was to identify at least one image captured during a semantic event which was described to the user. To accomplish this task, each prototype relied on generating filter queries containing a date, a time and a selection of keywords. Each query was generated using three separate interface screens: a date picker, an hour picker and a keyword picker. Once the query was submitted, the user was presented with a selection of images they could navigate which matched their filter query.

On the conventional platform, the user constructed each query using a standard mouse and web interface (see Figure 1). On the virtual reality platform, the user constructed each query using wireless controllers and a virtual interface that exists in 3D space. This interface differs slightly on each of the virtual reality prototypes we developed. In one version, the menu is attached to a wireless controller and the user interacts with it using the opposing controller, similar to how one might interact with a clipboard. For this reason we referred to this interface as the 'clipboard' style menu

Figure 1: Conventional Lifelog Prototype - Mouse

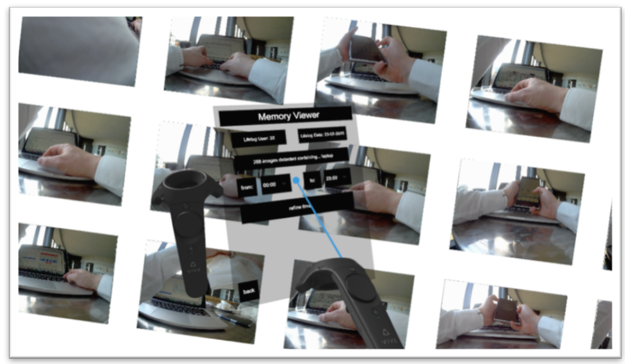

Figure 2: VR Lifelog Prototype - Clipboard

Figure 3: VR Lifelog Prototype - Billboard

(see Figure 2). In the other virtual reality prototype, the menu is twice it's normal size and hovers in space a set distance in front of the user. This menu can be re-positioned at any time and is still interacted with in the same fashion as the previous menu, just at a further distance. We referred to this interface as the 'billboard' style menu (see Figure 3). Both these approaches to menu interaction in virtual reality were observed as current mainstream paradigms for consumers (see figures 5 and 4) for examples of such interfaces in use.

A total of 12 participants were recruited from the university student body for this pilot study and all of them were regular computer

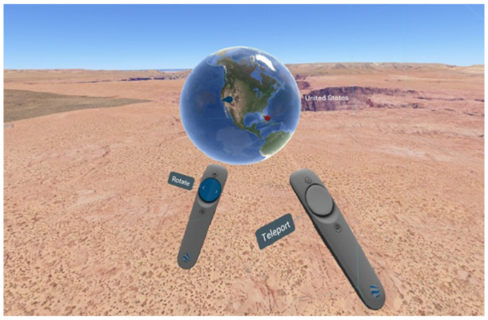

Figure 4: Google Earth VR

Figure 5: SteamVR Home

users. It was noted that about one third of the users had little to some experience using virtual reality but no one identified themselves as an advanced user. Each participant was instructed in the operation of the three lifelog exploration prototypes to reduce the potential of a learning bias. This involved a thorough walkthrough of each prototype and its features followed by the execution of a test query.

For our experiment, the dataset of 15 topics were divided into 3 groups of 5 and labelled A, B and C. This enabled us to organise the topics accordingly to ensure each group of events was searched for on each of the three prototypes equally. The order in which each user used the prototypes was randomised, as was the order of the 5 topics they were assigned to search for on each prototype. With 12 users, this resulted in each topic being searched for on each prototype a total of 4 times. Each participant was given a maximum of 180 seconds to locate each topic assigned to them. If the participant exceeded 180 seconds, the examiner would inform them and they would proceed to the next task and that topic was flagged as unsolved. Since the topics described semantic events which could span a large amount of time (e.g. riding a train) there was often a large pool of acceptable images to choose from for any given topic. For the purposes of our experiment it was only necessary for a user to locate a single image that was representative of the described topic. Once located, the examiner would inform the user if the image was a correct result and the experiment would proceed accordingly.

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11	T12	T13	T14	T15
U1	43	146	58	74	36	178	180	180	122	173	88	98	94	180	37
U2	9	21	38	24	17	51	144	129	76	64	82	180	15	35	18
U3	25	128	51	74	46	180	172	153	131	54	48	105	39	51	78
U4	152	92	47	80	29	81	180	29	64	51	65	180	67	58	29
U5	73	40	136	31	45	42	150	106	123	47	47	180	20	28	21
U6	117	60	18	28	46	140	180	44	37	40	52	180	137	180	19
U7	20	15	55	62	40	180	180	48	128	25	28	103	99	42	23
U8	14	32	35	31	28	26	73	105	44	19	15	112	11	14	9
U9	27	123	32	25	30	87	157	80	41	22	63	180	29	180	28
U10	47	180	109	117	24	180	180	105	39	147	57	50	179	92	67
U11	17	35	30	44	13	58	180	178	45	36	28	180	20	23	35
U12	42	38	26	31	40	94	103	177	81	49	58	180	66	180	28

PC   VR – Clipboard   VR – Billboard

Figure 6: Seconds taken for each of the 15 topics by our 12 participants (the cell colour and intensity indicate the prototype used and how many seconds the user took)

4 RESULTS

The results of this experiment can be seen in Figure 6 where each user is plotted vertically and the seconds they took for each topic are plotted horizontally. The colour of each result indicates the prototype used and the intensity of the colour correlates with how many seconds were taken. Where a user exceeded 180 seconds, the result is framed by a red border. In observing these results, we identified no notable correlation between the length of time it took to locate an event and the prototype that it was performed on. The primary factor appeared to be the difficulty of the topic itself. For example, some topics targeted very short time-spans in the lifelog which made them more difficult to query. This is particularly evident on topics 7 and 12 (see Figure 7) which were failed by many users, regardless of what prototype they used. It is important to acknowledge that this was only a small pilot study and it is likely more notable correlations could emerge in a larger experiment. Yet we believe this initial study suggests that visual lifelog exploration is not negatively impacted by the use of virtual reality as a platform. This conclusion leads us to continue our research into the development of known-item search and browsing tools for lifelog data in VR environments.

5 CONCLUSION

In this pilot study we set out to investigate whether virtual reality was a feasible platform for exploring visual lifelog datasets. Our small experiment consisting of 12 participants indicated to us that performing familiar lifelog interaction queries in virtual reality was a viable alternative and there appeared to be no noticeable drawback between using a virtual reality platform over a conventional platform. Subjective experience reported by some users also suggested exploring a lifelog in virtual reality felt more engaging and tactile than using a conventional platform. However, some users who had never used virtual reality acknowledged that there was a learning curve. The results of this pilot study do not indicate that virtual reality is a superior platform for visual lifelog exploration

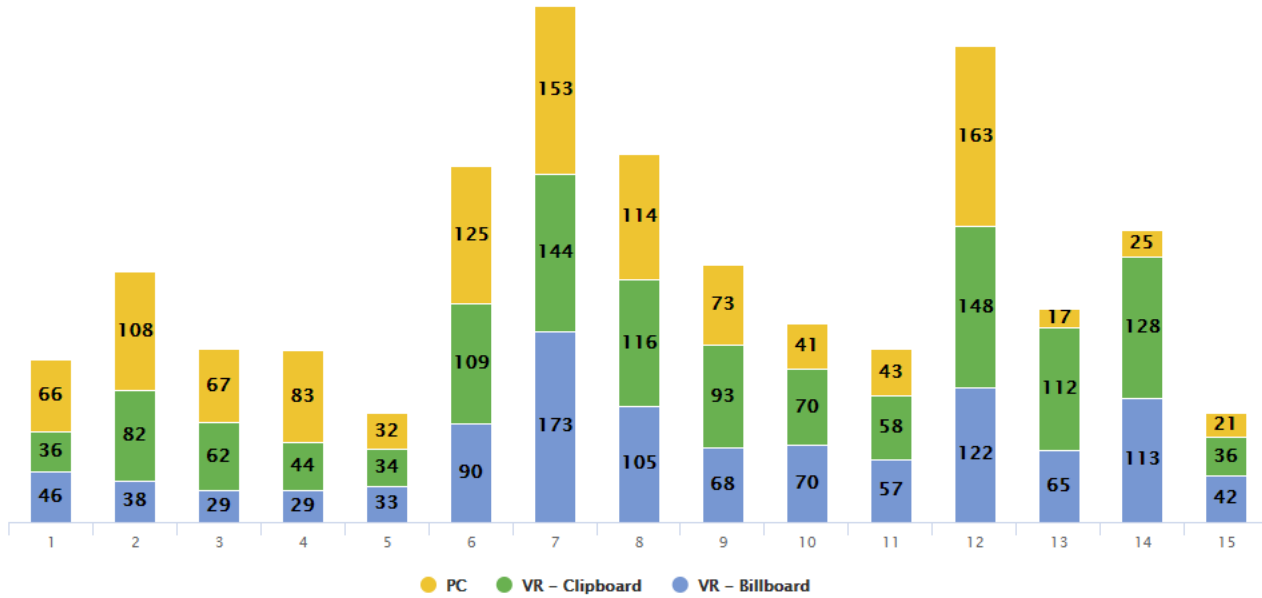

Figure 7: Average time in seconds taken for each topic on each prototype

but this initial study is part of a larger research effort to explore this possibility.

ACKNOWLEDGMENTS

This publication has emanated from research conducted with the financial support of Science Foundation Ireland (SFI) under grant number SFI/12/RC/2289.

REFERENCES

[1] Martin Dodge and Rob Kitchin. 2007. 'Outlines of a world coming into existence': Pervasive computing and the ethics of forgetting. *Environment and Planning B: Planning and Design* 34, 3 (2007), 431–445. https://doi.org/10.1068/b32041t

[2] Gurrin, Cathal; Joho, Hideo; Hopfgartner, Frank; Zhou, Liting; Albatal, Rami. 2016. NTCIR Lifelog: The First Test Collection for Lifelog Research. (2016), 705–708 pages. https://doi.org/10.1145/2911451.2914680

[3] ImageCLEF. [n. d.]. Conference Website, Last Accessed: April 2017. http://www.imageclef.org/. ([n. d.]).

[4] NTCIR. [n. d.]. Conference Website, Last Accessed: April 2017. http://research.nii.ac.jp/ntcir/index-en.html. ([n. d.]).

[5] NTCIR-12. [n. d.]. Conference Website, Last Accessed: April 2017. http://research.nii.ac.jp/ntcir/ntcir-12. ([n. d.]).

[6] SAS. [n. d.]. Blog, Last Accessed: August 2017. http://blogs.sas.com/content/sascom/2014/04/25/using-virtual-reality-understand-big-data/. ([n. d.]).

[7] 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