

LifeSeeker - Interactive Lifelog Search Engine at LSC 2019

Tu-Khiem Le*
Van-Tu Ninh*
Dublin City University
Ireland

Minh-Triet Tran
Ho Chi Minh University of Science
Vietnam

Duc-Tien Dang-Nguyen
University of Bergen
Norway

Liting Zhou, Pablo Redondo
Sinead Smyth, Cathal Gurrin
Dublin City University
Ireland

ABSTRACT

In this paper we present a first version of the LifeSeeker interactive lifelog retrieval engine that is under development at Dublin City University. This retrieval engine has been designed as a platform onto which future lifelog annotation and retrieval engines will be built. The first implementation of LifeSeeker has been designed for the LSC'19 comparative benchmarking challenge and it takes the form of a faceted search and browsing interface with the addition of query expansion to help solve the lexical-gap between novice users and the concept annotation tools employed for annotating the collection.

CCS CONCEPTS

- **Information systems** → **Search interfaces**; *Multimedia databases*;
- **Human-centered computing** → Interactive systems and tools.

KEYWORDS

lifelog, interactive retrieval, information system

ACM Reference Format:

Tu-Khiem Le, Van-Tu Ninh, Duc-Tien Dang-Nguyen, Minh-Triet Tran, Liting Zhou, Pablo Redondo, and Sinead Smyth, Cathal Gurrin. 2019. LifeSeeker - Interactive Lifelog Search Engine at LSC 2019. In *Lifelog Search Challenge'19 (LSC'19)*, June 10–13, 2019, Ottawa, ON, Canada. ACM, New York, NY, USA, 4 pages. <https://doi.org/10.1145/3326460.3329162>

1 INTRODUCTION

All information retrieval systems designed for use by individuals are inherently interactive. Users are a key component of the system operation in that they convert an information need from a thought process into a textual format and then implicitly judge the output of a retrieval system when browsing a ranked list [9]. Interactive retrieval poses a number of challenges that extend beyond the choice of retrieval model or annotation engine, because the user

*Both authors contributed equally to this research.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.

LSC'19, June 10–13, 2019, Ottawa, ON, Canada

© 2019 Association for Computing Machinery.

ACM ISBN 978-1-4503-6781-3/19/06...\$15.00

<https://doi.org/10.1145/3326460.3329162>

is a key component of the system. The Lifelog Search Challenge (LSC) is an international competitive benchmarking activity with the aim of supporting the fair and accurate comparison of different approaches to interactive retrieval from lifelog datasets [6].

While interactive information retrieval is a topic that has been active for decades, lifelogging and specifically lifelog retrieval is a recent topic that has been receiving increasing research attention since the seminal MyLifeBits[4] lifelog database in 2006. Consequently, in recent years we note that a number of comparative benchmarking exercises have been organised and the Lifelog Search Challenge (LSC) is the most recent of these.

In this paper, we describe the LifeSeeker prototype, motivate its construction and illustrate how it operates as well as the novel features that it incorporates. LifeSeeker extends prior research into interactive lifelog retrieval systems by the incorporation of novel features such as automatic query expansion, visual similarity retrieval and a wide-range of faceted querying options. These features were developed in order to assist a novice user in using the system, because it is our conjecture that a retrieval system for lifelog data would need to be designed with a novice-user in mind. In this paper, we will describe these novel features and illustrate how LifeSeeker operates.

2 RELATED RESEARCH

There have been a number of lifelog retrieval systems developed in the past, such as the MyLifeBits system [4], or the Sensecam browser from Microsoft [8]. These early systems, though pioneering in their application, were relatively straightforward data storage and browsing systems. While there have been many such lifelog browsing engines, one of the first identifiable multimodal lifelog search engines was developed by Doherty et al [2] as a means to support user experiments into human memory by allowing a user to generate and modify faceted queries for an interactive lifelog retrieval system. In their study, they showed that such an faceted interactive retrieval system significantly increased the success rate for a user to locate desired content, when compared to a conventional browsing interface. Following these, we can identify a number of interactive retrieval systems that were designed to support interactive retrieval from lifelogs.

The LeMoRe system[1] was a lifelog search engine designed to be used in an interactive manner and was deployed at the Lifelog Semantic Access Task (LSAT) of the the NTCIR-12 Lifelog challenge [5]. Being the only interactive lifelog retrieval system developed

for that challenge, it is not possible to compare its performance. More recently, the advent of the Lifelog Search Challenge, it became possible to directly compare interactive retrieval systems, operating over the same test collection and under the same experimental setting. For LSC'18, six lifelog systems were compared, with three of them performing well, which we highlight here. The lifeXplore [15], a lifelog browser supported interactive visual exploration and retrieval and metadata filtering, which presented the user with an adjustable multi-level feature map grouping together similar shots according to machine learning descriptors or hand-crafted features. A second system called SIRET [13] was based on an existing content-based video retrieval tool and viewed the lifelog as a image-sequence video and relied on a visual search and browsing metaphor. A third system was the virtual reality lifelog explorer [3] which ported a faceted search lifelog system to a VR-platform. These three systems performed highest in the 2018 running of the LSC [7] with the VR-platform narrowly achieving a higher score. One notable aspect of these systems was the heavy reliance on visual-based retrieval, which utilised computer vision technologies such as concept-detection and image similarity to aid the retrieval process. Indeed two of the systems [13, 15] were directly ported from existing video retrieval engines utilised at the VBS - Video Browser Showdown [12], a similar activity to the LSC, though aimed at interactive video search.

The contribution of the LifeSeeker system described in this paper, is that it incorporates a strong emphasis on faceted search using multiple sources of evidence from the LSC dataset, and it employs a prototype query expansion mechanism that addresses the lexical gap between user queries and the terms/concepts in the search engine index.

3 LIFELOG DATA FOR THE EXPERIMENT

The LSC'19 challenge used the same dataset as the LSC'18 challenge. The dataset is described in detail in [6], but we will briefly describe it here. The LSC dataset was a 27-day multimodal lifelog dataset gathered by one individual who wore multiple sensors and utilised smartphone and computer software to capture a continuous 24/7 lifelog. The lifelog data was redacted to remove faces and readable textual content. This data was then enhanced by the addition of various forms of metadata such as the output of a computer-vision concept detection toolset. This data was made available to download¹ for participants in the challenge. This dataset is then employed in the live search challenge with newly generated topics (expert and novice) being used.

4 OVERVIEW OF LIFESEEKER

The LifeSeeker interactive retrieval system has been designed primarily to support a user to input and refine queries, while also browsing through the retrieved ranked list in a fast and effective manner. The system is a progression of our previous NTCIR-14-Lifelog3 interactive search engine, but has been enhanced based on feedback from the NTCIR experimentation [16].

Since the LifeSeeker retrieval engine was aimed at novice users, we engaged in a small-scale qualitative user study with four novice users to both create a set of runs for our NTCIR14-Lifelog3 task

submissions as well as gain feedback into the usability of the system. Each user evaluated twelve of the twenty-four topics with the sort-order of topics being reversed for the 3rd and 4th user. For this initial user study, the following experimental protocol was employed; each user got 15 minutes of testing time with sample queries and once ready, each user processed their twelve ad-hoc queries on the NTCIR14-Lifelog3 dataset, with a time-limit of five minutes per query. After processing all twelve queries, feedback was sought via a standardised user experience questionnaire [10] and via task-observation. The findings suggested that we needed to enhance the system by:

- Taking measures to reduce the lexical gap.
- Integrating content-similarity to allow the user to find similar looking content.
- Narrowing search results with a filter panel.
- Replacing pagination with continuous loading to reduce user effort.

The LSC challenge scores teams based on the speed at which relevant content is located; hence the intuitiveness and speed of the interaction mechanism is extremely important. Since LifeSeeker is designed around the user, we first describe the user interface before we describe the underlying technical components.

4.1 User Interface Description

LifeSeeker supports two type of user queries. A user can choose to use a Google-style text box, which they can use to write free-text queries, or they can engage a faceted filtering mechanism which allows for detailed filtering of search results across many axes.

Initially LifeSeeker presents an intuitive user interface with the free-text box on the top of the screen and a result panel immediately below it, as shown in Figure 1 which shows the result for a simple query 'dog'. We refer to this as the free-text query mechanism.

After submitting a query, the user browses the ranked list of images (with associated metadata content) and can select any image for submission to the evaluation engine. Typically 100 images are displayed, but this is configurable. Alternatively it is also possible to explore within the temporal or visual context of any given image by selecting it and browsing within a small panel that appears as an interface layer above the selected image, as shown in Figure 2. This panel allows the user to browse forward and back temporally to observe the user actions before and after the selected image. Additionally, it is also possible to view similar images to the selected image using this panel.

Should this ad-hoc query approach to interactive search prove unsuccessful, then the user has the option to open the faceted search panel which contains a detailed faceted query selection interface, as shown in Figure 3. This panel allows for a number of facets to be added to the query as filters. The facets include 'days of the week', date-range, time-range, physical activity, location (category and name), biometrics (target heart rate and calorie burn), semantic location type, and visual concepts). Each of these can be selected independently or grouped together; our conjecture is that they become increasingly important as the size of the lifelog dataset increases and the size of the ranked list increases accordingly.

¹LSC Website: <http://lsc.dcu.ie>

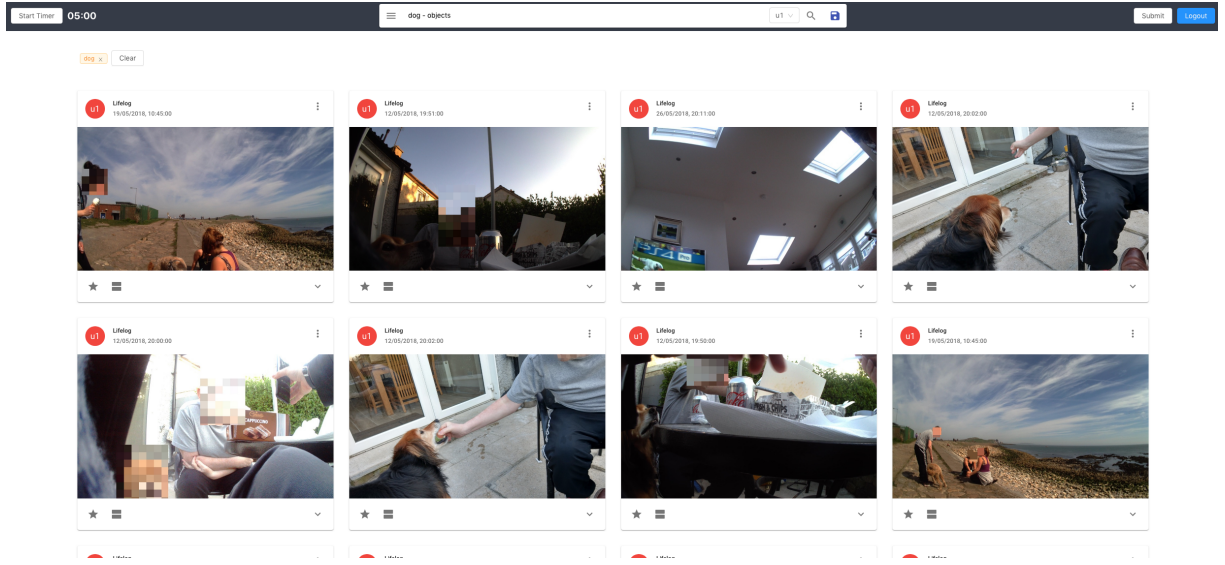


Figure 1: The result of a search for the query 'dog'

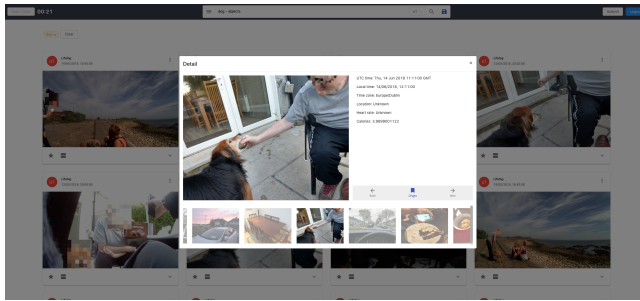


Figure 2: Illustrating the Contextual detail of a selected image

4.2 System Components

The LifeSeeker interactive retrieval system is designed as a standard retrieval system with a web-based user interface and employing MongoDB for back-end storage. For the sake of brevity, we don't describe the architecture here, rather we focus on highlighting the novel aspects.

4.2.1 Additional Concept Detection. As visual concepts play an important role in indexing and retrieving images, it would be advantageous for us to utilize them to increase the efficiency of our search engine. To increase the usefulness of visual concepts, we utilise SNIPER - an object detection network trained on MSCOCO dataset [11, 18, 19] to enhance the descriptions. Additionally, we predict the scene attributes and scene categories of images to enhance the landscape concepts by employing PlacesCNN [20]. In initial experimentation, these additional concepts are proven to work effectively on the NTCIR14-Lifelog-3 dataset.

4.2.2 Free-text Ranking Algorithm. The free-text ranking engine implemented in the system indexes all textual content associated

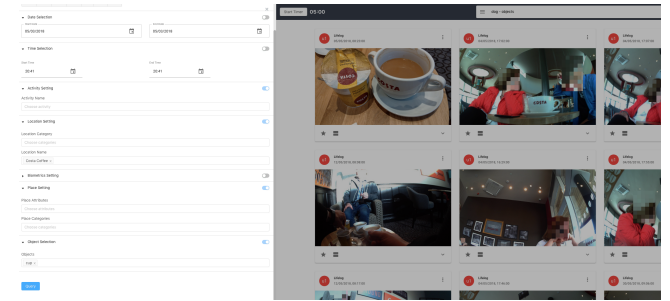


Figure 3: Illustrating the Faceted Search Panel (on left)

with any image within the collection. In order to reduce the architectural complexity and latency of the system, we choose to use a standard approach to term weighting [17] for the indexed data in MongoDB. This was appropriate given the small size of the collection. For the purposes of this interactive system, both stemming and stopwords were employed. The maximum number of results returned was 1,000, although in a standard configuration, only 100 were displayed to the user in the interface. The top 1,000 images was necessary for the ranking system to support faceted filtering.

4.2.3 Index Expansion. Initial experiences from teams at the first LSC'18 highlighted a lexical gap between user queries and the indexed concepts from the provided image annotation. This was further highlighted by qualitative feedback [16] from the initial testing of the LifeSeeker interface for NTCIR-14. Consequently, for this LSC search engine, we enrich the output of the visual and biometric concept detectors using a term-expansion (thesaurus-lookup) approach. For example the concept *seaside* would include the following synonyms; *shore, coast, sands, margin, strand, seaside, shingle, lakeside, water's edge, lido, foreshore, seashore, plage, littoral, sea*. It was our conjecture that this would make it easier for a novice

user, who is not familiar with the dataset annotation, to successfully query the system in a natural manner.

4.2.4 Image Similarity. One feature that was used successfully by participants in LSC'18 was visual similarity. In order to retrieve visually similar images, we utilise the Bag-of-Words model to transform visual features into a vector representation for comparing and returning similar images. Extracting visual features from image was done thanks to the Scale-Invariant-Feature-Transform (SIFT) [14] detector. Since parts of the images could contain the same content, we need to group all visual features of these parts into the same cluster so that it fully describes the content of the images. We apply K-Means Algorithm to classify all extracted features into K clusters, then perform vector quantization for all images in the dataset based on these clusters to transform them into data histograms. We can search for similar images by comparing the similarity in their histogram representation, which can be calculated using vector cosine distance.

5 CONCLUSION

In this paper, we presented an overview of a prototype interactive lifelog retrieval engine called LifeSeeker. LifeSeeker was designed to support efficient and effective retrieval of content from lifelog archives. LifeSeeker incorporated a number of novel features which were included based on experiences at LSC'18 and an initial user study [16]. We note that LifeSeeker is still an early-stage prototype and we hope to continue its development in the coming years to be in a position to compete in the Lifelog Search Challenge in 2020 and beyond.

For future enhancements, we will replace the simple index expansion methodology employed with a word-embeddings approach to query expansion [21] which we expect will capture higher order relationships between query words and indexed content. We will also enhance the quality of the visual features because this has been shown to be an important differentiate for system performance.

The free-text search system can be enhanced and we propose to re-use the BM25 term weighting algorithm [17], but with optimised parameters in future versions of LifeSeeker. Finally, we will continue to refine the user interaction supported by LifeSeeker in an effort to reduce the search time for a user to locate relevant content.

ACKNOWLEDGMENTS

This publication has emanated from research supported in part by research grants from Irish Research Council (IRC) under Grant Number GOIPG/2016/741 and Science Foundation Ireland under grant numbers SFI/12/RC/2289 and SFI/13/RC/2106. We wish to also thank the four users who tested our retrieval system for the NTCIR14-Lifelog3 task.

REFERENCES

- [1] Gabriel DeOliveira-Barra, X Giró-i Nieto, Alejandro Cartas-Ayala, and Petia Radeva. 2015. LEMoRe: A Lifelog Engine for Moments Retrieval at the NTCIR-Lifelog LSAT Task. *The 12th NTCIR Conference, Evaluation of Information Access Technologies* (2015), 0–5.
- [2] Aiden R. Doherty, Katalin Pauly-Takacs, Niamh Caprani, Cathal Gurrin, Chris J. a. Moulin, Noel E. O'Connor, and Alan F. Smeaton. 2012. Experiences of Aiding Autobiographical Memory Using the SenseCam. *Human-Computer Interaction* 27, 1-2 (2012), 151–174.
- [3] Aaron Duane, Cathal Gurrin, and Wolfgang Hürst. 2018. Virtual Reality Lifelog Explorer: Lifelog Search Challenge at ACM ICMR 2018. In *Proceedings of the 2018 ACM Workshop on The Lifelog Search Challenge (LSC '18)*. ACM, New York, NY, USA, 20–23. <https://doi.org/10.1145/3210539.3210544>
- [4] Roger Gemmell, Jim; Bell, Gordon; Lueder. 2006. My lifebits: a personal database for everything. *Commun. ACM* 49, 1 (2006), 88–95. <https://doi.org/10.1145/1107458.1107460>
- [5] Cathal Gurrin, Hideo Joho, Frank Hopfgartner, Liting Zhou, and Rami Albatal. 2016. Overview of NTCIR-12 Lifelog Task. (2016), 354–360. <http://eprints.gla.ac.uk/131460/> The authors acknowledge the financial support of Science Foundation Ireland (SFI) under grant number SFI/12/RC/2289 and the input of the DCU ethics committee and the risk & compliance officer. We acknowledge financial support by the European Science Foundation via its Research Network Programme ?Evaluating Information Access Systems?.
- [6] Cathal Gurrin, Klaus Schoeffmann, Hideo Joho, Bernd Munzer, Rami Albatal, Frank Hopfgartner, Liting Zhou, and Duc-Tien Dang-Nguyen. 2019. A Test Collection for Interactive Lifelog Retrieval. In *MultiMedia Modeling*. Ioannis Kompatsiaris, Benoit Huet, Vasileios Mezaris, Cathal Gurrin, Wen-Huang Cheng, and Stefanos Vrochidis (Eds.). Springer International Publishing, 312–324.
- [7] Cathal Gurrin, Klaus Schoeffmann, Hideo Joho, Liting Zhou, Aaron Duane, Andreas Leibetseder, Michael Riegler, Luca Piras, Minh-Triet Tran, Jakub Lokoč, and Wolfgang Hürst. 2019. Comparing Approaches to Interactive Lifelog Search at the Lifelog Search Challenge (LSC2018). *ITE Transactions on Media Technology and Applications* 7, 2 (2019), 46–59.
- [8] Steve Hodges, Lyndsay Williams, Emma Berry, James Srinivasan, Gavin Smyth, Narinder Kapur, and Ken Woodberry. 2006. SenseCam: A Retrospective Memory Aid. In *Proceedings of the 8th International Conference of Ubiquitous Computing (UbiComp 2006)* (proceedings of the 8th international conference of ubiquitous computing (ubicom 2006) ed.). Springer Verlag, 177–193. <https://www.microsoft.com/en-us/research/publication/sensecam-a-retrospective-memory-aid/>
- [9] Diane Kelly. 2009. Methods for Evaluating Interactive Information Retrieval Systems with Users. *Foundations and Trends in Information Retrieval* 3, 1&A\$2 (2009), 1–224. <https://doi.org/10.1561/15000000012>
- [10] Bettina Laugwitz, Theo Held, and Martin Schrepp. 2008. Construction and Evaluation of a User Experience Questionnaire. In *HCI and Usability for Education and Work*. Andreas Holzinger (Ed.). Springer Berlin Heidelberg, Berlin, Heidelberg, 63–76.
- [11] Tsung-Yi Lin, Michael Maire, Serge J. Belongie, Lubomir D. Bourdev, Ross B. Girshick, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Lawrence Zitnick. 2014. Microsoft COCO: Common Objects in Context. *CoRR* abs/1405.0312 (2014). arXiv:1405.0312 <http://arxiv.org/abs/1405.0312>
- [12] Jakub Lokoč, Gregor Kovalčik, Bernd Münzer, Klaus Schöffmann, Werner Bailer, Ralph Gasser, Stefanos Vrochidis, Phuong Anh Nguyen, Sitapa Rujikietgumjorn, and Kai Uwe Barthel. 2019. Interactive Search or Sequential Browsing? A Detailed Analysis of the Video Browser Showdown 2018. *ACM Trans. Multimedia Comput. Commun. Appl.* 15, 1, Article 29 (Feb. 2019), 18 pages. <https://doi.org/10.1145/3295663>
- [13] Jakub Lokoč, Tomáš Souček, and Gregor Kovalčik. 2018. Using an Interactive Video Retrieval Tool for LifeLog Data. In *Proceedings of the 2018 ACM Workshop on The Lifelog Search Challenge (LSC '18)*. ACM, New York, NY, USA, 15–19. <https://doi.org/10.1145/3210539.3210543>
- [14] David G. Lowe. 2004. Distinctive Image Features from Scale-Invariant Keypoints. *Int. J. Comput. Vision* 60, 2 (Nov. 2004), 91–110. <https://doi.org/10.1023/B:VISI.0000029664.99615.94>
- [15] Bernd Münzer, Andreas Leibetseder, Sabrina Kletz, Manfred Jürgen Primus, and Klaus Schoeffmann. 2018. lifeXplore at the Lifelog Search Challenge 2018. In *Proceedings of the 2018 ACM Workshop on The Lifelog Search Challenge (LSC '18)*. ACM, New York, NY, USA, 3–8. <https://doi.org/10.1145/3210539.3210541>
- [16] Van-Tu Ninh, Tu-Khiem Le, Liting Zhou, Graham Healy, Minh-Triet Tran, Duc-Tien Dang-Nguyen, Sinead Smyth, and Cathal Gurrin. 2019. A Baseline Interactive Retrieval Engine for the NTCIR-14 Lifelog-3 Semantic Access Task. In *The Fourteenth NTCIR conference (NTCIR-14)*.
- [17] S. E. Robertson and K. Sparck Jones. 1997. *Simple, Proven Approaches to Text Retrieval*. Technical Report.
- [18] Bharat Singh and Larry S Davis. 2018. An analysis of scale invariance in object detection-snip. *CVPR* (2018).
- [19] Bharat Singh, Mahyar Najibi, and Larry S Davis. 2018. SNIPER: Efficient Multi-Scale Training. *NIPS* (2018).
- [20] Bolei Zhou, Agata Lapedriza, Aditya Khosla, Aude Oliva, and Antonio Torralba. 2017. Places: A 10 million Image Database for Scene Recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence* (2017).
- [21] Guido Zucco, Bevan Koopman, Peter Bruza, and Leif Azzopardi. 2015. Integrating and Evaluating Neural Word Embeddings in Information Retrieval. In *Proceedings of the 20th Australasian Document Computing Symposium (ADCS '15)*. ACM, New York, NY, USA, Article 12, 8 pages. <https://doi.org/10.1145/2838931.2838936>