LIFER 2.0: Discovering Personal Lifelog Insights using an Interactive Lifelog Retrieval System

Van-Tu Ninh[1], Tu-Khiem Le[1], Liting Zhou[1], Luca Piras[2], Michael Riegler[3], Mathias Lux[4], Minh-Triet Tran[5], Cathal Gurrin[1], and Duc-Tien Dang-Nguyen[6]

[1] Dublin City University, Dublin, Ireland
[2] Pluribus One & University of Cagliari, Cagliari, Italy
[3] Simula Research Laboratory, Oslo, Norway
[4] Klagenfurt University, Klagenfurt, Austria
[5] University of Science, VNU-HCM, Ho Chi Minh City, Vietnam
[6] University of Bergen, Bergen, Norway

Abstract. This paper describes the participation of the Organiser Team in the ImageCLEFlifelog 2019 Solve My Life Puzzle (Puzzle) and Lifelog Moment Retrieval (LMRT) tasks. We proposed to use LIFER 2.0, an enhanced version of LIFER, which was an interactive retrieval system for personal lifelog data. We utilised LIFER 2.0 with some additional visual features, obtained by using traditional visual bag-of-words, to solve the Puzzle task, while with the LMRT, we applied LIFER 2.0 only with the provided information. The results on both tasks confirmed that by using faceted filter and context browsing, a user can gain insights from their personal lifelog by employing very simple interactions. These results also serve as baselines for other approaches in the ImageCLEFlifelog 2019 challenge to compare with.

1 Introduction

An increasingly wide range of personal devices, such as smartphones, video cameras, and wearable devices allow individuals to capture pictures, videos, and audio clips for every moment of their lives. Considering the huge amount of data created, questions on how to design and develop an automatic system for fast and accurate data retrieval and understanding are becoming increasingly important.

In this work, we highlight the state-of-the-art techniques adopted for ImageCLEFLifelog 2019 [3] at ImageCLEF2019 [7], which include *Solve My Life* (PUZZLE) and *Lifelog Moment Retrieval* (LMRT). For ImageCLEF LMRT task, considering the multi-modality of lifelog data, we pre-processed the images to remove noisy data as a first step and then focused on the exploitation of associated metadata (time, activities, location, etc.) from moments of daily life. Inheriting the structure of the interactive search engine from [15], we developed a new facets filter and context browsing interface, with additional visual concepts and criteria expansion for ImageCLEF2019 LMRT. For the Puzzle task, we interpreted this task as a clustering problem and applied the state-of-the-art Visual Bag-of-Words [2] method for both reordering lifelogger's moments and predicting the part-of-day.

Building on prior research, we extended the retrieval system and optimised it for the domain of lifelogging. The main contributions of this paper are thus:

- A short survey of the current and state-of-the-art work in relevant domain.
- An introduction and discussion of the schema and functions of our baseline interactive search engine.
- A presentation, analysis, and discussion of the results obtained from the official competition.

2 Related Work

Interactive Lifelog Retrieval System: In recent years, a large volume of work has been performed on developing information retrieval approaches to increasingly commonplace personal digital collections, such as lifelogs. This has been supported by a number of international benchmarking efforts, the most recent of which is the Lifelog Search Challenge(LSC) [5], which is a multi-annual, real-time retrieval challenge that evaluates different approaches to interactive retrieval from lifelog collections.

For benchmarking systems, Zhou el al. [15] provided an efficient retrieval system in 2018, based primarily on faceted querying using captured metadata, which served as a baseline for other systems, and provided the basis for the LIFER 2.0 system presented in this paper. For Interactive Retrieval, the LEMORE[11] system, integrates classical image descriptors with high-level semantic concepts and designs a graphical user interface that uses natural language to process a user's query. For a more complete review of interactive retrieval systems, we refer the reader to [5], which highlights six different interactive lifelog retrieval systems. More recently, we have noted the development of novel retrieval approaches, that transcend the desktop, such as the Virtual Reality interactive retrieval system [4] that combines visual concepts and dates/times as the basis for a faceted filtering mechanism that presents results in a novel VR-interface.

Image Retrieval: The description of the puzzle task is to rearrange the massive image data (without time stamps) in in chronological order and predict the correct day (Monday or Sunday) and part of day (morning, afternoon, or evening). One possible computer vision-based approach is to detect and extract features for efficient image retrieval. Visual Bag-of-Words is a well-known approach for this kind of solution. There are many visual features that can be used for visual Bag-of-Words such as SIFT [9], root-SIFT, SURF [1], etc. Another proper approach is to use deep feature from deep neural network like ResNet [6] to classify the part of day of an image and retrieve the most similar images to rearrange the images.

3 ImageCLEFlifelog2019 LMRT Task: LIFER 2.0- Baseline Interactive Retrieval Search Engine

For ImageCLEFlifelog2019 LMRT task, we exploit LIFER 2.0- baseline interactive search engine which was initiated in [15], and improved in [10]. In this section, we provide a description of the interactive retrieval system and how it can be used to solve information needs. Our system, as described in [10], is a criteria matching engine which is built mainly from five categories: date/time, location, activity, biometrics, and visual concepts.

1. **Date/Time**: Date/time is an important feature in our search engine system because it can narrow down the scope of moment searching. For instance, time is specifically useful in query 6: "Having breakfast at home" (must have breakfast at home from 5:00 am to 9:00 am". It could also be useful for result filtering and lifelogger's behaviour guessing. In our system, date/time criteria include week days, date, and time.
2. **Location**: Location criteria contain location categories and location names, which are also advantageous for user to retrieve the relevant images in topic 1, 5, and 6. These topics depend on mostly on location filtering to find the proper moments and increase the variety of chosen images.
3. **Activity**: Although activity metadata in ImageCLEFlifelog 2019 dataset is not diverse, it is a potential criterion to be integrated to our system to improve the search engine with user actions/behaviours when it is ready.
4. **Biometrics**: Due to the lack of activity information, biometric data provide us the means to guess the moments when lifelogger is eating, walking, moving by heart rate and calories changes.
5. **Visual Concepts**: These concepts play the key roles in finding the proper images for topics owing to the diversity of concepts, annotations, and keywords. They include place attributes, place categories, and objects' name. Place attributes and categories are extracted from places365-CNN [14] with top 10 extracted attributes and top 5 place category predictions. Objects in image are detected using Faster R-CNN [12] trained on MSCOCO dataset [8].

These five sources of information are instantiated in the user interface as facets of a user query, as shown in Figure 1.

The interface of our system was divided into two parts: facets filter and context browsing. For the facets filter, a user could adjust his/her choice of five aforementioned criteria to retrieve the desired moments. In each criterion, except for location, the keywords and tags are combined into query condition using the OR operator to expand the diversity of returned results. Finally, all the conditions from each criterion are merged into one final query by utilising the AND operator. For context browsing, the keywords and annotations from location, visual concepts, activity are added into an auto-complete search bar. The user then types and chooses the proper tags which are suitable for current context of each topic. The query processing of this function is the same as the facets filter. The interface of LIFER 2.0- baseline interactive search engine is demonstrated in Figure 2.

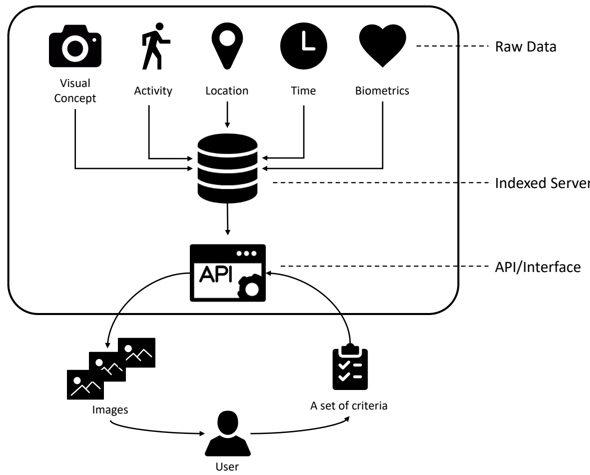

Fig. 1. Schema of LIFER 2.0, an improved interactive lifelog search engine.

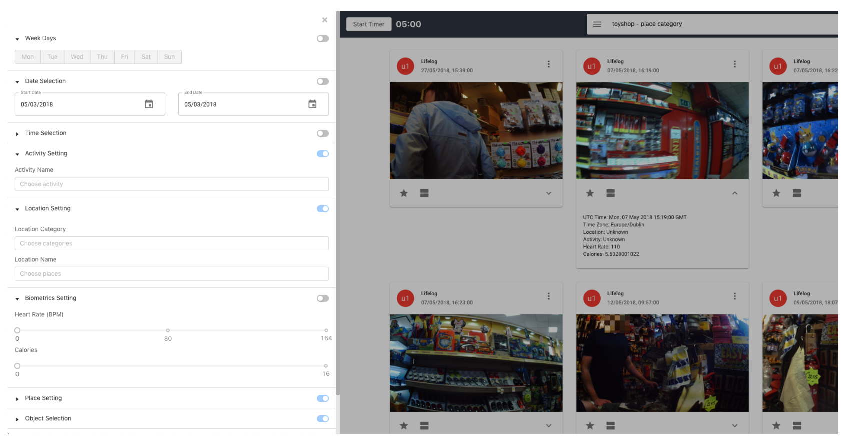

Fig. 2. The facets filter (left) and context browsing interface (right) of LIFER 2.0-baseline interactive search engine with an example of topic 1 results.

4 ImageCLEFlifelog2019 Puzzle Task: Lifelogger's Activity Mining Approach

In ImageCLEFlifelog2019 Puzzle Task, by utilising our baseline interactive search engine, we could review the provided training data and study lifelogger's activity. Because habit and daily routine of lifelogger's activity do not change much in lifelogger's life, we use only visual information to reconstruct the order of images in test set. We propose to utilise Visual Bag-of-Words [2] method to retrieve the proper time of images in each query and predict part-of-day based on the retrieved time. For this, we employ SIFT feature extraction [9] and conduct experiments on the number of visual clusters - k using the K-Means algorithm. The aim is to measure the effect of our proposed method while increasing the parameter k. The remaining steps are similar to the Bag-of-Words algorithm for text retrieval [13]. The way how we handle the rank list to choose the final time for each image in test set is presented in section 5.

5 Experiment and Results

5.1 LMRT Task

For the Lifelog Moment Retrieval Task, we conducted an interactive search experiment with the participation of two novice users. Each person was trained to use the search engine for 10 minutes and was then given a further 10 more minutes to get used to the system by performing 2 sample queries. Following this, the experiment began and the user executed 10 queries from test set. Table 1 displays the result of our two runs from the participant. As can be seen from the table, we achieved 41% in terms of precision, with cluster recall of 31% and 29% in F1 score.

Table 1. Submitted Runs for LMRT task.

RunID	P@10	CR@10	F1@10
LMRT Run 1	**0.41**	**0.31**	**0.29**
LMRT Run 2	0.33	0.26	0.24

Figure 3 and Figure 4 give us a precise look into multiple cut-off positions of the returned ranking for each query of both runs. We observe that the system has its stability across users as both graphs share the same pattern over three metrics.

5.2 Puzzle Task

In order to obtain the timestamp of each image in the test set, we established the majority vote among Top-N retrieved images from the returned ranking. The final time would be the average time of Top-N images. The accuracy of the Lifelogger's Activity Mining Approach also depends closely on the configuration of the Bag-of-Words model, especially the number of K clusters for visual features extracted from SIFT detector. Therefore, we submitted 8 runs in total, with 2 configurations of majority vote (Top-1 and Top-3) and 4 configurations of K clusters (512, 1024, 2048, 4096), which are summarised in Table 2

We achieved the overall score of 26.8% which shows that the best configuration is using the highest number of clusters and taking the time of the most relevant image as the final time.

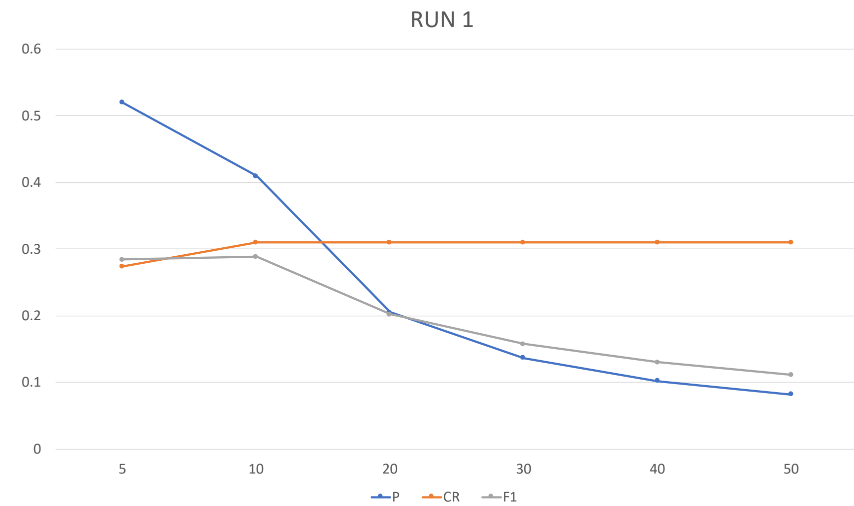

Fig. 3. Result of Run 1 in various cut-off positions

Fig. 4. Result of Run 2 in various cut-off positions

6 Discussions and Conclusions

In this paper, we introduced a baseline interactive search engine which uses faceted filtering and context browsing for the ImageCLEFlifelog2019 LMRT task. We also presented our proposed method for ImageCLEFlifelog2019 Puzzle Task to re-order the lifelogger's moments by using visual Bag-of-Words based on the belief of the minor change of his/her daily routine.

Table 2. Submitted Runs for Puzzle task.

RunID	Majority Vote	Number of Clusters	Kendall's Tau Score	Part of Day Accuracy	Primary Score
Puzzle Run 1	Top 1	512	0.055	0.308	0.182
Puzzle Run 2	Top 1	1024	0.034	0.352	0.193
Puzzle Run 3	Top 1	2048	0.033	0.336	0.184
Puzzle Run 4	Top 1	4096	0.048	**0.488**	**0.268**
Puzzle Run 5	Top 3	512	0.065	0.464	0.265
Puzzle Run 6	Top 3	1024	0.049	0.344	0.196
Puzzle Run 7	Top 3	2048	**0.071**	0.396	0.233
Puzzle Run 8	Top 3	4096	0.059	0.380	0.219

For the LMRT task, the analysis demonstrates that our search engine increased the F1 score by increasing cluster recall through valid experiment criteria. However, for novice users, the system still needs more annotation data of activities, object names, in order to increase the effectiveness of the search engine.

For the Puzzle task, it could be inferred that our proposed method could segment the images into correct clusters for part-of-days. However, our method could not solve the problem of re-ranking the moments in each cluster to increase the Kendall's Tau score. This shows that reconstruction the moments in each part of day still remains to be a challenge and requires further study.

7 Acknowledgement

This publication has emanated from research supported in party 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.

References

1. Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (surf). Comput. Vis. Image Underst. 110(3), 346–359 (Jun 2008), http://dx.doi.org/10.1016/j.cviu.2007.09.014
2. Csurka, G., Dance, C., Fan, L., Willamowski, J., Bray, C.: Visual categorization with bags of keypoints. In: Workshop on statistical learning in computer vision, ECCV. vol. 1, pp. 1–2. Prague (2004)
3. Dang-Nguyen, D.T., Piras, L., Riegler, M., Tran, M.T., Zhou, L., Lux, M., Le, T.K., Ninh, V.T., Gurrin, C.: Overview of ImageCLEFlifelog 2019: Solve my life puzzle and Lifelog Moment Retrieval. In: CLEF2019 Working Notes. CEUR Workshop Proceedings, CEUR-WS.org <http://ceur-ws.org>, Lugano, Switzerland (September 09-12 2019)
4. Duane, A., Gurrin, C.: User interaction for visual lifelog retrieval in a virtual environment. In: International Conference on Multimedia Modeling. pp. 239–250. Springer (2019)

5. Gurrin, C., Schoeffmann, K., Joho, H., Leibetseder, A., Zhou, L., Duane, A., Dang-Nguyen, D.T., Riegler, M., Piras, L., Tran, M.T., et al.: [Invited papers] Comparing Approaches to Interactive Lifelog Search at the Lifelog Search Challenge (LSC2018). ITE Transactions on Media Technology and Applications 7(2), 46–59 (2019)
6. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. CoRR abs/1512.03385 (2015), http://arxiv.org/abs/1512.03385
7. Ionescu, B., Müller, H., Péteri, R., Cid, Y.D., Liauchuk, V., Kovalev, V., Klimuk, D., Tarasau, A., Abacha, A.B., Hasan, S.A., Datla, V., Liu, J., Demner-Fushman, D., Dang-Nguyen, D.T., Piras, L., Riegler, M., Tran, M.T., Lux, M., Gurrin, C., Pelka, O., Friedrich, C.M., de Herrera, A.G.S., Garcia, N., Kavallieratou, E., del Blanco, C.R., Rodríguez, C.C., Vasillopoulos, N., Karampidis, K., Chamberlain, J., Clark, A., Campello, A.: ImageCLEF 2019: Multimedia retrieval in medicine, lifelogging, security and nature. In: Experimental IR Meets Multilinguality, Multimodality, and Interaction. Proceedings of the 10th International Conference of the CLEF Association (CLEF 2019), LNCS Lecture Notes in Computer Science, Springer, Lugano, Switzerland (September 9-12 2019)
8. Lin, T.Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., Dollár, P., Zitnick, C.L.: Microsoft coco: Common objects in context pp. 740–755 (2014)
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. Int. J. Comput. Vision 60(2), 91–110 (Nov 2004), https://doi.org/10.1023/B:VISI.0000029664.99615.94
10. Ninh, V.T., Le, T.K., Zhou, L., Healy, G., Tran, M.T., Dang-Nguyen, D.T., Smyth, S., Gurrin, C.: A Baseline Interactive Retrieval Engine for the NTICR-14 Lifelog-3 Semantic Access Task. In: The Fourteenth NTCIR conference (NTCIR-14) (2019)
11. de Oliveira Barra, G., Cartas Ayala, A., Bolaños, M., Dimiccoli, M., Giró Nieto, X., Radeva, P.: Lemore: A lifelog engine for moments retrieval at the ntcir-lifelog lsat task. In: Proceedings of the 12th NTCIR Conference on Evaluation of Information Access Technologies (2016)
12. Ren, S., He, K., Girshick, R., Sun, J.: Faster r-cnn: Towards real-time object detection with region proposal networks. In: Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 1. pp. 91–99. NIPS'15, MIT Press, Cambridge, MA, USA (2015), http://dl.acm.org/citation.cfm?id=2969239.2969250
13. Robertson, S.E., Jones, K.S.: Simple, proven approaches to text retrieval. Tech. rep. (1997)
14. Zhou, B., Lapedriza, A., Khosla, A., Oliva, A., Torralba, A.: Places: A 10 million image database for scene recognition. IEEE Transactions on Pattern Analysis and Machine Intelligence (2017)
15. Zhou, L., Hinbarji, Z., Dang-Nguyen, D.T., Gurrin, C.: Lifer: An interactive lifelog retrieval system. In: Proceedings of the 2018 ACM Workshop on The Lifelog Search Challenge. pp. 9–14. ACM (2018)