Electronic Letters on Computer Vision and Image Analysis 19(2):1-4, 2020

Behaviour understanding through the analysis of image sequences collected by wearable cameras

Estefanía Talavera Martínez[*,◇]

[*] *Johann Bernoulli Institute for Mathematics and Computer Science, University of Groningen, Netherlands*

[◇] *Departament de Matemàtiques i Informàtica, University of Barcelona, Spain*

This work was carried out at the University of Groningen (Netherlands) and at the University of Barcelona (Spain)

Abstract

Describing people's lifestyle has become a hot topic in the field of artificial intelligence. Lifelogging is described as the process of collecting personal activity data describing the daily behaviour of a person. Nowadays, the development of new technologies and the increasing use of wearable sensors allow to automatically record data from our daily living. In this paper, we describe our developed automatic tools for the analysis of collected visual data that describes the daily behaviour of a person. For this analysis, we rely on sequences of images collected by wearable cameras, which are called egocentric photo-streams. These images are a rich source of information about the behaviour of the camera wearer since they show an objective and first-person view of his or her lifestyle.

Key Words: Computer Vision, Image Classification, Egocentric Vision, Lifelogging, Temporal Segmentation, Behaviour Understanding, Lifestyle Tracking, Food-scenes classification, Sentiment Retrieval, Social patterns, Visual Pattern Recognition, Image Sequence Analysis.

1 Introduction

The study and analysis of people's lives and behaviour have become a trend in several disciplines. The automatic evaluation of someone's habits would help ordinary people and people with needs. In the latter case, practitioners can highly benefit from automatic tools for the objective quantification of behaviour.

The process of tracking and recording the personal daily activity data of a person is called *lifelogging*, and appeared in the 1960s. Nowadays, we can easily and automatically record data from our daily living with the development of new wearable technologies. Wearable devices are light-ware and affordable, which shows potential for the increase of their use by our society. Egocentric images are recorded by wearable cameras and show a first-person view of the life of the camera wearer. These collected images show an objective view of the daily life of a person and thus are a rich source of information about her or his habits. However, there is a lack of tools for the analysis of collections of egocentric photo-sequences and thus room for progress.

The goal of our research is to introduce appropriate tools for the analysis of egocentric photos-streams with the final goal of understanding the behaviour of the camera wearer. This analysis can help for memories

Correspondence to: <e.talavera.martinez@rug.nl>

Recommended for acceptance by <name>
ELCVIA ISSN:1577-5097
Published by Computer Vision Center / Universitat Autònoma de Barcelona, Barcelona, Spain

retrieval [3], self-monitoring, or health improvement. The recognition of events or moments in the collected photo-streams is needed for their later analysis and quantification. Therefore, we address the following topics in the field of egocentric vision. (1) The temporal segmentation for the analysis of cues of visual data and the detection of temporal boundaries between events. We describe an event as a sequence of images that describe an environment where the users spend time. (2) When studying the healthy lifestyle, the discovery of routine-related days plays a relevant role in the understanding of someone's routine. Moreover, (3) food-related activities and (4) social interactions are good descriptors of the nutritional and social daily habits of people, respectively, and thus, of high interest to our research. Finally, the (5) sentiment associated with an image is analysed in order to be able to find positive moments to be retrieved.

Deep learning has had a huge impact in the computer vision community for the classification and description of images. In particular, in our work, we have relied on the use of these techniques for the above-mentioned applications. Due to the limited amount of collected data, we used transfer learning theory in the different classification problems that we have addressed. Moreover, we made use of detected objects, places, and faces for the semantic description of the images. This obtained information allowed us to build on top of pre-trained models for the understanding of the lifestyle and behavioural patterns of the camera wearer.

In this paper we report the main achievements included in the doctoral thesis titled 'Lifestyle Understanding through the analysis of egocentric photo-streams', in which we proposed novel models for analysis of image sequences collected by wearable cameras [9].

2 Brief methods for behaviour analysis

We introduced novel automatic tools for the analysis of egocentric photo-streams with the ultimate goal of getting an understanding of the lifestyle of the camera wearer. Moreover, we release data sets* as a benchmark of egocentric image sequence analysis and classification and pattern recognition methods. Given sets of photo-streams collected by a camera wearer, our proposed models address five main topics in the field of egocentric vision. We visually depict the pipelines in Fig.1.

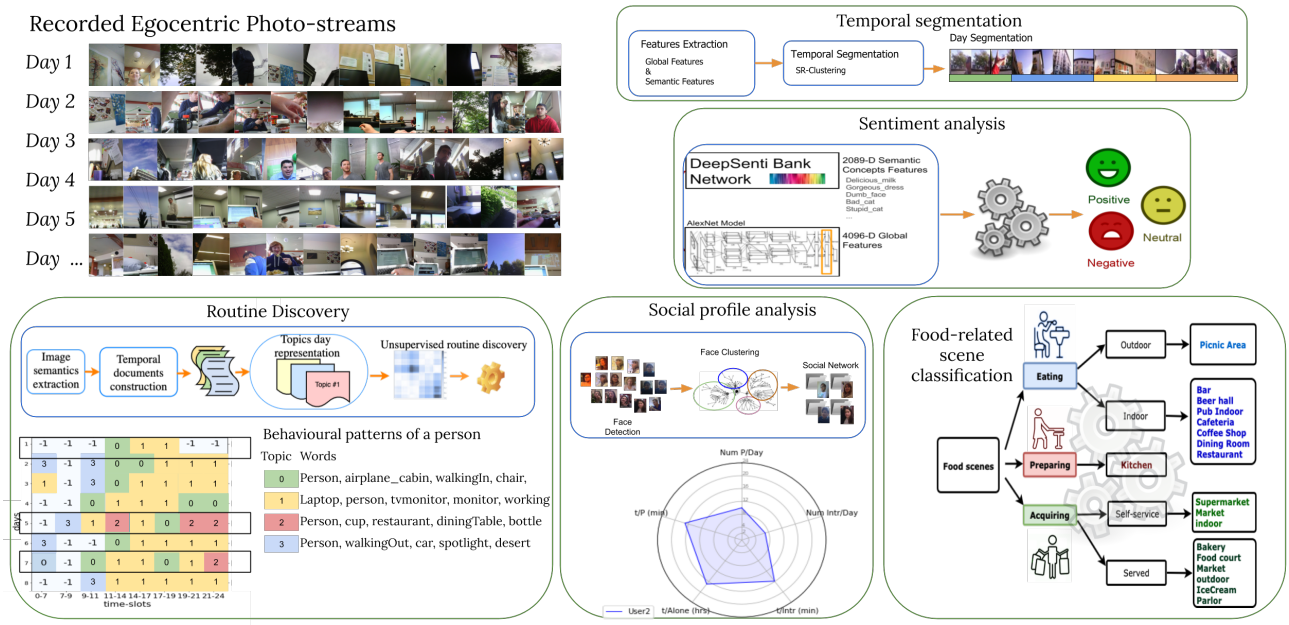
1. *Temporal photo-sequences segmentation:* We introduce the Semantic Regularized Clustering automatic model (SR-Clustering) [1] for the definition of temporal boundaries for the division of egocentric photo-sequences into moments, which are sequences of images describing the same environment. This model takes into account semantic concepts in the image together with the global image context for event representation. The model achieves a 66% F-score over our introduced *EDUB-Seg* dataset.

2. *Routine discovery:* On one side, we make publicly available our collected *EgoRoutine* dataset, composed of more than 100,000 images, from 104 days collected by 7 different users. On the other side, We propose automatic tools for the discovery of routine-related days and the visualization of patterns of behaviour. Firstly, we propose an unsupervised and automatic model for the discovery of routine days following a novelty detection approach [5]. Secondly, we introduce an unsupervised pipeline for the automatic discovery of routines from egocentric photo-streams in [8]. The model relies on topic modelling over semantic concepts extracted from the photo-sequences, and it is able to classify days into routine- and non-routine related with an accuracy of 80%.

3. *Food-related scenes recognition:* In [2], we address the task of food-related scene classification as a step forward for the automatic analysis of the nutritional behaviour of a person. We define a taxonomy that relates the food-related environments that describe the studied classes and organize them taking into account the main daily activities related to food consumption, acquisition, and preparation. Moreover, we introduce a hierarchical model composed of different layers of deep neural networks for classification.

*The data sets are available in http://www.ub.edu/cvub/dataset/

Figure 1: Given a collection of egocentric photo-streams describing the lifestyle of the camera wearer, we have developed automatic tools for the temporal segmentation into events [1], sentiment classification [6, 7], routine discovery [5, 8], social pattern analysis [4], and food-related scene recognition [2]



We adapt this model to the introduced taxonomy for the recognition of visually highly similar food-related images into 15 different classes. Finally, we introduce the *EgoFoodScenes* dataset composed of 33000 images and 15 food-related environments. Our model achieves a classification accuracy of 68%.

4. *Sentiment retrieval:* We explore the classification of egocentric images based on their associated sentiment, Positive, Neutral, and Negative. To this end, we propose two different analytic tools. On one side, in [7] we study how semantic concept with sentiment label can help to describe the sentiment associated with an image and how they can be combined to describe the sentiment related to an event. On the other side, in [6] we also evaluate their combination with global descriptors extracted with pre-trained convolutional neural networks, proving that the combination of descriptors enhances the performance of the classifier. We test our method on a new egocentric dataset of 12,088 pictures with ternary sentiment values acquired from 3 users in a total of 20 days. From the results achieved by these two works, we conclude that positive images tend to relate to outdoor environments or social interactions, while neutral images describe. work-related environments. In contrast, negative images are those that are non-informative or visually not clear.

5. *Social pattern characterization:* In [4], we propose a model that characterizes the social behaviour of the camera wearer based on the occurrence of people that the camera wearer meets throughout her/his photo-streams collection. To do so, we perform person re-identification throughout the collected photo-sequences by a user. In order to address social behaviour understanding, we propose social parameters that allow us to draw a radar chart for the comparison among individuals. The detected social interactions shape the social habits of the users, describing when, with whom, and for how long they interact.

3 Conclusions

The introduced and made publicly available egocentric data sets and the obtained results in the different performed experiments indicate that human behaviour can be identified and studied. We conclude that the developed automatic algorithms for the analysis of egocentric images allow a better understanding of the lifestyle of the camera wearer. Further developments in this field and the later analysis of discovered behavioural trends can lead to new applications that will result in an improvement of the quality of life of people and thus, are worth to be exploring. Following this line, we foresee that the characterization of the behavioural patterns of a person can be addressed by exploring multi-modal analysis and pattern mining by aggregating data coming from different types of sensors.

References

[1] Mariella Dimiccoli, Marc Bolaños, Estefania Talavera, Maedeh Aghaei, Stavri G Nikolov, and Petia Radeva. Sr-clustering: Semantic regularized clustering for egocentric photo streams segmentation. *Computer Vision and Image Understanding*, 155:55–69, 2017.

[2] Estefania Talavera Martinez, Maria Leyva-Vallina, Mostafa Kamal Sarker, Domenec Puig, Nicolai Petkov, and Petia Radeva. Hierarchical approach to classify food scenes in egocentric photo-streams. *IEEE journal of biomedical and health informatics*, 2019.

[3] Gabriel Oliveira-Barra, Marc Bolaños, Estefania Talavera, Olga Gelonch, Maite Garolera, and Petia Radeva. Lifelog retrieval for memory stimulation of people with memory impairment. In *Multimodal Behavior Analysis in the Wild*, pages 135–158. Elsevier, 2019.

[4] Estefania Talavera, Alexandre Cola, Nicolai Petkov, and Petia Radeva. Towards egocentric person re-identification and social pattern analysis. In *Applications of Intelligent Systems: Proceedings of the 1st International APPIS Conference 2018*, volume 310, page 203. IOS Press, 2018.

[5] Estefania Talavera, Nicolai Petkov, and Petia Radeva. Unsupervised routine discovery in egocentric photo-streams. In *International Conference on Computer Analysis of Images and Patterns*, pages 576–588. Springer, 2019.

[6] Estefania Talavera, Petia Radeva, and Nicolai Petkov. Towards egocentric sentiment analysis. In *International Conference on Computer Aided Systems Theory*, pages 297–305. Springer, 2017.

[7] Estefania Talavera, Nicola Strisciuglio, Nicolai Petkov, and Petia Radeva. Sentiment recognition in egocentric photostreams. In *Iberian Conference on Pattern Recognition and Image Analysis*, pages 471–479. Springer, 2017.

[8] Estefania Talavera, Carolin Wuerich, Nicolai Petkov, and Petia Radeva. Topic modelling for routine discovery from egocentric photo-streams. *Pattern Recognition*, page 107330, 2020.

[9] Estefanía Talavera Martínez. *Lifestyle understanding through the analysis of egocentric photo-streams*. PhD thesis, University of Groningen, 2020.