Automatically Detecting Important Moments from Everyday Life using a Mobile Device

Zhengwei Qiu
CLARITY;Centre for Sensor Web Technologies,
School of Computing
Dublin City University, Ireland
zqui@computing.dcu.ie

Cathal Gurrin
CLARITY;Centre for Sensor Web Technologies,
School of Computing
Dublin City University, Ireland
cgurrin@computing.dcu.ie

Aiden Doherty
CLARITY;Centre for Sensor Web Technologies,
School of Computing
Dublin City University, Ireland
adoherty@computing.dcu.ie

Alan F. Smeaton
CLARITY;Centre for Sensor Web Technologies,
School of Computing
Dublin City University, Ireland
Alan.smeaton@dcu.ie

Abstract

This paper proposes a new method to detect important moments in our lives. Our work is motivated by the increase in the quantity of multimedia data, such as videos and photos, which are capturing life experiences into personal archives. Even though such media-rich data suggests visual processing to identify important moments, the oft-mentioned problem of the semantic gap means that users cannot automatically identify or retrieve important moments using visual processing techniques alone. Our approach utilises on-board sensors from mobile devices to automatically identify important moments, as they are happening.

Keywords

important moments, novelty, lifelog

Introduction

One good quality image may communicate experiences, concepts or situations in far greater detail than a textual description. With such consideration, we have witnessed an increase in the research effort for the gathering of visual lifelogs (vast digital diaries) of a person's activities in recent years. One example of this is the SenseCam, from Microsoft, which can capture about 3 photos every minute automatically and in so doing build up a rich visual diary of up to 5,000 photos a day. In additon to the SenseCam, an ever increasing number of new devices and applications can capture extended periods of our lives in video and image format.

When faced with the challenge of organising and searching through such vast personal media archives, an efficient and effective way of judging the importance of captured moments within these archives is one essential challenge to be overcome. We already know that an entire day, captured visually, generates at least 30 such moments [1], each of which should be indexed for later retrieval and judged important or not. The challenge in automatically identifying important moments in real time is in accurately judging moment importance without relying on user input or annotation. Requiring a person to actually spend time annotating the content, brings a number of drawbacks; manual annotations vary from person-to-person and from day-to-day and there is no efficient way to annotate such a volume of daily multimedia annotations.

Algorithms for visual understanding could hold a key to idetnifying important moments from visual media streams by utilising computer vision software to attempt to semantically 'understand' the data. Thus far, however, this has proven to be both very time expensive and not accurate enough. Both annotation and automatic visual processing can help us to locate moments from vast digital archives, but this does not help in the disambiguation of what is an important event from the normal daily events or moments. In this paper, we propose an approach for automatically detecting important moments using contextual data.

Event Segmentation into Moments

Recent research on multimedia lifelog retrieval has addressed the challenge of identifying life events or moments automatically [1][2] to find the important information from videos and photos. But when the archive grows by 30 or more significant moments per day, the number of moments ranked highly for any particular query becomes a concern.

Consider for a moment the conventional written diary. The most important moments from a day are manually identified and recorded in this manual lifelog. This is not a challenge for us because our senses can identify when a moment is important by scanning environmental context and self-driven goals. Such allocation of importance to moments then regulates how our memories encode the moment, either consigning it to memory or discarding it. It is our conjecture that by mining data streams from wearable sensors on a mobile smartphone, it becomes possible to automatically identify when moments that are likely to be judged important are occurring. We now provide a description of how this is possible.

Detect Important Moments using Real and Virtual Sensors

To detect important moments, we need to understand what makes a moment important. Obviously, users will not think their routine life important. Events such as driving to work or having lunch with regular colleagues will not usually be considered as important. However when an extraordinary event occurs, such as meeting a friend, having a dinner date or going to a new place, such moments should be considered important, because of the novelty of the moment when compared to the routine live events.

As mentioned in [3], a moment is the smallest unit used to describe change in life events. As the fundamental element of life, changes in time, location, people and activities are the important symbols of novelty. As suggested above, before we can detect the novelty in users' lives, we should know what their routine live events and event sequencing are. By collecting data streams from mobile device sensors, the system can automatically define the different life patterns for different people. Recent generations of smart phones are equiped with many types of sensors that can help to define moment importance. Take for example, the Google Nexus One smartphone (which we employ in our research); this phone incorporates a number of sensors, including a 3-axis accelerometer, GPS receiver, bluetooth adapter, WiFi adapter and compass. Figure 1 (on the next page) shows the different kinds of sensor information we can gather by the mobile phone. On the left side are the raw sensor readings and on the right side are semantic sensor readings from virtual (software) sensors. The most influential sensed data that we gather are:

Time: Every sensor reading and virtual sensor reading has timestamp (date/time). Some sensor readings (accelerometer) occur every second while others (photos, bluetooth) occur a number of times per minute. We consider time to be the key binding agent for sensor readings.

Location information: We can acquire users location information by GPS. For any situation when there is no GPS available (e.g.. indoor), we will combine bluetooth, WiFi and cell tower ids to approximately locate the user [4][5].

Activities: With a tri-axial accelerometer & compass, we can infer users' activities in the real world (e.g. walking, sitting, standing) by classifying users activities using machine learning techniques [6][7]. Other more subtle activities can be infered by screen status and the compass (e.g. playing games or making phone calls).

Social relationships: Users' relationships can be mined automatically from Bluetooth device encounters and phone communication logs [8][9]. As various types of time, location, activity and social relationship data are continuously gathered, it is important to identify inherent patterns that span across large temporal archives of location, activity and social relationship data. From these patterns we propose to recognise important moments by firstly seeking novel moments and later by automatically learning the signature of an important moment from the sensor streams. For example, through pattern analysis of time, location, activity and social relationship data, we may find that a user has dinner in his friend's home every week, but on one particular week a new person appears at dinner, and this is an indication of a moment of meeting a potential new friend.

Due to the diversity in users' interests and daily activities, we apply pattern analysis to every user and every data source from each user. This approach helps us to quickly clean out noisy and infrequent items in each category and focus on identifying and combining patterns that are truly frequent and span multiple data sources. In addition, the pattern analysis process is incremental as new data are added into the system, they are analysed and integrated into previously-discovered patterns or are used to generate new patterns. By compared with life pattern, all unusual moments can be detected by the changing of time, location, activity and relationship, along with the basic sensors.

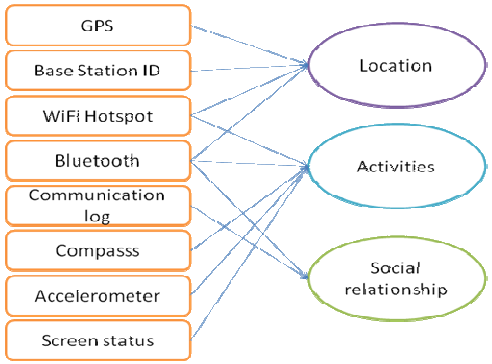

Figure 1 Concepts (right) mined from the onboard sensors (left)

Conclusions and Future work.

We present a model to be explored towards detecting users' important moment from lifelog data. Several challenges need to be addressed in building such a system and our planned evaluation will require ecxtensive user experimentation in the real world.

Acknowledgements

This material is based upon works supported by the Science Foundation Ireland under Grant No. 07/CE/I1147.

References

[1] A. R. Doherty and A. F. Smeaton, "Automatically segmenting lifelog data into events," in *Ninth International Workshop on Image Analysis for Multimedia Interactive Services*, pp. 20–23, 2008.

[2] B. Clarkson, K. Mase, and A. Pentland, "Recognizing user context via wearable sensors," in *Wearable Computers, 2000. The Fourth International Symposium on*, pp. 69-75, 2000.

[3] J. F. Allen and P. J. Hayes, "Moments and points in an interval-based temporal logic," *Computational Intelligence*, vol. 5, no. 3, pp. 225–238, 1989.

[4] M. S. Bargh and R. de Groote, "Indoor localization based on response rate of bluetooth inquiries," in *Proceedings of the first ACM international workshop on Mobile entity localization and tracking in GPS-less environments*, pp. 49–54, 2008.

[5] D. Madigan, E. Einahrawy, R. P. Martin, W. H. Ju, P. Krishnan, and A. S. Krishnakumar, "Bayesian indoor positioning systems," in *INFOCOM 2005. 24th Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings IEEE*, vol. 2, pp. 1217–1227, 2005.

[6] C. Bouten, K. Koekkoek, M. Verduin, R. Kodde, and J. Janssen, "A triaxial accelerometer and portable data processing unit for the assessment of daily physical activity," *Biomedical Engineering, IEEE Transactions on*, vol. 44, no. 3, pp. 136-147, 1997.

[7] N. Ravi, N. Dandekar, P. Mysore, and M. L. Littman, "Activity recognition from accelerometer data," in *Proceedings of the National Conference on Artificial Intelligence*, vol. 20, p. 1541, 2005.

[8] D. Quercia, J. Ellis, and L. Capra, "Nurturing Social Networks Using Mobile Phones," *IEEE Pervasive Computing*, 2010.

[9] J. K. Min, S. H. Jang, and S. B. Cho, "Mining and Visualizing Mobile Social Network Based on Bayesian Probabilistic Model," *Ubiquitous Intelligence and Computing*, pp. 111–120, 2009.