Application of Statistical Physics for the Identification of Important Events in Visual Lifelogs

Na Li, Martin Crane, Heather J. Ruskin and Cathal Gurrin
School of Computing, Dublin City University, Ireland
Email: na.li@dcu.ie, {mcrane, hruskin, cgurrin}@computing.dcu.ie

***Abstract*—Dementia is one of the most common diseases in the elderly people. Experience shows that Microsoft's SenseCam can be an effective memory-aid device, as it helps users to improve recollecting an experience by creating visual lifelogs. Given the vast amount of images that are maintained in a visual lifelog, it is a significant challenge to deconstruct a sizeable collection of images into meaningful events for users. In this paper, random matrix theory (RMT) is applied to a cross-correlation matrix *C*, constructed using SenseCam lifelog data streams to identify such events. The analysis reveals a number of eigenvalues that deviate from the spectrum suggested by RMT. The components of the deviating eigenvectors are found to correspond to "distinct significant events" in the visual lifelogs. Finally, the cross-correlation matrix *C* is cleaned by separating the noisy part from the non-noisy part. Overall, the RMT technique is shown to be useful to detect major events in SenseCam images.**

I. Introduction

Dementia is an important public health issue as it is one of the most common diseases in the elderly people. There are an estimates suggest that Ireland have 41,470 people with dementia in 2010 [1] and approximately 5.1 million people live with dementia in the European Union [2]. Microsoft Research has contributed a device, the SenseCam, that has potential to aid dementia patience to cope with their life situation. SenseCam [3] was first developed to help people with memory loss, but the camera is currently being tested to aid those suffering from serious cognitive memory loss [4]. The SenseCam is a small, wearable camera that takes images, automatically, in order to document the events of a wearers's day. The SenseCam also contains several electronic sensors, including those which record light-intensity and light-colour, a passive infrared (body heat) detector, a temperature sensor, and a multiple-axis accelerometer for monitoring the wearer's environment. The device takes pictures at VGA resolution, (480x640 pixels), and stores these as compressed .jpg files on internal flash memory. SenseCam can thus collect a large amount of data, even over a short period of time, with a picture typically taken every 30 seconds, an average of 2,000 images captured in any given day, together with associated sensor readings. The SenseCam produces images which are very similar to one's memory, particularly episodic memory, which is usually in the form of visual imagery [3]. By reviewing the day's filmstrip, patients have found it much easier to retrieve lost memories [5]. However, given the large size of the dataset that is created by the SenseCam, refreshing one's memory just by browsing the vast corpus is a tedious, if not unacceptable task. Hence, techniques are required for all users to manage, organise and analyse these large image collections, e.g., by automatically highlighting *key episodes* and, ideally, classifying them in order of importance to the life logger. Doherty et al. [6] address this challenge by identifying distinct events within a full day, *e.g. breakfast, working on PC, meeting, etc.*. However, their approach still contains a significant percentage of routine events. Our previous research [7] tackle the challenge by treating SenseCam images as time series. They show that these time series exhibit a strong long-range correlation, concluding that the time series is not a random walk, but is cyclical, with continuous low levels of background information picked up constantly by the device. Further, they adopt a cross-correlation matrix to highlight *key episodes*, thus identifying boundaries between different daily events.

However, due to the finite length of time series available to estimate cross correlations, the matrix contains much which corresponds to "random" contributions [8, 9]. As a consequence, their technique results in the identification of a high percentage of noise or routine events. This phenomenon can also be observed in other domains such as the analysis of financial data, wireless communications and many other fields. A well-proven technique to handle this issue is the application of random matrix theory (RMT) [10]. In this paper, we investigate whether RMT can be used to distinguish routine events from important events. We argue that such routine events can then be removed from the cross-correlation matrix by applying RMT. Our goal is to segment the content of the cross-correlation matrix into two: (a) the part of the correlation matrix that conforms to the properties of random correlation matrices ("noise") and (b) the part of the correlation matrix that deviates from random (i.e. has "information" on important events).

This paper is organised as follow: in Section II we describe the data used, methods are reviewed in Section III, Section IV details the results obtained, conclusions and outlook on future work are given in Section V.

II. Data

For this study, we analyze 2096 lifelog images that have been recorded using a SenseCam over the period of one day. The wearer of the camera, i.e., the lifelogger, experienced an average day of her life: commuting to the office in the morning, sitting and working in the office at a desk, talking with colleagues and sharing lunch in the cafeteria, as well as commuting back home in the evening and so on. In order to create a ground truth, the user reviewed her collection and manually classified the whole day into 12 events. Data statistics are reported in Table I. Given the size of the test corpus and its content, we argue that it is a typical visual lifelogging collection depicting a typical day of the lifelogger's life. As discussed above [6], a user will experience approximately 20 events per a day, but when exploring one's lifelog, reviewing routine or "boring" events has only limited interest, depending on the device purpose [5]. Efforts to determine automatically which events is most important or unusual (e.g., talking with a colleague as opposed to working in front of a computer), is an open research challenge. In order to distinguish routine or "boring" events from important events, we apply RMT methods to the cross-correlation matrix of the dataset, where such noise filtering has proved successful in many fields [10–14]. In the next section, successful the method is outlined.

TABLE I. DATA STATISTICS

Event Number	Event Series	Number of Images
1	Go to work	38
2	Arriving in the office	21
3	Working	136
4	Chatting with people	107
5	Working	157
6	Walking in the building	29
7	Working	177
8	Go to the bank	108
9	Working	412
10	Lunch	148
11	Working	668
12	Leaving the office	38
		Total: 2096

III. METHODS

A. Random Matrix Theory

In order to optimize the calculation process and reduce the amount of memory required for our calculations, we first adopt an averaging method to decrease the image size from 480×640 pixels to 60×80 pixels. Given pixels $G_i(t)$, $i = \{1, .., N\}$, of a collection of images. The equal-time cross-correlation matrix [15] may be expressed as follows:

$$C_{ij} \equiv \frac{\langle G_i G_j \rangle - \langle G_i \rangle \langle G_j \rangle}{\sigma_{(i)} \sigma_{(j)}} \tag{1}$$

where $\sigma_{(i)}$ is the standard deviation of G_i for image numbers $i = \{1, .., N\}$, and $< \ldots >$denotes a time average over the period studied.

In matrix notation, the correlation matrix can be expressed as $C = \frac{1}{T} G G^{\tau}$, where τ is the transpose of a matrix, G is an $N \times T$ matrix with elements g_{it}. N is the number of images and T is the pixel size of an image. The spectral properties of C may be compared to those of a "random" Wishart correlation matrix $R = \frac{1}{T} A A^{\tau}$ [10], where A is an $N \times T$ matrix with each element random, by distributed with zero mean and unit variance.

In particular, the limiting property for the sample size $N \to \infty$ and sample length $T \to \infty$, providing that $Q = T/N \geq 1$ is fixed, has been analysed to give the distribution of eigenvalues λ of the random correlation matrix R, given by:

$$P_{rm}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \tag{2}$$

where λ_- and λ_+ are the minimum and maximum eigenvalues of R. Hence, by comparing the empirical distribution of the eigenvalues of the correlation matrix to the distribution for a random matrix, as given in Eq. (2), we can identify those key eigenvalues which can be used to identify the specific information relating to the system. Eigenvector analysis enables identification of the specific information present, in terms of contributory components.

B. Eigenvector Analysis

Differences between the eigenvalues $P(\lambda)$ of C and RMT eigenvalues, $P_{rm}(\lambda)$ should also be displayed, therefore, in the statistics of the corresponding eigenvector components. In order to interpret this deviation of the eigenvectors, we note that the largest eigenvalue is an order of magnitude larger than the others, which constrains the remaining $N-1$ eigenvalues, since the trace of C, $Tr[C]$ sums to N. Hence, in order to analyse the contents of the remaining eigenvectors, we need first remove the effect of the largest eigenvalue. To do this we can use the linear regression [10]

$$G_i(t) = \alpha_i + \beta_i G^{large}(t) + \epsilon_i(t) \tag{3}$$

where $G^{large} = \sum_1^N u_i^{large} G_i(t)$ and N is the number of images in our sample. Here u_i^{large} corresponds to the components of the largest eigenvector. The cross-correlation matrix C is then recalculated using the residuals $\epsilon_i(t)$. If we quantify the remainder variance, (i.e., of the part not explained by the largest eigenvalue) as $\sigma^2 = 1 - \lambda_{large}/n$, this value can be used to recalculate our values of $\lambda_{\pm}$.

IV. RESULTS

A. Eigenvalue Analysis

As stated above, our aim is to separate information (major events) and noisy parts from the cross-correlation matrix C. In order to approach this task, we compare the eigenvalue distribution $P(\lambda)$ of C with $P_{rm}(\lambda)$. Fig. 1 compares the probability distribution $P(\lambda)$ with $P_{rm}(\lambda)$. We note the presence of a well-defined "bulk" of eigenvalues which fall within the bounds $[\lambda_-, \lambda_+]$ for $P_{rm}(\lambda)$. We also note deviations for a number of ($\approx 80\%$) largest and smallest eigenvalues. Fig. 1 hence suggests that the cross-correlation matrix captured most of the major events from the data stream, but still contains a small part of noise ($\approx 20\%$).

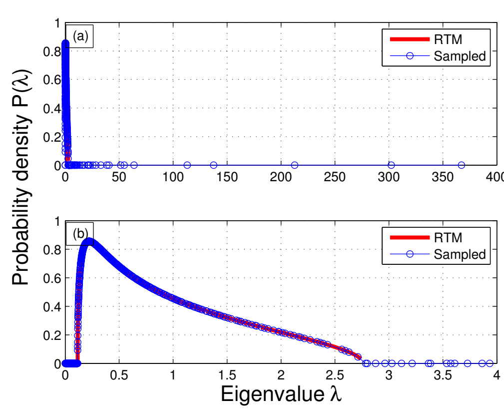

Fig. 1. Eigenvalue Distribution for the Correlation Matrix C for SenseCam data, Full spectral distribution (a) Partial spectral distribution (b)

B. Bootstrapping

In order to study whether there is no dependency on the choice of image series or the length of the pixel series we split the image series into two segments and compare the eigenvalue spectrum of the cross-correlation matrix C with that of a random Wishart matrix. As can be seen in Fig. 2, the eigenvalue contributions are very similar for both periods chosen which implies independence from the choice of image series, i.e., for both image series studies we observe the same amount of eigenvalues that deviate from the RMT prediction. We conclude from this, that independent from the image series that is analyzed, the proposed technique can be used to identify major events.

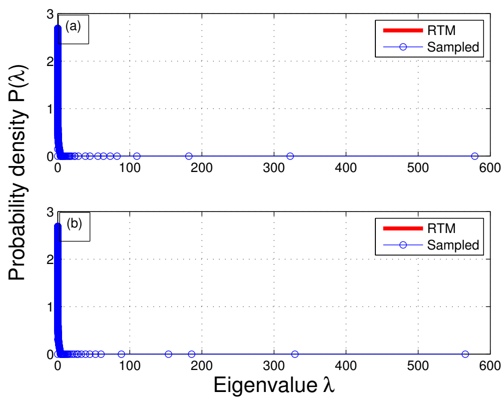

Fig. 2. Eigenvalue Distribution for the Correlation Matrix C for SenseCam data, 2096 images using 1st 2400 pixels series (a) 2096 images using 2nd 2400 pixels series (b)

C. Eigenvector Analysis

The deviations of $P(\lambda)$ from the RMT result $P_{rm}(\lambda)$ suggest that these deviations should also be displayed in the statistics of the corresponding eigenvector components [15]. Accordingly, in this section, we analyse the distribution of eigenvector components. Fig. 3 shows the spectrum of the components of the largest, the second largest and the third largest eigenvector. Our previous research [7] has shown that the largest Eigenvalue presents information from the image that reflects the largest change in the SenseCam recording. Similarly, we expect that the largest eigenvector can be interpreted as major events or key sources in the data corpus. As Fig. 3 illustrates, the different features can be found at various points in the eigenvectors spectrum, suggesting that the eigenvectors represent different major events or key sources with different eigenvector components. Compared to the third largest eigenvector, the first and second largest eigenvector have larger fluctuations in the eigenvector spectrum. This implies that the first and second eigenvector captured some different major events or key sources.

Fig. 4 shows that $P(C_{ij})$ is asymmetric, with a long positive tail and has a high peak, implying that positively correlated behaviour is more prevalent than negatively correlated positively. This is consistent with our previous research [7]. We argue that the tail represents significant or unusual events in the data stream. In order to remove the effects of the largest eigenvalue we use the techniques described in Section III

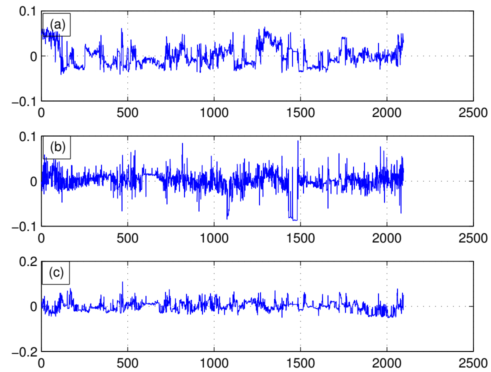

Fig. 3. Eigenvector components spectrum, the largest eigenvalue (a), 2nd largest eigenvalue (b) and 3rd largest eigenvalue (c)

(B). In Fig. 4 the distribution $P(C_{ij})$ thus obtained has a smaller significantly average value $<C_{ij}>$, showing that a degree of cross correlations contained in C can be attributed to the influence of the largest eigenvalue and its corresponding eigenvector.

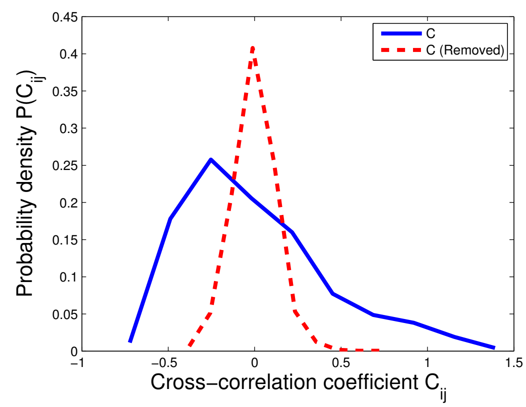

Fig. 4. Probability distribution P of the cross-correlation coefficients for data before (blue) and after (red) removing the effect of the largest eigenvalue by linear regression method

The distribution of the components of the largest remaining deviating eigenvectors shows some distinctive clustering in Fig. 5(a). In particular, Events $2, 6\&12$, Events 4&5 and Events 3&7 (in Table I) are the major contributors here. By examination of the images, we find that each clustering reflects quite similar light levels for grouping of events. For instance in the clustering for Events $2, 6\&12$: Event 2 described the user going from outside to the office; The office is dark; When the user switched on the lights, the light level suddenly changed. Also, quite similar situations occurred in Events 6 and 12, when the user was walking in the building, sometimes the camera captured the lights and sometimes not. When the user was preparing to leave the office, she was packing her stuff and stood up, and the camera captured the lights form the roof. Also several images captured the user's bag (black colour) and this makes a big difference in the sequence of images. Events 3&7 describe the same scenario: the user was sitting in front of her laptop, with laptop, lights and seating position unchanged over on extended period, contributing same pixel values in this sequence of images. However, although Event 11 also described the same scenario (working), it have not been grouped in the same cluster. By re-examination of these images, we note that this scenario is different to Events 3&7. This working process lasted for a very long time when the user was mainly working, but several times the camera was totally blocked by the user's hair or the desk. A similarly clustered distribution also emerges for the other deviating eigenvalues, shown in Fig 5 (b) and (c). However, these eigenvalues show no variation when compared to the largest remaining deviating eigenvector. Events $2, 5, 6, 8, 9\&12$, $1, 3\&10$ and $1, 2, 5, 8\&12$, $3, 6\&10$ mainly contribute for the second and third largest remaining deviating eigenvector, representatively. This implies that these eigenvectors could carry additional information on the description of the events such as possible lead-in, lead out.

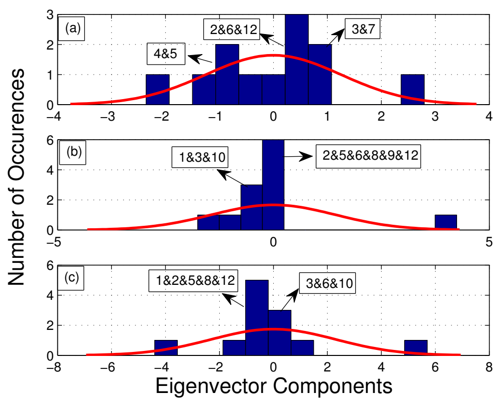

Fig. 5. Distribution of eigenvector components, largest remaining eigenvalue (a), 2nd largest remaining eigenvalue (b) and 3rd largest remaining eigenvalue (c)

V. Conclusions

A promising method to assist dementia patients in coping with their disability to retrieve autobiographical events is the creation of visual lifelogs that allow them to re-live these events. Visual lifelogs are personal image collections that are recorded on a frequent basis to capture the wearer's life activities. Facing vast amounts of images that have been recorded automatically, an open research challenge is the identification of key events in the data stream, i.e., events that are potentially more memorable than other events. In this paper, we studied whether random matrix theory (RMT) can be applied to extract meaningful information and noise from such data corpus. Significant deviations from RMT predictions are observed. Further, we analyze the deviations from RMT, and find that (a) different eigenvectors have different features, suggesting that different major events are captured by various eigenvectors. (b) By examining the eigenvectors corresponding to the images, we find that alternating light levels as key source are always picked up by the device for the largest remaining deviating eigenvector. In addition, the same events have been grouped together. The distinctive clustering captured the same events. (c) The second and third largest remaining deviating eigenvectors have a similarly-clustered distribution to the largest remaining deviating eigenvector. However, the different clusters consist of different events. This imply that these eigenvectors carry additional information on the description of the events for leading-in or leading out to major events. (d) We also note that even quite similar scenarios but key sources are slightly different, these events have been classified by different clustered distributions. This implies that key sources have played a major part for the classification by the cross-correlation matrix. The proposed cleaning technique of separating the noisy part from the non-noisy part has been proved useful. Overall, the RMT technique provides a powerful tool to analyze cross correlations across whole data streams. Future work includes evaluation large of datasets and assessment of the eigenvalues of C within the RMT bound for universal properties of random matrices, in order to confirm initial results and further explore the detailed features of the SenseCam images.

Acknowledgment

NL would like to acknowledge generous support from the Sci-Sym Centre Small Scale Research Fund, as well as additional support from the School of Computing, Dublin City University.

References

[1] E. O'Shea and S. O'Reilly, "The economic and social cost of dementia in ireland," *International Journal of Geriatric Psychiatry*, vol. 15, no. 3, pp. 208–218, 2000.

[2] C. Berr, J. Wancata, and K. Ritchie, "Prevalence of dementia in the elderly in europe," *European Neuropsychopharmacology*, vol. 15, no. 4, pp. 463 – 471, 2005.

[3] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. R. Wood, "Sensecam: A retrospective memory aid," in *Ubicomp*, 2006, pp. 177–193.

[4] P. Piasek, K. Irving, and A. F. Smeaton, "Sensecam intervention based on cognitive stimulation therapy framework for early-stage dementia." in *PervasiveHealth*. IEEE, 2011, pp. 522–525.

[5] E. Berry, N. Kapur, L. Williams, S. Hodges, P. Watson, G. Smyth, J. Srinivasan, R. Smith, B. Wilson, and K. Wood, "The use of a wearable camera, sensecam, as a pictorial diary to improve autobiographical memory in patient with limbic encephalitis: A preliminary report." *Neuropsychological Rehabilitation*, vol. 17, no. 4-5, pp. 582–601, 2007.

[6] A. R. Doherty and A. F. Smeaton, "Automatically segmenting lifelog data into events," in *WIAMIS*, 2008, pp. 20–23.

[7] N. Li, M. Crane, H. J. Ruskin, and C. Gurrin, "Multiscaled cross-correlation dynamics on sensecam lifelogged images," in *MMM (1)*, 2013, pp. 490–501.

[8] E. P. Wigner, "On the statistical distribution of the widths and spacings of nuclear resonance levels," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 47, pp. 790–798, 10 1951.

[9] F. J. Dyson, "Statistical theory of the energy levels of complex systems i," *Journal of Mathematical Physics*, vol. 3, pp. 140–157, 1962.

[10] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral, T. Guhr, and H. E. Stanley, "Random matrix approach to cross correlations in financial data," *Physical Review E*, vol. 65, no. 6, pp. 066 126+, Jun. 2002.

[11] A. M. Tulino and S. Verdú, "Random matrix theory and wireless communications," *Foundations and Trends in Communications and Information Theory*, vol. 1, no. 1, 2004.

[12] G. S. Chirikjian, "Multivariate statistical analysis and random matrix theory," *Applied and Numerical Harmonic Analysis*, vol. 2, pp. 229–270, 2012.

[13] M. O. Ulfarsson and V. Solo, "Dimension estimation in noisy pca with sure and random matrix theory," *IEEE Transactions on Signal Processing*, vol. 56, no. 12, pp. 5804–5816, 2008.

[14] T. Conlon, H. Ruskin, and M. Crane, "Random matrix theory and fund of funds portfolio optimisation," *Physica A: Statistical Mechanics and its Applications*, vol. 1, no. 2, pp. 565–576, 2010.

[15] T. Conlon, H. J. Ruskin, and M. Crane, "Multiscaled cross-correlation dynamics in financial time-series," *Advances in Complex Systems (ACS)*, vol. 12, no. 04, pp. 439–454, 2009.