View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by DCU Online Research Access Service

Random Matrix Ensembles of Time Correlation Matrices to Analyze Visual Lifelogs

Na Li, Martin Crane, Heather J. Ruskin, Cathal Gurrin

School of Computing, Dublin City University, Ireland
na.li@dcu.ie, {mcrane,hruskin,cgurrin}@computing.dcu.ie

Abstract. Visual lifelogging is the process of automatically recording images and other sensor data. Such lifelogs are usually created using wearable cameras. 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.

Keywords: Random Matrix Theory, Cross-correlation Matrix, Eigenvalues and Eigenvectors, SenseCam

1 Introduction

Lifelogging is the process of automatically recording aspects of one's life in digital form. This includes visual lifelogging using wearable cameras such as the SenseCam. The SenseCam [1], developed by Microsoft Research in Cambridge, UK, is a small wearable device that is worn via a lanyard suspended around the neck. The SenseCam incorporates a digital camera and multiple feeds, including sensors to detect changes in light levels, an accelerometer to detect motion, a thermometer to detect ambient temperature, and a passive infrared sensor to detect the presence of a person. The device takes pictures at VGA resolution, (480x640 pixels), and stores these as compressed JPEG files on internal flash memory. All sensor data and captured SenseCam images can be downloaded to a standard PC via a USB cable. SenseCam can 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 (logged every second).

Experience shows that the SenseCam can be an effective memory-aid device [2], as it helps users to improve recollection of an experience. 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. [3] address this challenge by identifying distinct events within a full day, (which typically consists of 2,000 lifelog images) *e.g. breakfast, working on PC, meeting, etc.*. However, their approach still contains a significant percentage of routine events. Li et al. [4] 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 [5–8]. 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).

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).

In detail, we address this challenge as follows. First, we analyze the distribution of the correlation coefficients of the cross-correlation matrix C, which reveals an asymmetric, long positive tail with a high peak, implying that some background information is picked up constantly by the device. Next we apply RMT methods to analyse the cross-correlation matrix C and show that a number ($\approx 20\%$) of the eigenvalues of C agree with RMT predictions, suggesting that C may have a "random band" matrix structure, i.e., "noise" element. Further, using the inverse participation ratio concept, we analyze the eigenvectors of the cross-correlation matrix and find that both edges of the eigenvalue spectrum of C (smallest and largest eigenvalues) deviate from RMT prediction. We argue that these deviant components represent significant or unusual events in the data stream. Finally, the cross-correlation matrix is cleaned by separating the noisy part from the non-noisy part of C in order to remove the user's routine events.

This paper is organized as follows: In Section 2, we discuss related work in the fields of lifelogging and random matrix theory. Section 3 provides a brief overview of the data that are used within this work to study the research hypothesis. In Section 4, we discuss the use of random matrix theory in the extraction of information from a correlation matrix of SenseCam image time series. In Section 5, we discuss the statistics

of cross-correlation coefficients and compare the eigenvalue distribution of the cross-correlation matrix with RMT results. In addition, we detail the analysis of the contents of eigenvectors that deviated from RMT. Finally, Section 6 provides conclusions on the work and anticipates future directions.

2 Related Work

This work builds on two research streams: (1) The creation and analysis of visual lifelogs and (2) the application of random matrix theory. In the remainder of this section, we briefly introduce both aspects.

2.1 Visual Lifelogging

As stated, visual lifelogging is the process of automatically capturing images and storing them in a personal repository. Although technologies for visual lifelogging has existed for several years, the use of such devices so far has been explored mainly by early adopters and researchers, in terms of studying its role as memory aid [2]. However, a recent study [9] suggests that lifelogging is an emerging trend with wider applications and an increasing member of devices will be available in the near future. A prominent example is Google Glass [10], which has received considerable media attention, since its first introduction to the public.

The majority of past research in the visual lifelogging domain has focused on issues of hardware miniaturisation [11] and also storage of images [12]. However, these challenges have been comparatively well addressed [1], resulting in improved wearability of devices and inexpensive storage [2]. The challenge is now that of retrieving *relevant* information from the vast quantities of captured data [2, 13–15].

A previous method, used to review images captured by the SenseCam is SenseCam Image View [1]. However, it can take two minutes or more to browse through a day's worth of SenseCam images, which means to 15 minutes to review all the images from one week. Therefore, we propose creation of one page 'visual summaries' of a day containing different images representing activities and events experienced by the user. Other research in this area has seen segmentation of the lifelog of SenseCam images into approximately 20 distinct events in a wearer's day, which translates to over 7,000 events per year [3]. Nevertheless, this large collection of personal information still contains a significant percentage of routine events. The objective is thus to determine which events are the most important or unusual for the lifelogger. In this paper, we address this issue by applying random matrix theory, described in the next subsection.

2.2 Random Matrix Theory

Random matrix theory (RMT) was first introduced by Wigner, Dyson, Mehta, and others [5–8] who aimed to study the energy levels of complex atomic nuclei [16]. Devi-

ations from the *universal* predictions of RMT can be used to identify system specific, non-random properties of the system under consideration, providing clues about the underlying interactions [6–8].

Its successful application to atomic physics has stimulated its use in many other fields, including number theory and combinatorics, wireless communications [17], and in multivariate statistical analysis and principal components analysis [18, 19], as well as for financial and other large dimensional data analysis [16, 20] . Applications of RMT methods to analysis of the properties of the cross-correlation matrix C show that a large proportion of eigenvalues of C agree with RMT predictions, indicating a considerable degree of randomness in measured cross correlations. Deviations from RMT predictions are, however also observed.

In the context of visual lifelogging, it has been shown that SenseCam image time series reflect strong long-range correlation [4] which suggests continuous low levels of background information picked up constantly by the device. In this paper, we investigate the use of RMT applied to a cross-correlation matrix C to detect details on SenseCam lifelog data streams.

3 Data

For this study, we analyzed 2101 lifelog images, recorded using a SenseCam over the period of one day. The wearer of the camera, i.e., the lifelogger, experienced an average day: 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. Fig. 1 shows some examples of SenseCam images that have been recorded on that day. Given the size of the test corpus and its content, we can argue that it is a typical visual lifelogging collection depicting a typical day of the subject's life.

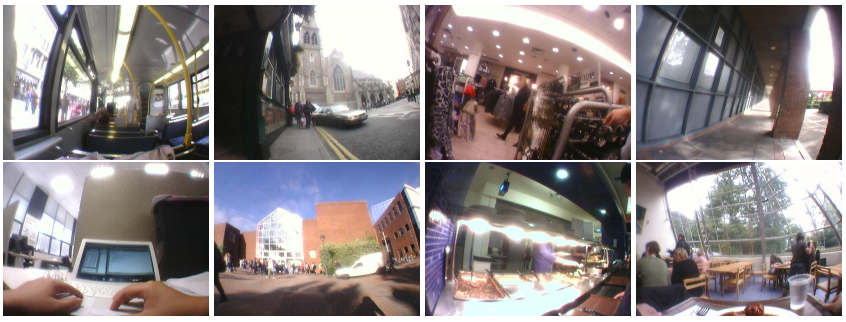

Fig. 1. Example of SenseCam Images

As discussed before, 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 [21]. Efforts to determine automatically which events are 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 [17–19, 16, 20]. In the next section, successful the method is outlined.

4 Methods

4.1 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 image size from 480×640 pixels to 60×80 pixels. Given pixels $G_i(t)$, $i = \{1, .., N\}$, of a collection of images, we normalized $G_i(t)$ in order to standardize the different pixels for the images as follows:

$$g_i(t) = \frac{G_i(t) - \overline{G_i(t)}}{\sigma_{(i)}} \tag{1}$$

Where $\sigma_{(i)}$ is the standard deviation of G_i for image numbers $i = \{1, .., N\}$ and $\overline{G_i}$ is the pixels' average of G_i over total pixel values T.

Then, the equal-time cross-correlation matrix [22] may be expressed in terms of $g_i(t)$

$$C_{ij} \equiv \left\langle g_i(t) g_j(t) \right\rangle \tag{2}$$

The elements of C_{ij} are limited to the domain -1≤C_{ij}≤1,where C_{ij}=±1 defines perfect positive/negative correlation and C_{ij}=0 corresponds to no correlation. In matrix notation, the correlation matrix can be expressed as

$$C = \frac{1}{T} G G^{\tau} \tag{3}$$

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 [16],

$$R = \frac{1}{T} A A^{\tau} \tag{4}$$

Where A is an $N \times T$ matrix with each element randomly 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:

$$P_{rm}(\lambda) = \frac{Q}{2\pi\sigma^2} \frac{\sqrt{(\lambda_+ - \lambda)(\lambda - \lambda_-)}}{\lambda} \tag{5}$$

for $\lambda_- \leq \lambda_i \leq \lambda_+$, where λ_- and λ_+ are the minimum and maximum eigenvalues of R, respectively, given by

$$\lambda_\pm = \sigma^2(1 + \frac{1}{Q} \pm 2\sqrt{\frac{1}{Q}}) \tag{6}$$

Then, σ^2 is the variance of the elements of G and $\lambda_\pm$ are the bounds of the theoretical eigenvalue distribution. Eigenvalues that fall outside this region are said to deviate from the expected values of the Random Matrix. Hence, by comparing the empirical distribution of the eigenvalues of the correlation matrix to the distribution for a random matrix, as given in Eq. (5), 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.

4.2 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 first remove the effect of the largest eigenvalue. To do this we use the linear regression [16]

$$G_i(t) = \alpha_i + \beta_i G^{large}(t) + \epsilon_i(t) \tag{7}$$

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 $\lambda_\pm$.

As suggested in [16], we also aim to assess whether random effects are less marked further away from RMT upper boundary λ_+. To do this we use the Inverse Participation Ratio (IPR). The IPR allows quantification of the number of components that participate significantly in each eigenvector and tells us more about the level and nature of the deviation from RMT. The IPR of the eigenvector u^k is given by $I^k \equiv \sum_{l=1}^{N} [u_l^k]^4$ and allows us to compute the inverse of the number of eigenvector components that contribute significantly to each eigenvector.

5 Results

5.1 Statistics of Correlation Coefficients

In order to quantify correlations, we first analyse the distribution $P(C_{ij})$ of the elements $\{C_{ij} : i = j\}$ of the cross-correlation matrix *C* calculated by Eq. (3) and distribution $P(R_{ij})$ of the elements $\{R_{ij} : i = j\}$ of the random matrix calculated by Eq. (4). Fig. 2 shows that $P(R_{ij})$ is consistent with a Gaussian with zero mean, in contrast to $P(C_{ij})$. We note 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 [4]. We argue that the tail represents significant or unusual events in the data stream. In addition, we see that the $P(C_{ij})$ falls within the Gaussian curve for the control, suggesting the possibility that observed similarities with *R* in the cross-correlation matrix *C* may be an effect of randomness.

5.2 Eigenvalue Analysis

As stated above, our aim is to distinguish between information (major events) and noise in the cross-correlation matrix *C*, so we compare the eigenvalue distribution $P(\lambda)$ of C with $P_{rm}(\lambda)$ for N = 2101 images, each containing T = 4800 pixels, Thus, $Q = T/N$ = 2.28, and we obtain λ_- = 0.11 and λ_+ = 2.76 from Eq. (6). We compute the eigenvalues λ_i of C, where λ_i are rank ordered ($\lambda_{i+1}>\lambda_i$). Fig. 3 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 [λ_-, λ_+] for $P_{rm}(\lambda)$. We also note deviations for ($\approx 80\%$) largest and smallest eigenvalues. Thus suggests that the cross-correlation matrix has captured most major events from the data streams, but still contains some percentage of noise ($\approx 20\%$).

5.3 Eigenvector Analysis

The deviations of $P(\lambda)$ from the RMT result $P_{rm}(\lambda)$ suggests that these deviations should also be observed in the statistics of the corresponding eigenvector components [23]. Accordingly, in this section, we analyse the distribution of eigenvector components. The distribution of the components $\{u_l^k; l = 1, ..., N\}$ of eigenvector U^k of a random correlation matrix R should conform to a Gaussian distribution with zero mean and unit variance. First, we compare the distribution of eigenvector components of C with a Gaussian distribution. We analyse $P(u)$ for *C* computed for the total of 2101 images. We choose one typical eigenvalue λ_k from the bulk ($\lambda_- \leq \lambda_k \leq \lambda_+$) defined by $P_{rm}(\lambda)$ from Eq. (5). It shows that $P(u)$ for a typical U^k from the bulk shows nearly total agreement with the RMT result $P_{rm}(u)$. Similar analysis on the other eigenvectors, belonging to eigenvalues within the bulk, yields consistent results, (in agreement with those of the previous sections) to random matrix predictions. We test the agreement of

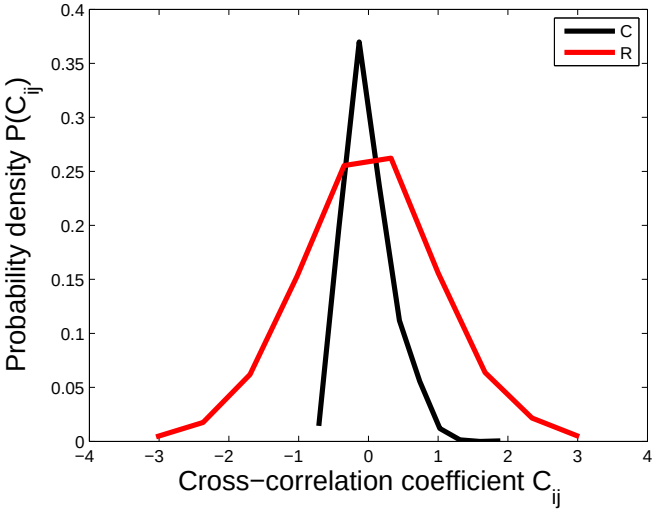

Fig. 2. Correlation Coefficients Distribution for Correlation Matrix C for SenseCam data (black) and Random Matrix R (red).

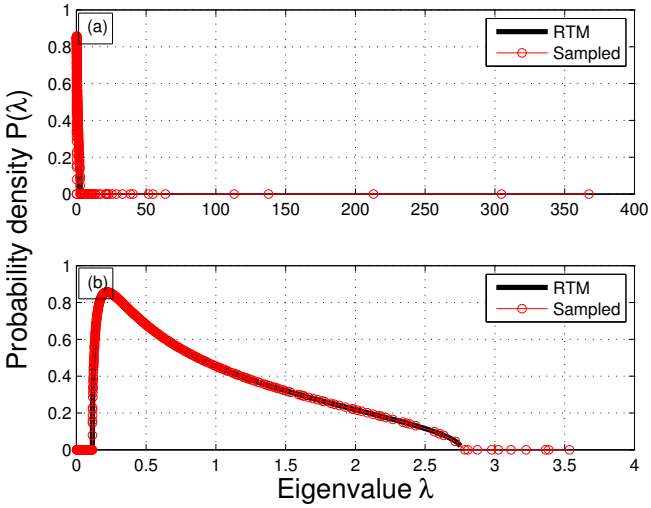

Fig. 3. Eigenvalue Distribution for the Correlation Matrix C for SenseCam data, Full spectral distribution (a) Partial spectral distribution (b)

the distribution $P(U)$ with $P_{rm}(u)$ by calculating the kurtosis, which for a Gaussian has the value 3. We find that the largest eigenvector (≈ 4.07) significant deviates from the Gaussian value. The second and third Eigenvectors ($\approx 3.7, 3.8$) are also larger than the

Gaussian value. The Eigenvector from the bulk is however consistent with the Gaussian value 3. These findings suggest that the largest eigenvalue (corresponding to the largest eigenvector) present information from the image that reflects the largest change in the SenseCam recording.

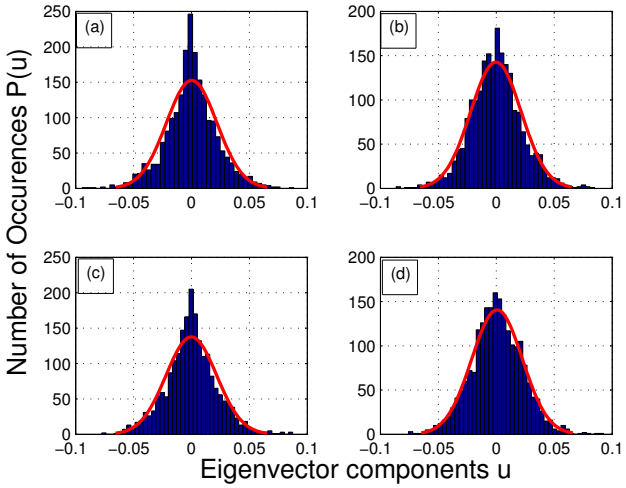

Fig. 4. Comparison of Eigenvector Components, largest Eigenvector (a), second largest Eigenvector (b), third largest Eigenvector (c) and Eigenvector from the bulk (d)

In order to remove the effects of the largest eigenvalue we use the techniques described in Section 3.2. We remove the contribution of $G^{large}(t)$ to each time series $G_i(t)$, and construct C from the residuals $\epsilon_i(t)$ of Eq. (7). Fig. 5 shows that the distribution $P(C_{ij})$ thus obtained has a smaller 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.

Having studied the largest eigenvalue and noting that if deviates significantly from RMT results, we conclude that it reflects the largest change in the SenseCam data streams. So next focus on the remaining eigenvalues to see whether these relate also to key sources or major events and what information these contribute additionally to the images. The Inverse Participation Ratio (IPR) quantities the reciprocal of the number of eigenvector components that contribute significantly. Fig. 6 (a) shows I^k for the case of the control of Eq. (4). The average value of I^k is $<I>\approx 0.0014 \approx 1/N$ with a very narrow spread, indicating that the vectors are extended [24] - i.e., almost all components elements (of the vector) contribute. Fluctuations around this average value are confined to a narrow range. Fig. 6 (b) show I^k for the cross-correlation matrix constructed from the 2101 SenseCam images. The edges of the eigenvalue spectrum of C show significant deviations of I^k from $<I>$, indicating that there are major events contributing to these

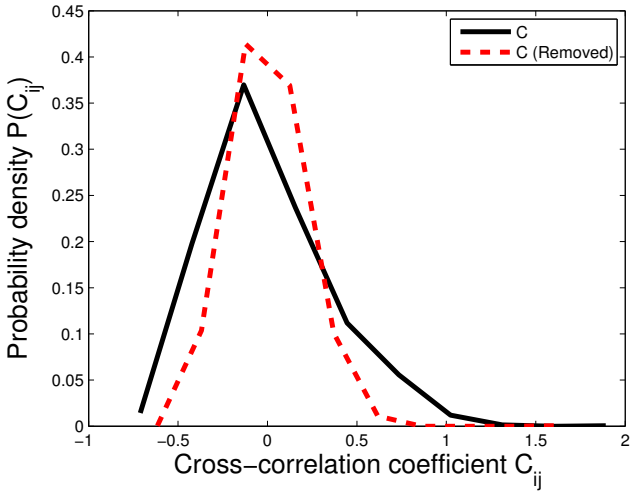

Fig. 5. Probability distribution P of the cross-correlation coefficients for data before (black) and after (red) removing the effect of the largest eigenvalue by linear regression method

eigenvectors. In addition, we also find that there are a number of small eigenvalue deviations from the control case, which suggests that the vectors are localized [24] - i.e. only a few images contribute to them.

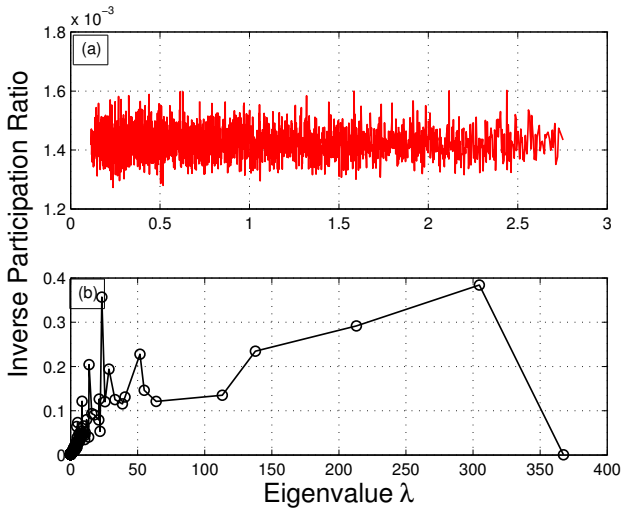

Fig. 6. Inverse Participation Ratio (IPR) as function of eigenvalue λ for the random cross-correlation matrix R (a) and cross-correlation matrix C (b)

Examination, of the eigenvalue and eigenvector content, indicates that the percentage of ($\approx 20\%$) noise for that period is described by the wearer working in front of the laptop for a long time without performing any other activities. The deviating eigenvalues from the RMT upper bound are described by the wearer morning from in front of the laptop and preparing to go home, with every image capturing different moments of this event. For example, considerable light is captured at the moment of standing up, different colours appear when she turns around, etc. Although all these images are visually very diverse, all have been captured at the same space, i.e., in the office. The deviating eigenvalues from the RMT lower bound involve several different activities, events such as commuting from home to the work place, the wearer talking with her colleague, the wearer sharing lunch with her colleague etc. Only a few images depict the same environment, but these appear on few images only. We argue that this confirms our observation that the IPR shows that the smallest eigenvalue deviations are localized, when only a few images contribute to them.

6 Conclusions

To summarise, we have illustrated that RMT, even with rather limited data, (2101 lifelog images depicting a typical day for a lifelogger) can be applied to extract the information (major events) and noise from a cross-correlation matrix. Significant deviations from RMT predictions are observed. In analysing these deviations find that (a) the largest eigenvalue and its corresponding eigenvector present information from the image that reflects the largest change in the SenseCam recording, and (b) the smallest eigenvalues, and their corresponding eigenvectors, represent short duration major events from the SenseCam recording. The 'cleaning technique' (of separating the noisy part from the non-noisy part) is demonstrated to be useful. Overall, RMT provides a powerful tool to analyse 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.

Acknowledgments

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, DCU.

References

1. Hodges, S., Williams, L., Berry, E., Izadi, S., Srinivasan, J., Butler, A., Smyth, G., Kapur, N., Wood, K.R.: Sensecam: A retrospective memory aid. In: Ubicomp. (2006) 177–193

2. Bell, G., Gemmell, J.: A digital life. Scientific American (2007)
3. Doherty, A.R., Smeaton, A.F.: Automatically segmenting lifelog data into events. In: WIAMIS. (2008) 20–23
4. Li, N., Crane, M., Ruskin, H.J., Gurrin, C.: Multiscaled cross-correlation dynamics on sensecam lifelogged images. In: MMM (1). (2013) 490–501
5. Wigner, E.P.: On the statistical distribution of the widths and spacings of nuclear resonance levels. Mathematical Proceedings of the Cambridge Philosophical Society **47** (10 1951) 790–798
6. Dyson, F.J.: Statistical theory of the energy levels of complex systems i. Journal of Mathematical Physics **3** (1962) 140–157
7. Dyson, F.J., Mehta, M.L.: Statistical theory of the energy levels of complex systems iv. Journal of Mathematical Physics **4** (1963) 701–712
8. Mehta, M.L., Dyson, F.J.: Statistical theory of the energy levels of complex systems v. Journal of Mathematical Physics **4** (1963) 713–719
9. Daskala, B., Askoxylakis, I., Brown, I., Dickman, P., Friedewald, M., Irion, K., Kosta, E., Langheinrich, M., McCarthy, P., Osimo, D., Papiotis, S., Pasic, A., Petkovic, M., Price, B., Spiekermann, S., Wright, D.: Risks and benefits of emerging life-logging applications. Final report, European Network and Information Security Agency (ENISA) (11 2011)
10. Goldman, D.: Google unveils project glass virtual-reality glasses. Money(CNN) (2012)
11. Mann, S.: Wearable computing: a first step toward personal imaging. Computer **30** (1997)
12. Gemmell, J., Bell, G., Lueder, R.: Mylifebits: a personal database for everything. Commun. ACM **49**(1) (2006) 88–95
13. Ashbrook, D., Lyons, K., Clawson, J.: Capturing experiences anytime, anywhere. IEEE Pervasive Computing **5**(2) (2006) 8–11
14. Lin, W.H., Hauptmann, A.: Structuring continuous video recordings of everyday life using time-constrained clustering. Multimedia Content Analysis, Management, and Retieval SPIE-IST Electronic Imaging **6073** (2006) 111–119
15. Lee, M.L., Dey, A.K.: Providing good memory cues for people with episodic memory impairment. In: ASSETS. (2007) 131–138
16. Plerou, V., Gopikrishnan, P., Rosenow, B., Nunes Amaral, L.A., Guhr, T., Stanley, H.E.: Random matrix approach to cross correlations in financial data. Physical Review E **65**(6) (June 2002) 066126+
17. Tulino, A.M., Verdú, S.: Random matrix theory and wireless communications. Foundations and Trends in Communications and Information Theory **1**(1) (2004)
18. Chirikjian, G.S.: Multivariate statistical analysis and random matrix theory. Applied and Numerical Harmonic Analysis **2** (2012) 229–270
19. Ulfarsson, M.O., Solo, V.: Dimension estimation in noisy pca with sure and random matrix theory. IEEE Transactions on Signal Processing **56**(12) (2008) 5804–5816
20. Conlon, T., Ruskin, H.J., Crane, M.: Random matrix theory and fund of funds portfolio optimisation. Physica A: Statistical Mechanics and its Applications (2) (2010) 565–576
21. Berry, E., Kapur, N., Williams, L., Hodges, S., Watson, P., Smyth, G., Srinivasan, J., Smith, R., Wilson, B., Wood, K.: The use of a wearable camera, sensecam, as a pictorial diary to improve autobiographical memory in a patient with limbic encephalitis. Neuropsychological Rehabilitation **17**(2) (2007) 580–601
22. Conlon, T., Ruskin, H.J., Crane, M.: Multiscaled cross-correlation dynamics in financial time-series. Advances in Complex Systems (ACS) **12**(04) (2009) 439–454
23. Conlon, T., Ruskin, H.J., Crane, M.: Cross-correlation dynamics in financial time series. Papers 1002.0321, arXiv.org (February 2010)
24. Lee, P.A., Ramakrishnan, T.V.: Disordered electronic systems. Rev. Mod. Phys. **57** (Apr 1985) 287–337