Finding Motifs in Larger Personal Lifelogs

Na Li, Martin Crane, Cathal Gurrin and Heather J. Ruskin
School of Computing, Dublin City University, Ireland
na.li3@mail.dcu.ie, {mcrane,cgurrin,hruskin}@computing.dcu.ie

ABSTRACT

The term *Visual Lifelogging* is used to describe the process of tracking personal activities by using wearable cameras. A typical example of wearable cameras is Microsoft's SenseCam that can capture vast personal archives per day. A significant challenge is to organise and analyse such large volumes of lifelogging data. State-of-the-art techniques use supervised machine learning techniques to search and retrieve useful information, which requires *prior* knowledge about the data. We argue that these so-called rule-based and concept-based techniques may not offer the best solution for analysing large and unstructured collections of visual lifelogs. Treating lifelogs as time series data, we study in this paper how motifs techniques can be used to identify repeating events. We apply the Minimum Description Length (MDL) method to extract multi-dimensional motifs in time series data. Our initial results suggest that motifs analysis provides a useful probe for identification and interpretation of visual lifelog features, such as frequent activities and events.

CCS Concepts

•**Information systems** → **Data analytics;** *Data management systems; Clustering; Summarization;* •**Human-centered computing** → Personal digital assistants;

Keywords

Lifelogging, TimeSeries, Motif, Minimum Description Length, equal-time Correlation Matrices; Maximum Overlap Discrete Wavelet Transform

1. INTRODUCTION

As wearable technology has become significantly cheaper, people increasingly rely on such devices to record profiles of individual behaviour by capturing and monitoring personal activities. This activity is often referred to as *lifelogging*, i.e., the process of automatically, passively and digitally recording aspects of our life experience. Considering the heterogeneous nature of the data created, as well as its appearance in form of constant data streams, lifelogging shares features that are usually attributed to big data. A special case of lifelogging is *visual lifelogging*, where lifeloggers wear cameras mounted on the head [14, 26] or chest [3, 35], that capture personal activities through the medium of images or video. Despite its relative novelty, visual lifelogging is gaining popularity due to projects such as the Microsoft SenseCam [13]. The SenseCam is a small, lightweight wearable device that automatically captures a wearer's every moments as a series of images and sensor readings. Normally, the SenseCam captures an image at the rate of one every 30 seconds and collects about 4,000 images in a typical day. However, since the SenseCam generates a very large amount of data for a single day, a significant challenge is to manage, organise and analyze the large lifelogging data set in order to automatically categorise a wearer's characteristics.

To date, various aspects of lifelogging have been studied, such as the development of sensors, efficient capture and storage of data, processing and annotating the data to identify events [8], improved search and retrieval of information [41], assessment of user experience, design of user interfaces for applications of memory aids [13], diet monitoring [33], or analysis of activities of daily living (ADL) [27].

Given the relative success of these efforts, the research challenge has now shifted from broader aspects of data management to that of retrieving *refined and relevant* information from the vast quantities of captured data [10, 2, 24]. Current applications address this by employing automatic classifiers for segmenting a whole day's recording into events and searching the historical record [7], or by building ontology-based multi-concept classifiers and searching for specific events [40]. More recent research suggests use of statistical mapping from low-level visual features to semantic concepts of personal lifelogs [41]. It is important to note that these approaches are based on training classifiers from a set of annotated ground truth images. Although supervised methods can lead to more accurate outcome in terms of detecting known patterns, they require prior knowledge from a domain expert to be fed into the system. In addition, the result for the classifier depend heavily on the quality and quantity of the training data, i.e. are biased to detection of activities that are defined and known to the domain expert a priori. Given that visual lifelogs usually consist of large and often unstructured collections of multimedia information, such a 'concept-based' and 'rule-based method' for analysing lifelogging data is not suitable for all use-cases. Ideally, an algorithm should be able to detect unknown phenomena occurring at different frequencies in such data. In our previous works, we introduced and evaluated the use of sophisticated time series analysis methods for management of large lifelogging data sets [20, 21, 19]. The results of this evaluation suggest that strong correlations do exist in these time series, with

AH'16 February 25-26, Geneva, Switzerland

ACM ISBN 978-1-4503-2138-9.
DOI: 10.1145/1235

recognisable cycles representing specific events.

In this paper, we build on this observation to address the challenge of refining the analysis of lifelogs by studying time series motifs. Such motifs are frequently occurring, but often previously unknown, subsequences of longer time series [22]. Similar analyses feature widely in various medical applications, including examining data from on-body monitoring sensors [28] and selecting maximally informative genes [1], protein sequence identification [30] and others. Time series motifs are used also for finding patterns in sports motion capture data [38] and in video surveillance applications [12]. Many researchers have studied the extraction of characteristics features from multi-dimensional time series data. In Tanaka et al. [38], Principle Component Analysis (PCA) is used to transform multidimensional time series to one dimensional to detect motifs that are common to all. More recently, Minnen et al. [29] extended a motif discovery method for single time series to detection of motifs occurring across several dimensions of a multi dimensional signal. Visual lifelogs contain records of a wearer's activities and events that occur over different time periods. Consequently, we argue that motifs can represent activities of different length and timing in data resembling attributes of big data. We explore this idea by analysing high frequency patterns in multi-dimensional visual lifelogging data.

This paper is organized as follow: In Section 2, we introduce and describe our method of motifs detection in lifelogging data. Section 3 presents an experiment performed to asses the technique and present results obtained. Conclusions are given in Section 4.

2. METHODS

In prior work [20] we studied the spectral dynamics of SenseCam images by applying the multi-scaled cross-correlation matrix technique, for which time series exhibit atypical or non-stationary characteristics, symptomatic of "Distinct Significant Events" in the data. Our study suggests that we can use such *key episodes* to identify boundaries between different daily events. Results to date indicate that different distinct events or activities can be detected at different scales through wavelet analysis. Building on this observation, we aim in this paper to extract the motifs in different wavelet scales using the Minimum Description Length (MDL) principle. In this section, therefore, we first give a review of the Cross-correlation matrix structure and the Maximum Overlap Discrete Wavelet Transform (MODWT). Then, we introduce the Symbolic Aggregate approXimation (SAX) algorithm for discretization of time series data into symbolic strings. Finally, we detail our motif extraction algorithm, based on the MDL principle.

2.1 Equal-time cross-correlation matrix and Maximum Overlap Discrete Wavelet Transform

The behaviour of the largest eigenvalue of a cross-correlation matrix over small windows of time, has been studied for financial series [37, 15, 34, 6, 18, 32, 31], electroencephalographic (EEG) recordings [11], magnetoencephalographic (MEG) recordings [39] and a variety of other multivariate data. Similar techniques are used here to investigate the dynamics of SenseCam images.

To reduce the size of the calculation further, and thus the amount of memory used, we first adopt an averaging method to decrease the image size from 480x640 pixels to 6x8 pixels. Hence the correlation matrix is made up of 48 time series over the total number of images. The equal-time cross-correlation matrix, between time series of images, is calculated using a sliding window, where the number of pixels in one image, N, is smaller than the window size T. Given pixels $G_i(t)$, i={1,...,N}, of a collection of images, we normalise G_i within each window in order to standardise 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 time average of G_i over a time window of size T. Then the equal-time cross-correlation matrix 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$\leq C_{ij} \leq$1,where $C_{ij}=\pm 1$ defines perfect positive/negative correlation and C_{ij}=0 corresponds to no correlation. In matrix notation, the correlation matrix can be expressed as $C = \frac{1}{T} GG^t$ where t is the transpose of a matrix and G is an $N \times T$ matrix with elements g_{it}.

The eigenvalues λ_i and eigenvectors $\overline{v}_i$ of the correlation matrix C are found from the eigenvalue equation $C\overline{v_i} = \lambda_i \overline{v_i}$.

The eigenvalues are then ordered by size, such that $\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_N$. Given that the sum of the diagonal elements of a matrix (the Trace) remains constant under linear transformation [11], $\sum_i \lambda_i$ must always equal the Trace of the original correlation matrix. Hence, if some eigenvalues increase then others must decrease, to compensate, and vice versa, (a feature known as *Eigenvalue Repulsion* [9]).

There are two limiting cases for the distribution of the eigenvalues: (i) perfect correlation, $C_i \approx 1$, when the largest is maximised with value N, (all others taking value zero). (ii) when each time series consists of random numbers with average correlation $C_i \approx 0$ and the corresponding eigenvalues are distributed around 1, (where any deviation is due to spurious random correlations). Between these two extremes, the eigenvalues at the lower end of the spectrum can be much smaller than λ_{max}. To study the dynamics of each of the eigenvalues using a sliding window, we normalise each eigenvalue in time using

$$\tilde{\lambda}_i(t) = \frac{(\lambda_i - \overline{\lambda})}{\sigma^{\lambda}} \tag{3}$$

where $\overline{\lambda}$ and σ^{λ} are the mean and standard deviation of the eigenvalues over a particular reference period. This normalisation allows us to visually compare eigenvalues at both ends of the spectrum, even if their magnitudes are significantly different. The reference period used to calculate the mean and standard deviation of the eigenvalue spectrum can be chosen to be a low volatility sub-period, (which helps to enhance the visibility of high volatility periods), or the full time-period studied.

The wavelet transform (WT) is a mathematical tool that can be applied in many areas such as image analysis [43], meteorology [5], signal processing [25] and financial time series [4] and is used to decompose a signal into different time horizons. For example, wavelets allow us to decompose a signal on a Scale-by-Scale basis, e.g., in measuring the correlation between equities over different time scales (values at hourly intervals, two hourly intervals etc.). This allows characterization of the impact of the different trading

strategies or horizons of traders, on correlations between such equities. In particular, the discrete wavelet transform (DWT) [42] is useful in dividing the data series into components of different frequencies, so that each component can be studied separately in order to investigate the data series in depth. In our case, where we wish to compare different pixel time series values, we may do so over a range of time scales.

The Maximum Overlap Discrete Wavelet Transform, (MODWT) [42], is a linear filter that transforms a series into coefficients related to variations over a set of scales. Like the DWT it produces a set of time-dependent wavelet and scaling coefficients with basis vectors associated with a location t and a unitless scale τ_j=2^{j-1} for each decomposition level j= {1,...,J_0}. Unlike the DWT, the MODWT has a high level of redundancy. However, it is *non-orthogonal* and can handle any sample size N, whereas the DWT restricts the sample size to a multiple of 2^j. MODWT retains downsampled[1] values at each level of the decomposition that would be discarded by the DWT. This reduces the tendency for larger errors at lower frequencies, when calculating frequency dependent variance and correlations, as more data are available.

Decomposing a signal to J levels, using the MODWT, theoretically involves the application of J pairs of filters. The filtering operation at the j^{th} level consists of applying a rescaled *father wavelet*[2] to yield a set of detail coefficients

$$\tilde{D}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{\varphi}_{j,l} f_{t-l} \tag{4}$$

and a rescaled *mother wavelet*[3] to yield a set of scaling coefficients

$$\tilde{S}_{j,t} = \sum_{l=0}^{L_j-1} \tilde{\phi}_{j,l} f_{t-l} \tag{5}$$

for all times $t = \{..., -1, 0, 1, ...\}$, where f is the function to be decomposed [36]. The rescaled mother, $\tilde{\varphi}_{j,t}$=$\frac{\varphi_{j,t}}{2^j}$, and father, $\tilde{\phi}_{j,t}$=$\frac{\varphi_{j,t}}{2^j}$, wavelets for the j^{th} level are a set of scale-dependent localised differencing and averaging operators and can be regarded as rescaled versions of the originals. The j^{th} level equivalent filter coefficients have a width $L_j = (2^j - 1)(L - 1) + 1$, where L is the width of the $j = 1$ base filter. In practice the filters for $j > 1$ are not explicitly constructed because the detail and scaling coefficients can be calculated, using an algorithm that involves the $j = 1$ filters operating recurrently on the j^{th} level scaling coefficients, to generate the $(j + 1)^{th}$ level scaling and detail coefficients [36]. Each of the sets of coefficients in a wavelet is called a 'crystal'.

The wavelet variance $\nu_f^2(\tau_j)$ is defined as the expected value of $\tilde{D}_{j,t}^2$ if we consider only the non-boundary coefficients[4]. An unbiased estimator of the wavelet variance is formed by removing all coefficients that are affected by boundary conditions and is given by

$$\nu_f^2(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{D}_{j,l}^2 \tag{6}$$

where $M_j = N - L_j + 1$ is the number of non-boundary coefficients at the j^{th} level [36]. The wavelet variance decomposes the variance of a process on a scale-by-scale basis (at increasingly higher resolutions of the signal) and allows us to explore how a signal behaves over different time horizons.

The wavelet covariance between functions $f(t)$ and $g(t)$ is similarly defined to be the covariance of the wavelet coefficients at a given scale. The unbiased estimator of the wavelet covariance at the j^{th} scale is given by

$$\nu_{fg}(\tau_j) = \frac{1}{M_j} \sum_{t=L_j-1}^{N-1} \tilde{D}_{j,l}^{f(t)} \tilde{D}_{j,l}^{g(t)} \tag{7}$$

where all the wavelet coefficients affected by the boundary are removed [36], and $M_j = N - L_j + 1$.

The MODWT estimate of the wavelet cross-correlation between functions $f(t)$ and $g(t)$ may be calculated using the wavelet covariance and the square root of the wavelet variance of the functions at each scale j. The MODWT estimator, of the wavelet correlation is given by

$$\rho_{fg}(\tau_j) = \frac{\nu_{fg}(\tau_j)}{\nu_f(\tau_j)\nu_g(\tau_j)} \tag{8}$$

where, at scale j, $\nu_{fg}(\tau_j)$ is the covariance between $f(t)$ and $g(t)$, $\nu_f(\tau_j)$ is the variance of $f(t)$ and $\nu_g(\tau_j)$ is the variance of $g(t)$.

2.2 Dimensionality Reduction and Discretization

We use a dimensionality reduction algorithm based on Piecewise Aggregate Approximation (PAA) [22, 17] called Symbolic Aggregate approXimation (SAX) [23]. We apply this technique to transform the largest eigenvalue time series into a sequence of symbols. For the largest eigenvalue time series λ_1 with number of images n, this time series can be reduced to a string of arbitrary length w, (where $w < n$) and the alphabet size of arbitrary length a, (where $a > 2$). The largest eigenvalue time series $\lambda_1(t) = \{x_1, ..., x_n\}$ of length n can be represented as a w-dimensional space by a vector $\bar{\lambda}$ = $\{\bar{x}_1, ..., \bar{x}_w\}$:

$$\bar{x_i} = \frac{w}{n} \sum_{j=\frac{w}{n}(i-1)+1}^{\frac{w}{n}i} x_i \tag{9}$$

In order to transform the vector of w dimension into a sequence of "PAA symbols", it is necessary also to determine "breakpoints" that determine the range of the PAA value for assigning unique PAA symbols. One approach is to determine the breakpoints that will produce an equal-sized area under a Normal distribution. Breakpoints are a sorted list of numbers $B = \{\beta_1, ..., \beta_{a-1}\}$ such that

[1] *Downsampling* or *decimation* of the wavelet coefficients retains half of the number of coefficients that were retained at the previous scale. Downsampling is applied in the Discrete Wavelet Transform

[2] Father wavelet, also called scaling function, the scaling function filters the lowest level of the transform and ensures all the spectrum is covered.

[3] Mother wavelet, also called wavelet function, the wavelet function is in effect a band-pass filter and scaling it for each level halves its bandwidth.

[4] The MODWT treats the time-series as if it were periodic using "circular boundary conditions". There are L_j wavelet and scaling coefficients that are influenced by the extension, and which are referred to as the boundary coefficients.

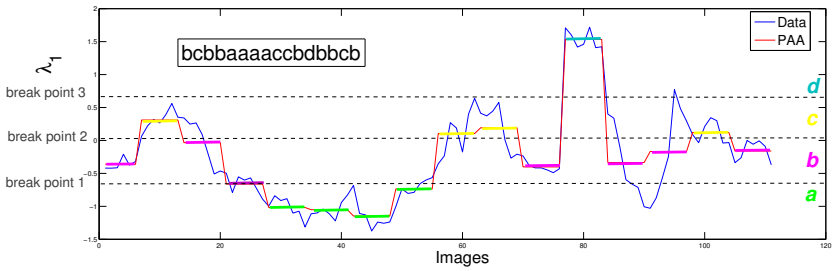

Figure 1: Example of time series transformation into SAX symbols. In this, with n=112, w=16 and a=4. the time series is mapped to the PAA symbols bcbbaaaaccbdbbcb.

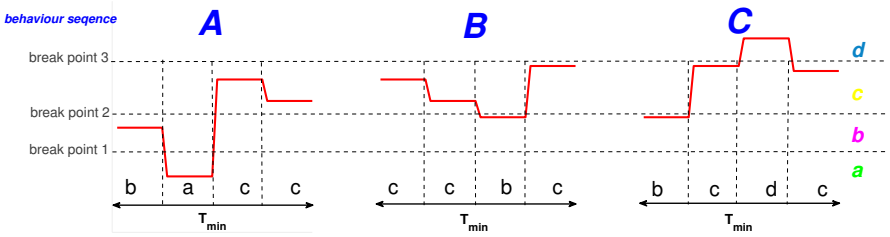

Figure 2: Example of "behaviour symbol" assignment for pattern order of PAA symbols. In this, A=bacc, B=ccbc and C=bcdc.

the area under a $N(0,1)$ Standard Normal distribution from β_i to $\beta_{i+1} = 1/a(\beta_0$ and β_a are defined as $-\infty$ and ∞, respectively. Once the breakpoints have been obtained we can discretize a time series as follows. We first obtain a PAA of the time series. All PAA coefficients that are below the smallest breakpoint are mapped to the symbol "a", all coefficients greater than or equal to the smallest breakpoint and less than the second smallest breakpoint are mapped to the symbol "b", etc. Fig. 1 illustrates the idea. Finally, a "behaviour symbol" is assigned for every subsequence of PAA symbols. An example is given in Fig. 2, for which the analysis window of T_{min} is the minimum length of motif for the time series.

2.3 Estimating Extracted Motif Candidate Based on MDL Principle

Several theoretical information theory principles from literature are relevant to the current analysis, including AIC (Akaike's Information Criterion), BIC (Bayesian Information Criterion) and MDL (Minimum Description Length) principles. The AIC estimates the best model based on "prediction capability", while BIC estimates the best model based on bayesian principles. Our approach is focused however, on finding frequent patterns, rather than prediction, for the time series. The MDL principle states that the best model to describe a set of data is that which minimises the description length of the entire data set. The underlying concept is the selection of the best model to compresses the data. Table 1 summarises the principal notation used in this sub-section.

As stated, we apply the SAX algorithm for transforming time series into behavior symbol sequences as BSS in Fig. 3 (a) for example. We can extract the pattern FPV from BSS. However, we found that from position 25 to 29 of the BSS, all sequence elements are 'A', so we can transform/compress these 'A' by noting BS length of 5 (as Fig. 3 (b)). This new sequence is the "modified BSS". There are three costs to define in calculating the MDL. The 'data encoding cost' is the lower bound of description length that is required to encode each segment. The 'parameter encoding cost' is the description length that is required to describe the order of BS in each segment. Finally, the 'segmentation cost' is required to describe the location of all segments. We summarise these costs in Table 2.

Table 1: A summarisation of the notation used in this subsection

BS	Behavior Symbol
BSS	Behavior Symbol Sequences
TSS	Time-Series Subsequence
T_{min}	Analysis window
$\tilde{C}$	A symbol representation of a time series
SC	Subsequence
DL	Description Length

In calculating the data encoding cost of the i-th segment, we calculate the length of the i-th segment t_i. For example, in Table 2, the length of the first segment is t_1=9, the length of the second segment is t_2=3 and so on. In addition, we assume that the j-th BS has a length l_{ij}. A data encoding cost for the j-th BS in the ith segment is calculated then as:

$$-l_{ij} log_2 \frac{l_{ij}}{t_i} \tag{10}$$

For example, in Table 2, the first BS of the second segment of the Modified BSS, Fig. 3 (a), is 'F', and the data encoding cost of 'F'is $-1 log_2 \frac{1}{3}$. By calculating the data encoding cost of all unique BSS in the i-th segment, we obtain the data encoding cost of the whole segment as:

$$\sum_j -l_{ij} log_2 \frac{l_{ij}}{t_i} \tag{11}$$

Using the following equation, we then calculate the data encoding

Table 2: Example of MDL calculation

Segment	Length	Data encoding cost	Parameter encoding cost
s1	9	$AX1 : -1log_2\frac{1}{9}$	$log_2 9$
s2	3	$FX1 : -1log_2\frac{1}{3}$ ...	$log_2 3$
s3	39	$TX1 : -1log_2\frac{1}{39}$ ... $AX5 : -5log_2\frac{5}{39}$ ...	$log_2 39$
...	...	...	...
s7	29	$SX1 : -1log_2\frac{1}{29}$	$log_2 29$
Sum	**98**	**DL1($\tilde{C}$|"FPV")**	**DL2($\tilde{C}$|"FPV")**

cost $DL1(\tilde{C}|SC)$ of $\tilde{C}$ that is segmented by the pattern SC:

$$DL1(\tilde{C}|SC) = \sum_{i}^{m}\sum_{j} -l_{ij}log_2\frac{l_{ij}}{t_i} \quad (12)$$

We calculate the complementary parameter encoding cost of each segment as $log_2 t_i$.

For example, in Table 2, the parameter encoding cost of the first segment is $log_2 9$, the second segment is $log_2 3$ and so on. Thus, we calculate $DL2(\tilde{C}|SC)$ of $\tilde{C}$ as:

$$DL2(\tilde{C}|SC) = \sum_{i}^{m} log_2 t_i \quad (13)$$

Next, we calculate the segmentation cost $DL3(\tilde{C}|SC)$ as:

$$DL3(\tilde{C}|SC) = mlog_2(\sum_{i}^{m} t_i) \quad (14)$$

Again, for example, in Table 2, the length of $\tilde{C}$ is 98, so the segmentation cost is $7log_2 98$.

Finally, based on this table, we obtain the description length of $\tilde{C}$ that is segmented by the pattern SC as follows:

$$MDL(\tilde{C}|SC) = DL1(\tilde{C}|SC) + DL2(\tilde{C}|SC) + DL3(\tilde{C}|SC) \quad (15)$$

We use Eq. (15) as the MDL estimation function for the MDL pattern detection algorithm.

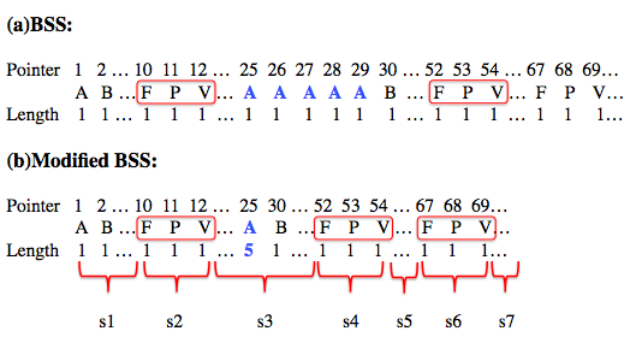

Figure 3: (a) BS sequence obtained from the time-series data. (b) Modified BS sequence.

3. EXPERIMENTAL EVALUATION

In order to evaluate our technique, we aim to identify motifs in the 'All I have Seen" dataset [16]. This data was generated by a lifelogger wearing a SenseCam camera that continuously captured an image every 20 seconds (on average), depicting activities such as walking, working, cooking, eating, driving, shopping, etc. In order to investigate the wearer's 'typical day' lifestyle, we studied data recorded on four Wednesdays. Selecting data recorded at the middle of the week was deliberate, given the likelihood of a more regular routine. The selected subset consists of 7549 images. Descriptive statistics are reported in Table 3.

First, the MODWT of the pixels for each image was calculated within each window of given size 400 images and the correlation matrix between pixels at each scale found. The Eigenvalues of the correlation matrix in each window were determined, and the Eigenvalue time series were normalised in time. Then, the largest Eigenvalue for different window sizes was analysed. Finally, the SAX algorithm was applied to transform the time series to PAA symbols. The results are illustrated in Fig. 4. The wavelet scales 1-4 correspond to a 1-2 minute period, a 2-4 minute period, a 4-8 minute period and a 8-16 minute period, respectively. The different features, found at various scales, suggest that the correlation matrix captured different major events with different time horizons. The largest Eigenvalue dynamics show that with increased wavelet scales comes increased smoothing, as expected. This removes some of the high frequency small-scales changes, typically associated with noise.

Table 3: The Data Set

User	Date	Week Day	Images
1	22-04-2009	Wednesday	1913
1	29-04-2009	Wednesday	1921
1	06-05-2009	Wednesday	2095
1	13-05-2009	Wednesday	1620
			Total: 7549

Fig. 5 depicts examples of motifs discovered from different wavelet scales using the method we introduced in Section 2. As depicted, different motifs were extracted from different wavelet scales. Since different wavelet scales represent different time horizons, the motifs, extracted from different scales, represent different 'events' that the wearer experienced each Wednesday.

Examining the images that are identified by the motifs analysis, we find that both a_1 and a_2 describe the combined event of driving back home in the afternoon, followed by watching TV. b_1 combines activities of eating, and then moving to the living room. b_2 corresponds to driving to the shopping mall. Both c_1 and c_2 are similar events where the wearer drives to work and then sits or some time in front of the computer. For d_1 the event sequence comprises sitting and watching TV, talking with family and then starting to cook, while d_2 comprises sitting in the shopping mall and then driving to an outdoor garden with children.

By examining the data set, we determine that most motifs discovered at the same wavelet scale represent similar scenarios. We note also that most describe the change from a static setting (e.g., sitting in front of a computer or TV) to a more dynamic activity. Light change reminds an important feature in identify/distinguishing between *key episodes* detected by our technique, in agreement with

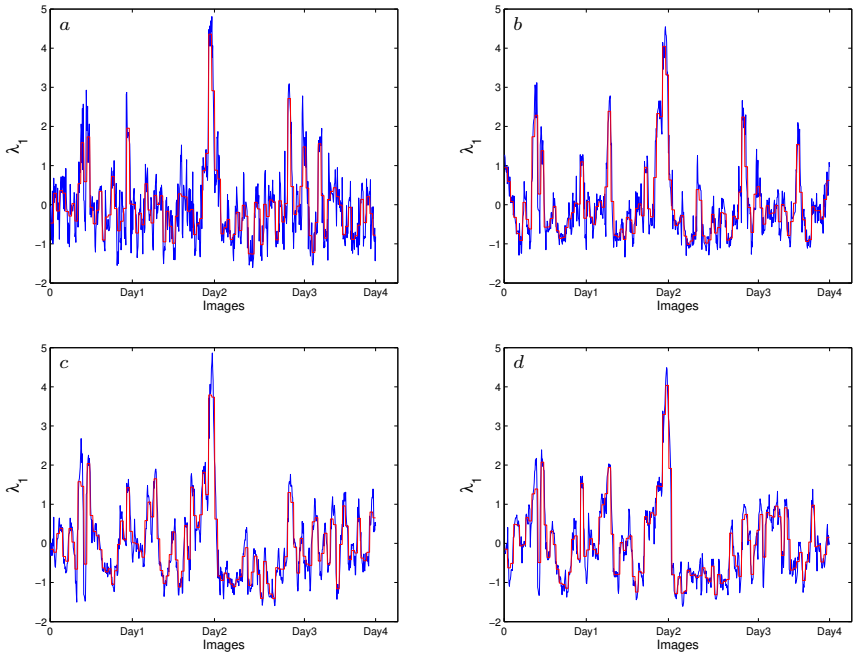

Figure 4: The largest Eigenvalue λ_1 (*blue*) dynamics together with the curve generated for the PAA (*red*). (a) Wavelet scale 1, (b) Wavelet scale 2, (c) Wavelet scale 3 and (d) Wavelet scale 4.

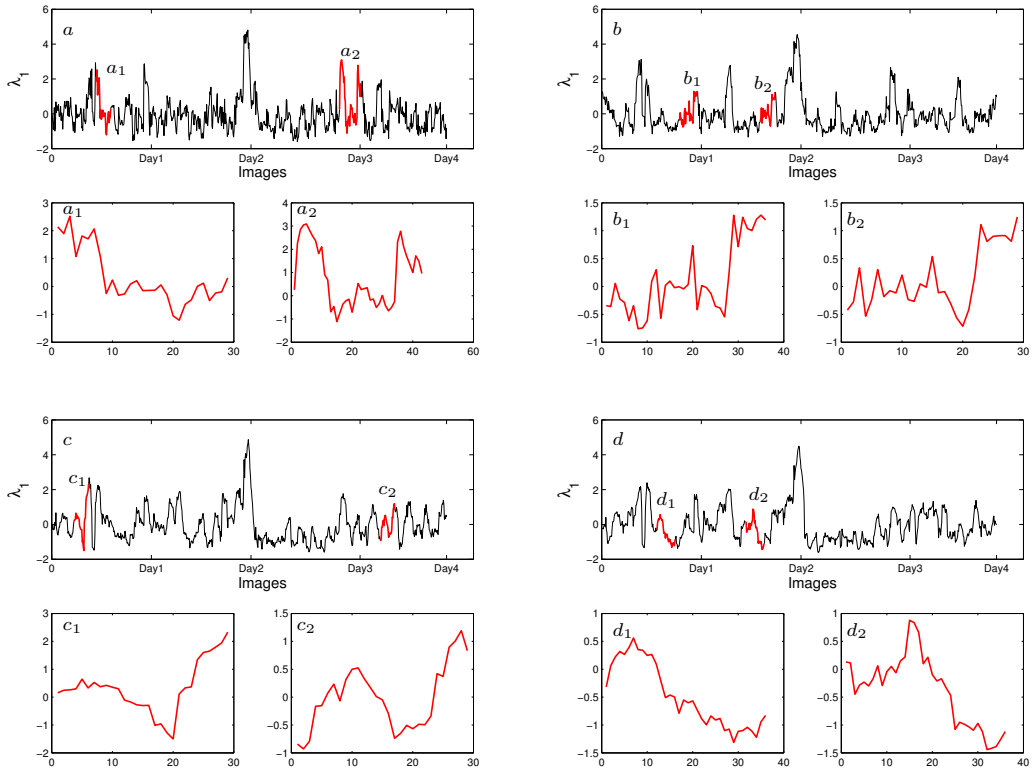

Figure 5: Example of Motifs extracted based on MDL principle from different wavelet scales. (a) Wavelet scale 1, (b) Wavelet scale 2, (c) Wavelet scale 3 and (d) Wavelet scale 4.

the prior study [21]. Added here is a way of aggregating related or sequential events in to a recognisable template, which can be used as potential marker for automatic extraction or comparison of similar episodes from an extended time series. Use of other information criteria with wavelet deconvolution may also be of value in future analysis.

4. CONCLUSIONS

In this paper, we introduced a non-supervised method based on motifs to identify events in visual lifelogs. The major contributions of this paper include demonstrating that motifs provide a good way to organise, structure and interpret vast amount of heterogeneous streams of visual data. Further, we show that the Minimum Description Length (MDL) can be used to extract such motifs from multi-dimensional time series Sensecam data. In particular, this method does not require any prior knowledge of the data. The motifs discovered provide prototype templates for identification of similar scenarios at specific time scales e.g. 'typical' lifestyle patterns of lifeloggers. Finally, results reported here provide supporting evidence in confirmation of the previous work, which indicated that light changes are natural markers for distinguishing key episodes. As future work, we intend to compare the overall performance of the method with existing supervised event recognition methods.

5. REFERENCES

[1] I. P. Androulakis. Selecting Maximally Informative Genes. *Computers & Chemical Engineering*, 29(3):535–546, 2005.

[2] D. Ashbrook, K. Lyons, and J. Clawson. Capturing Experiences Anytime, Anywhere. *IEEE Pervasive Computing*, 5(2):8–11, 2006.

[3] M. Blum, A. Pentland, and G. Tröster. InSense: Interest-Based Life Logging. *IEEE MultiMedia*, 13(4):40–48, 2006.

[4] W. Breymann. Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management (2nd ed.). Jean-Philippe Bouchaud and Marc Potters. *Journal of the American Statistical Association*, 101:850–852, 2006.

[5] Z. Can, Z. Aslan, O. Oguz, and A. H. Siddiqi. Wavelet transforms of meteorological parameters and gravity waves. *Annales Geophysicae*, 23(3):659–663, 2005.

[6] P. Carpena, P. Bernaola-Galván, A. V. Coronado, M. Hackenberg, and J. L. Oliver. Identifying characteristic scales in the human genome. *Physical Review E*, 75:032903, 2007.

[7] A. R. Doherty, C. J. A. Moulin, and A. F. Smeaton. Automatically assisting human memory: A SenseCam browser. *Memory*, 19(7):785–795, 2011.

[8] A. R. Doherty and A. F. Smeaton. Automatically segmenting lifelog data into events. In *WIAMIS 2008 - 9th International Workshop on Image Analysis for Multimedia Interactive Services*, 2008.

[9] A. Dumitru and D. Smith. Eigenvalue repulsion in an effective theory of SU(2) Wilson lines in three dimensions. *Physical Review D*, 77:094022, May 2008.

[10] J. Gemmell, G. Bell, and R. Lueder. MyLifeBits: a personal database for everything. *Commun. ACM*, 49(1):88–95, 2006.

[11] P. Gopikrishnan, B. Rosenow, V. Plerou, and H. E. Stanley. Identifying Business Sectors from Stock Price Fluctuations. *Physical Review E*, 64:35106, 2001.

[12] R. Hamid, S. Maddi, A. Johnson, A. Bobick, I. Essa, and C. Isbell. Unsupervised activity discovery and characterization from event-streams. In *In Proceedings of the 21st Conference on Uncertainty in Artificial Intelligence (UAI05*, 2005.

[13] S. Hodges, L. Williams, E. Berry, S. Izadi, J. Srinivasan, A. Butler, G. Smyth, N. Kapur, and K. Wood. Sensecam: A retrospective memory aid. In *Proceedings of the 8th International Conference of Ubiquitous Computing (UbiComp 2006)*, pages 177–193. Springer Verlag, September 2006.

[14] T. Hori and K. Aizawa. Context-based video retrieval system for the life-log applications. In *Proceedings of the 5th ACM SIGMM International Workshop on Multimedia Information Retrieval*, MIR '03, pages 31–38, New York, NY, USA, 2003. ACM.

[15] I. M. Jánosi and R. Müller. Empirical mode decomposition and correlation properties of long daily ozone records. *Physical Review E*, 71:056126, 2005.

[16] N. Jojic, A. Perina, and V. Murino. Structural epitome: a way to summarize oneâĂŹs visual experience. In J. Lafferty, C. K. I. Williams, J. Shawe-Taylor, R. Zemel, and A. Culotta, editors, *Advances in Neural Information Processing Systems 23*, pages 1027–1035. 2010.

[17] E. Keogh, K. Chakrabarti, M. Pazzani, and S. Mehrotra. Dimensionality reduction for fast similarity search in large time series databases. *Journal of Knowledge and Information Systems*, 3:263–286, 2000.

[18] L. Laloux, P. Cizeau, J.-P. Bouchaud, and M. Potters. Noise Dressing of Financial Correlation Matrices. *Physical Review Letters*, 83:1467–1470, 1999.

[19] N. Li, M. Crane, H. Ruskin, and C. Gurrin. Application of statistical physics for the identification of important events in visual lifelogs. In *Bioinformatics and Biomedicine (BIBM), 2013 IEEE International Conference on*, pages 589–592, Dec 2013.

[20] N. Li, M. Crane, H. Ruskin, and C. Gurrin. Multiscaled cross-correlation dynamics on sensecam lifelogged images. In S. Li, A. El Saddik, M. Wang, T. Mei, N. Sebe, S. Yan, R. Hong, and C. Gurrin, editors, *Advances in Multimedia Modeling*, volume 7732 of *Lecture Notes in Computer Science*, pages 490–501. Springer Berlin Heidelberg, 2013.

[21] N. Li, M. Crane, and H. J. Ruskin. Automatically Detecting "Significant Event" on SenseCam. *International Journal of Wavelets, Multiresolution and Information Processing*, 11(06):1350050, 2013.

[22] J. Lin, E. Keogh, S. Lonardi, and P. Patel. Finding Motifs in Time Series. In *In the 2nd Workshop on Temporal Data Mining, at the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, pages 53–68, 2002.

[23] J. Lin, E. Keogh, L. Wei, and S. Lonardi. Experiencing sax: A novel symbolic representation of time series. *Data Min. Knowl. Discov.*, 15(2):107–144, Oct. 2007.

[24] W.-H. Lin and A. Hauptmann. Structuring Continuous Video Recordings of Everyday Life Using Time-Constrained Clustering. *Multimedia Content Analysis, Management, and Retieval SPIE-IST Electronic Imaging*, 6073:111–119, 2006.

[25] S. Mallat. *A Wavelet Tour of Signal Processing, Third Edition: The Sparse Way*. Academic Press, 3 edition, 2008.

[26] S. Mann, J. Fung, A. S. Interaction, D. Chen, and H. Media. Designing EyeTap digital eyeglasses for continuous lifelong

capture and sharing of personal experiences. In *In Proceeding of CHI'15, ACM Press*, 2005.

[27] R. Mégret, V. Dovgalecs, H. Wannous, S. Karaman, J. Benois-Pineau, E. El Khoury, J. Pinquier, P. Joly, R. André-Obrecht, Y. Gaëstel, and J.-F. Dartigues. The IMMED Project: Wearable Video Monitoring of People with Age Dementia. In *MM'10*, pages 1299–1302, 2010.

[28] D. Minnen, T. Starner, I. Essa, and C. Isbell. Discovering Characteristic Actions from On-Body Sensor Data. In *Wearable Computers, 2006 10th IEEE International Symposium on*, pages 11–18, 2006.

[29] D. Minnen, T. Starner, I. A. Essa, and C. L. I. Jr. Improving activity discovery with automatic neighborhood estimation. In *IJCAI 2007, Proceedings of the 20th International Joint Conference on Artificial Intelligence, Hyderabad, India, January 6-12, 2007*, pages 2814–2819, 2007.

[30] C. G. Nevill-Manning, T. D. Wu, and D. L. Brutlag. Highly specific protein sequence motifs for genome analysis. *Proc. of the National Academy of Sciences*, 95(11):5865–5871, 1998.

[31] V. Plerou, P. Gopikrishnan, B. Rosenow, L. Amaral, and H. Stanley. A random matrix theory approach to financial cross-correlations . *Physica A: Statistical Mechanics and its Applications*, 287(3-4):374 – 382, 2000.

[32] V. Plerou, P. Gopikrishnan, B. Rosenow, L. A. Nunes Amaral, and H. E. Stanley. Universal and Nonuniversal Properties of Cross Correlations in Financial Time Series. *Physical Review Letters*, 83:1471–1474, 1999.

[33] S. Reddy, A. Parker, J. Hyman, J. Burke, D. Estrin, and M. Hansen. Image Browsing, Processing, and Clustering for Participatory Sensing: Lessons from a DietSense Prototype. In *Proceedings of the 4th Workshop on Embedded Networked Sensors*, EmNets '07, pages 13–17, New York, NY, USA, 2007. ACM.

[34] M. S. Santhanam, J. N. Bandyopadhyay, and D. Angom. Quantum spectrum as a time series: Fluctuation measures. *Physical Review E*, 73:015201, 2006.

[35] A. J. Sellen, A. Fogg, M. Aitken, S. Hodges, C. Rother, and K. Wood. Do life-logging technologies support memory for the past?: An experimental study using sensecam. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '07, pages 81–90, New York, NY, USA, 2007. ACM.

[36] S. Sharifi, M. Crane, A. Shamaie, and H. Ruskin. Random matrix theory for portfolio optimization: a stability approach. *Physica A: Statistical Mechanics and its Applications*, 335(3-4):629 – 643, 2004.

[37] Z. Siwy, M. Ausloos, and K. Ivanova. Correlation studies of open and closed state fluctuations in an ion channel: Analysis of ion current through a large-conductance locust potassium channel. *Physical Review E*, 65:031907, 2002.

[38] Y. Tanaka, K. Iwamoto, and K. Uehara. Discovery of time-series motif from multi-dimensional data based on MDL principle. *Machine Learning*, 58(2-3):269–300, 2005.

[39] A. Utsugi, K. Ino, and M. Oshikawa. Random Matrix Theory Analysis of Cross Correlations in Financial Markets. *Physical Review E*, 70:026110, 2004.

[40] P. Wang and A. F. Smeaton. Using visual lifelogs to automatically characterize everyday activities. *Information Sciences*, 230(0):147 – 161, 2013.

[41] P. Wang, A. F. Smeaton, and C. Gurrin. Factorizing time-aware multi-way tensors for enhancing semantic wearable sensing. In *MMM'15*, pages 571–582, 2015.

[42] D. Wilcox and T. Gebbie. On the analysis of cross-correlations in South African market data. *Physica A: Statistical Mechanics and its Applications*, 344(1-2):294 – 298, 2004. Applications of Physics in Financial Analysis 4 (APFA4).

[43] Z. Xizhi. The Application of Wavelet Transform in Digital Image Processing. In *Proceedings of the 2008 International Conference on MultiMedia and Information Technology*, MMIT '08, pages 326–329, Washington, DC, USA, 2008. IEEE Computer Society.