Computer Methods and Programs in Biomedicine 160 (2018) 65–74

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Determining Anxiety in Obsessive Compulsive Disorder through Behavioural Clustering and Variations in Repetition Intensity

Conor Cameron[a,*], Ibrahim Khalil[a], David Castle[b]

[a] School of Computer Science and Information Technology, RMIT University, Melbourne, Victoria, Australia
[b] Department of Psychiatry, The University of Melbourne, Australia

ARTICLE INFO

Article history:
Received 8 July 2017
Revised 18 February 2018
Accepted 20 March 2018

Keywords:
Compulsive behaviour
Obsessive compulsive disorder
SURF
Behaviour classification
Anxiety analysis

ABSTRACT

Background and objectives: Over the last decade, the application of computer vision techniques to the analysis of behavioural patterns has seen a considerable increase in research interest. One such interesting and recent application is the visual behavioural analysis of mental disorders. Despite the very recent surge in interest in this area, relatively little has been done thus far to assist individuals living with Obsessive Compulsive Disorder. The work proposed herein represents a proof of concept system designed to demonstrate the efficacy of such an approach, from the computational perspective. The specific focus of this work lies in demonstrating a mechanism for clustering different kinds of Obsessive Compulsive Disorder behaviours and then comparing new behaviours to existing behaviours to determine the approximate level of anxiety represented by a compulsive behaviour.
Methods: The proposed system uses Temporal Motion Heat Maps, SURF descriptors, a visual bag of words model and SVM-based classification to categorise representations of various behaviours commonly seen in OCD. Moreover, we apply a set of statistical measures to the images in a given category in order to derive an approximate anxiety level for a given compulsive behaviour. This proof of concept is an essential step in the direction of integrating computational approaches into the treatment of psychiatric conditions such as Obsessive Compulsive Disorder, for more effective recovery.
Results: Results gleaned from experimental simulations indicate that the proposed system is capable of correctly classifying different types of simulated Obsessive Compulsive Disorder behaviour classes 75% of the time, with the misclassifications almost exclusively occurring when two behavioural clusters appear highly similar. Based on this information the proposed system is then able to assign an approximate behavioural anxiety level to the compulsive behaviours that meets the approval of a mental health professional.
Conclusions: The proposed system demonstrates a good ability to categorise various types of simulated OCD behaviour, in addition to establishing an approximate anxiety level for a given compulsive behaviour. This research demonstrates strong potential for the use of such systems in assisting mental health professionals looking to better understand their patients' condition and treatment progress across time.

1. Introduction

Computer-vision-based behaviour analysis and classification is a topic that has been of interest to researchers for some time now [1–8]. Much of this interest stems from the wide variety of applications available to researchers studying the topic, from crowd analysis [9], to security-based surveillance and anomaly detection [10,11], to gait recognition [12,13], assisted living [14,15], emotion recognition [16] and more recently, mental health applications [17,18]. Arguably one of the most valuable applications of late, computer-vision-based mental health analysis has relied upon increases in computing power and algorithmic sophistication that have only relatively recently allowed researchers to discern the subtle behavioural cues that can be observed in everything from stress [19] to depression [20,21], dementia [22], bipolar disorder [23] and Autism Spectrum Disorder (ASD) [24]. Despite this keen interest, little has been done to exploit the naturally visible behavioural compulsions evident in Obsessive Compulsive Disorder (OCD). As these behavioural manifestations are related to the level of anxiety that an individual is experiencing at a given time we see strong potential in such research regarding its value to mental

* Corresponding author.
E-mail address: conor_cameron@student.rmit.edu.au (C. Cameron).

https://doi.org/10.1016/j.cmpb.2018.03.019
0169-2607/

health professionals in more deeply understanding their patients' needs [25].

Conventional approaches to treating and understanding patients with OCD have tended to focus on a combination of therapeutic interventions, such as Cognitive Behavioural Therapy (CBT), and self-report measures. While such approaches have been demonstrated to be quite successful, they nevertheless entail multiple drawbacks. Arguably chief among these drawbacks is the fact that patient feedback relies on human memory, which is not only highly subjective, but is also known to be highly volatile [26,27]. Conversely, the aim of the proposed system is to provide a proof-of-concept for analysing the behavioural compulsions typical of OCD in an objective manner with a view to the future application of such systems by mental health professionals to not only better understand their patients' individual compulsions, but also better understand the nature of the physical signs of the condition in greater depth. The proposed system aims to substantiate this approach by grouping different physical compulsions together based on their visual similarity and then ascribing an approximate anxiety level to each behaviour in turn, based on the intensity of its visual repetitiveness. Through this new method of visually analysing and understanding OCD compulsions and their related anxiety, we believe that the proposed system demonstrates the viability of computer vision based OCD monitoring to aid mental health professionals in improving their understanding and treatment of their individual OCD patients.

Traditional applications of visual behavioural analysis have focused on a wide-variety of behaviours, with many of these being common human behaviours for such applications as surveillance, workplace ergonomics, assisted living, detecting emotional states, and even driver vigilance [28–33]. However, only relatively recently has the analysis of abnormal behaviours characteristic of mental disorders become a topic of considerable research interest. Of this body of research, the primary focus has been on both the detection and analysis of behaviours characteristic of specific disorders. Much of this research has been driven by an attempt to understand the current progress and severity of the disorder, as well as to assist mental health professionals in early diagnosis or determining the most effective treatment strategies.

For example, MADRIM, proposed by Mugica et al. was designed to monitor patients with Major Depressive Disorder and analyse their progress during treatment [34]. In a similar vein, Joshi et al. proposed a system for the visual analysis of behavioural cues in individuals with depression, based on upper body movements and intra-facial muscle movements [35]. Additionally, Goodwin et al. used accelerometers and pattern recognition techniques to detect behaviours typical of individuals with ASD [36]. Amor et al. proposed a system for analysing individuals with Bipolar Disorder using Personalised Ambient Monitoring [23]. Finally, multiple computer-vision based approaches have been proposed to address the increasing burden of dementia, most notably Alzheimer's Disease [37,38]. Despite this notable interest in psychiatric interventions, no research that we have encountered has focused specifically on OCD and its associated anxiety-driven physical compulsions. Furthermore, no research has focused on using visual compulsion detection and analysis to produce anxiety ratings, which could be of great value to mental health professionals. These points are what the research presented herein is designed to address.

In total, the proposed system comprises the following contributions:

- **Demonstrated a method capable of classifying different patterns of simulated compulsive behaviour:** As part of the proposed system, we herein demonstrate a method that is capable of accurately clustering similar compulsive behaviours that are typical of OCD. This is achieved by deriving SURF descriptors from Temporal Motion Heat Maps (Temporal Motion Heat Maps) of compulsive behaviour from our prior research. The SURF descriptors are then used to classify various labelled instances of simulated compulsive behaviour using n-fold cross-validation. This results in a robust system that is scale and translation invariant and requires only a basic set of manually labelled behaviour instances to be trained.
- **Presented a useful method of establishing the relative anxiety of various compulsive behaviours:** We have provided a method that could, for example, be used by mental health professionals to quickly and easily view the relative anxiety one of their patient's compulsions. While the current system is a proof of concept, we see this research as demonstrating a strong potential for such future applications. We achieved this by producing a baseline average (mean) for each compulsive TMHM in a given behaviour cluster and then used this to compare each example of behaviour to the baseline for that cluster. This demonstrates the viability of such an approach to comparing compulsive behaviours, both across time and within a given time period, to determine whether any given example of compulsive behaviour is more or less anxiety-driven than usual for an individual.

The remainder of this paper is organised as follows: Section 2 presents the material that the reader will need to understand the content in subsequent sections. Section 3 provides an in-depth explanation of the proposed system and its components. Section 4 explains how the experiments were set up and executed as well as the parameters that were used. Section 5 presents the results from our experiments and discusses their implications for the utility of the proposed system. Section 6 provides a brief discussion of the pertinent literature and Section 7 recapitulates our findings, as well as the limitations and benefits of the proposed system, before briefly mentioning future research that we plan to undertake.

2. Background

This section provides a brief overview of the foundational material that is needed to understand the information in subsequent sections. It includes a Section 2.1 describing SURF descriptors and how they have been integrated with the proposed system in order to provide robust compulsive behaviour classification.

2.1. SURF descriptors

Speeded-up Robust Features (SURF) is a scale and rotation-invariant feature detector and descriptor, which is designed to be highly robust and efficient [39]. It achieves these goals via a simplified Hessian-based matrix measure. Thus, simply put, when an image is fed to SURF it will create a set of representative, orientation-invariant feature descriptors from the image, which can then be compared for similarity with other descriptors from other images. The efficiency and general orientation invariance of SURF make it highly useful when applied to certain problems in computer vision, including object recognition and tracking. The need for such an algorithm was to fill a void in the ability to detect individuals at different scales and slightly different positions during compulsive behaviours. For example, an individual may move closer to or further from the camera during the same compulsion performed at different times. By making the system scale invariant, it is better able to recognise similar behaviours regardless of scale, among other traits. This is why SURF was a valuable asset, as it is able to perform at a similar accuracy level to SIFT, but with more efficiency, thus making it apt for the proposed system. The formula for SURF is detailed below:

$$I\Sigma(x, y) = \sum_{i<x} i = 0 \sum_{i<y, j=0} I(i, j)$$

The aforementioned benefits of SURF indicated that it could be implemented in the proposed system to produce a set of robust feature descriptors from our dataset of TMHMs. These descriptors were then clustered into a vocabulary of visual words in a bag of words model and the result classified into different behaviour groups via an SVM. For the purposes of this research, MATLAB's implementations of SURF, bag of words model, and SVM were used [40].

3. Methodology

This section presents the methodology of the proposed system and is organised in the following manner: Section 3.1 gives a general picture of the proposed system so that the reader can understand how the components fit together. Section 3.2 explains how we used a combination of TMHMs, SURF descriptors and a bagging model in order to train a classifier to recognise different groups of labelled compulsive behaviours. Section 3.3 describes the metrics that we used to derive anxiety approximations from the classified groups of compulsive behaviour.

3.1. An overview

The proposed system is composed of two primary components:

- **A behaviour classification algorithm:** This is a MATLAB-based SURF descriptor combined with a bag of features model, used to represent the features. An SVM was then applied to classify the behaviours based on these features.
- **An anxiety correlation algorithm:** This is a set of statistical measures designed to compare new examples of compulsive behaviour to averages of the same behaviour exhibited by the same individual in the past. This approach expects some prior grouping of the behavioural examples to have been appropriately establish. In this case, this would be the achieved by the aforementioned classification algorithm.

The overall operation of the proposed system proceeds as follows:

To begin with, we took a set of 40 Temporal Motion Heat Maps from our previous research, which comprised eight examples of five different simulated compulsive behaviours. An example of these TMHMs can be seen in Fig. 1 We then labelled the 40 behaviours manually, and fed them into a MATLAB implementation of a bag of SURF features classification model. Due to the small size of our dataset, we modified the approach to use 4-fold cross validation. This resulted in each fold comprising ten behavioural examples, which could then be compared against the examples in every other fold. The behaviours were evenly distributed amongst the folds i.e. two examples of each behaviour per fold, in order to ensure an even balance of compulsive behaviour types being trained and tested.

As previously mentioned, in addition to training a classifier to recognise different compulsive behaviour groups, an anxiety algorithm was also used to ascribe an approximate anxiety level to each compulsive behaviour. This was achieved by comparing each compulsive behaviour to the average of the same type of compulsive behaviour in order to place it within an anxiety distribution. The standard deviation of a compulsive behaviour group could then be used to establish how far from average anxiety a given level of compulsivity in a behaviour is for an individual. In other words, any new example of a previously seen compulsive behaviour could be assessed automatically in order to determine how aberrant it is in terms of repetitiveness for that individual.

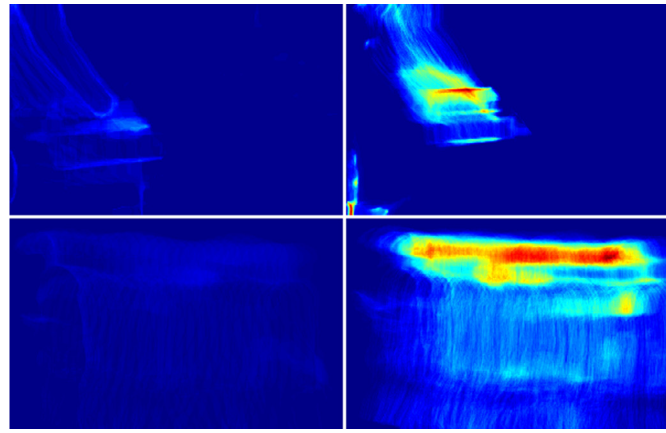

Fig. 1. Four TMHMs are depicted demonstrating two different behaviours. Moving clockwise from top-left: a 'cold', low-intensity, non-compulsive example of a drawer opening behaviour; a 'hot', high-intensity, compulsive example of a drawer opening behaviour; a 'hot', high-intensity, compulsive example of a back-and-forth walking behaviour; a 'cold', low-intensity, non-compulsive example of a walking behaviour.

Additionally, the proposed anxiety algorithm also takes into account the frequency of compulsive behaviours, in general, over a given time period. For the purposes of the proposed system, 24 hours was decided upon as an appropriate time frame within which to count behavioural frequency. However, any chosen value in hours can be substituted for the 24 hours in a day. Thus, not only is the intensity of each compulsive behaviour taken into account in comparison to its norms, but the number of compulsive behaviours of any type, exhibited over a period of 24 hours, was also included in the anxiety algorithm. Consequently, during times of high anxiety, an individual's increased compulsive behaviour frequency could be captured by the system as an objective indicator of the anxiety. Thus, while this system represents a proof-of-concept, future systems building upon information of this type could potentially be used by mental health professionals to better understand the condition and progress of their patients. An illustration of how the aforementioned system components fit together is presented in the flowchart in Fig. 2 Additionally, For the purpose of clarity, the aforementioned system components are also detailed in Algorithms 1 and 2.

3.2. Behavioural clustering

In order to derive an approximate level of anxiety from the compulsive behaviours of a given individual, a baseline first needed to be established for the same type of behaviour, or a very similar type of behaviour. This would then enable the proposed system to ascribe an average degree of compulsivity for that specific behaviour as exhibited by that individual. In order to achieve this, we first separated the test set of 40 TMHMs into their constituent five sets of behaviour. Note that half of all examples were of compulsive behaviour while the other half were of non-compulsive behaviour. Thus, for each of the five types of behaviour, four examples of said behaviour were compulsive and four examples of the same behaviour were non-compulsive.

Each example was then manually labelled with its behaviour type so that a k-means clustering algorithm could be used to create a bag of visual words based on the centroids of the set of TMHM-derived SURF descriptors as input data. We chose not to use the rotation-invariant principle of the SURF descriptors when they were being derived as it weakened the overall accuracy of the system, likely due to a greater potential for confusion of the points among the classes. A bag of 500 visual words were used for each

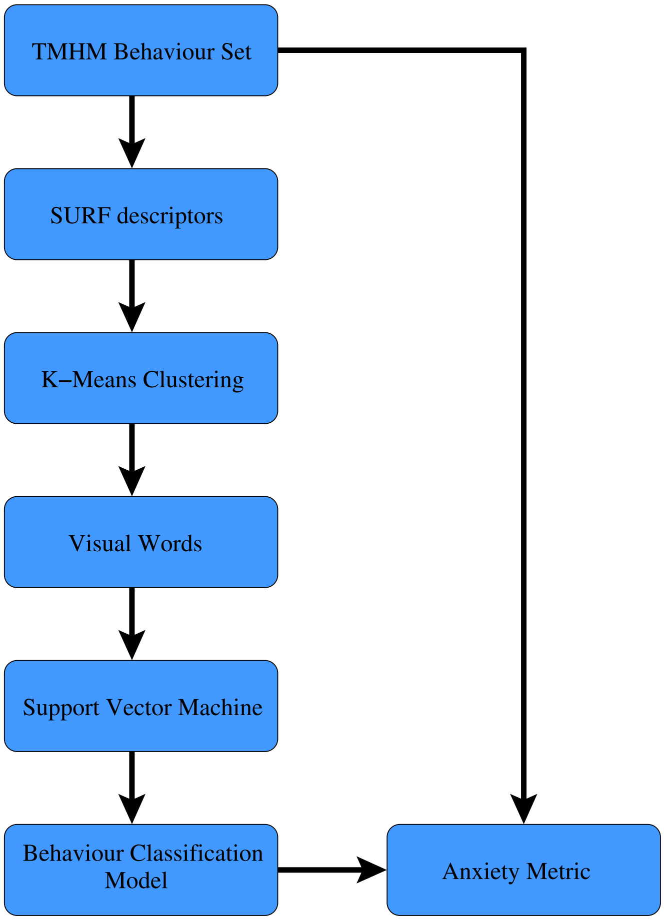

Fig. 2. A flowchart illustrating the system components and processes

behaviour instance, with 30 behaviour instances being used in total (10 were held out for testing). Once the bag of visual words had been computed based on the cluster centroids, an SVM with a linear kernel was then used to train and classify the instances. In order to optimise this process, we tuned the SVM cost parameter to a reasonable low value, of 0.1, in order to avoid overfitting while still retaining a good and reliable overall classification accuracy. In the process of training and testing the SVM classifier, the dataset was manually split into four folds, each comprising ten examples, with each fold having an even number of examples of any given behaviour. Once the folds were created, SURF descriptors were derived from each TMHM image. An example of this process can be seen in Fig. 3. Finally, four-fold cross validation was used when performing training and testing with the SVM classifier on the dataset.

Thus, the SVM model was trained to predict future examples of previously unseen types of behaviour, thus no longer requiring manual labelling, except in the case of new behaviours. Separating the behaviours in this way is of great value, as it allows each behaviour group to be viewed in isolation, which could then be used to gain a better understanding of the nature of such patterns of compulsive behaviour. Perhaps more importantly, as previously noted, separating compulsive behaviours into distinct groups allows direct comparisons to be made within behavioural groups, especially as new examples of that same behaviour come to light.

3.3. Anxiety metrics

Once the behaviours had been classified into distinct groups, an algorithm was then needed to ascribe anxiety levels to individual behaviour examples and behaviour groups. In considering this, two general directions present themselves, which will be referred to as *global* and *relative* anxiety metrics. A global anxiety metric can be thought of as providing an objective, universal measure of anxiety to any given behaviour in isolation, based on its overall level of repetitiveness. The main issue with this approach is that individuals will naturally differ in both how compulsive and dynamic their behaviours typically are under a given level of anxiety. Conversely, a relative anxiety metric would consider an example of a given behaviour, as manifested by a specific individual, in comparison to what is typical for that individual. For the purposes of this research, a relative anxiety metric was deemed more appropriate as it will not only provide a more fitting assessment for the individual, but could also provide a strong foundation for future research to capitalise on in informing a mental health professional of the change in an individual's compulsions across time.

```
// Get set of Temporal Motion Heat Maps
tmhm_set = getTMHMs();
numFolds = 4;

// Get SURF descriptors for each TMHM
tmhm_surf = getSURFDescs(tmhm_set);

// Divide set of TMHMs into four folds (ten examples per
// fold), whilst preserving a relatively equal number of
// compulsive behaviours per fold
for( i = 0; i < numFolds; i++)
   fold(i) = createBalancedFold(tmhm_surf,10);

for ( i = 0; i < numFolds; i++)
{
   // Get every fold besides this one
   trainData = getOtherFolds(i);
   testData = fold(i);

   resultTrain[i] = trainBagOfWords(trainData);
   resultTest[i] = testBagOfWords(testData);
}

// Derive an overall accuracy from the separate runs
resTest = mergeResults(resultTest);
```

Algorithm 1. Behaviour classification

```
// Get set of Temporal Motion Heat Maps
tmhm_set = getTMHMs();

for every tmhm in tmhm_set
    anxMet = mean(tmhm);

// Produce standard deviation from the current
// sample of anxiety points
anxStdDev = calcStdDev(anxMet);

// Compare new points with the existing distribution
// to see where they fall in terms of anxiety
for every tmhm in new_tmhm
{
    newAnxMet = mean(tmhm);
    anxVal = comparePoints(newAnxMet, anxStdDev);
}

// Combine anxiety values for all compulsive behaviours over
// 24 hours to produce a single anxiety rating
anxRating = sum(anxVals)/24;
```

Algorithm 2. The anxiety algorithm

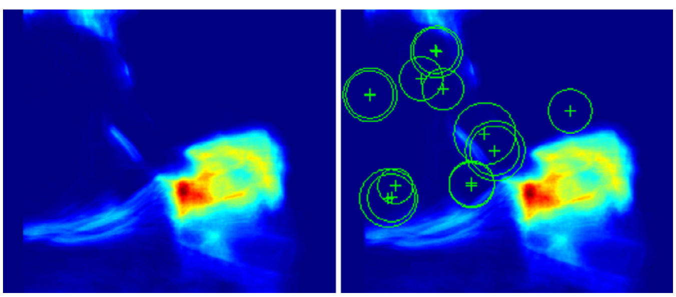

Fig. 3. An example Temporal Motion Heat Map for a reading behaviour (left), along with its set of strongest interest points (right)

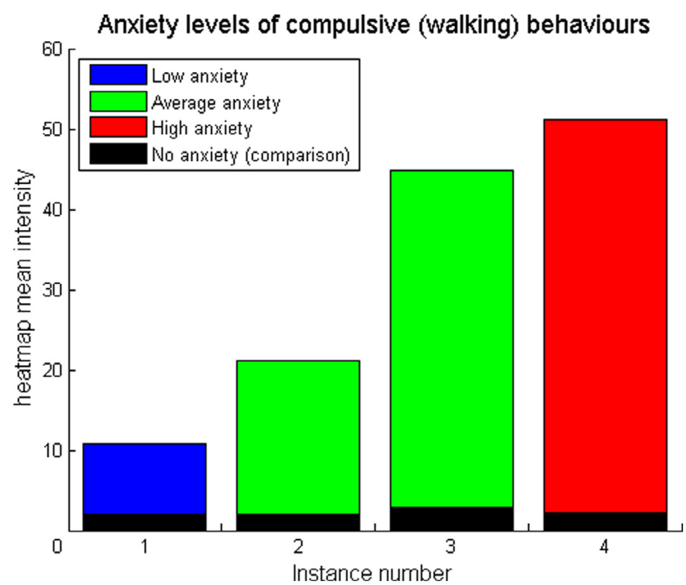

Fig. 4. An example of the relative anxiety metric as applied to the set of walking behaviours

Note that in order to isolate the results for the two components of the proposed system, the behavioural clustering and the anxiety metric, and thus provide a more accurate appraisal of each, the original labelled dataset of behaviours i.e. the pre-classified set of behaviours was used rather than the output of the behaviour classification system. To begin with, the mean of each labelled TMHM in a given behaviour group was calculated. A sample standard deviation was then calculated from the mean values of the compulsive examples of the given behaviour. The sample standard deviation was chosen because it took into account the fact that the given set of behaviours were only a small representative sample of how an individual's compulsions may appear overall. Once the standard deviation had been calculated for the distribution, each example could then be compared to the mean of all examples based on where it fell within the distribution. Naturally, all example data points falling within one standard deviation of the mean of all examples would be considered normal, all falling between one and two standard deviations above and below the mean would be considered indications of higher and lower than normal levels of anxiety, while everything greater than two standard deviations would be seen as considerably higher or lower anxiety than normal. An example of this approach can be seen in Fig. 4

To compare the proposed system to a manual approach, a mental health professional would need to view large quantities of recorded data of a patient and then assess their anxiety using their own statistical metrics. The proposed system instead is able to produce this metric from the grouped TMHMs automatically. This has the potential to allow mental health professionals to see aspects of their patients' compulsions in a new light. Note, however, that the current research is only concerned with producing a viable 'proof of concept' simulation of OCD compulsion analysis. Any direct use of the proposed system by mental health professionals in a clinical context is beyond the scope of the current work and is planned for future work. Finally, all compulsive behaviours are then considered as a whole and viewed in terms of frequency over a period of 24 h. In combining these two factors we produce the following formula which can be used to calculate the relative anxiety of an individual, based on their OCD compulsions, over a given time period:

$$\mathbf{AnxI} = \sum_{b=1}^{B} mean(b)\sigma \sum_{n=1}^{N} mean(Im)$$

Further value of this approach can be seen in the research which would build upon this proof-of-concept system. During the course of treatment, an individual may be experiencing more anxiety than can be detected via their compulsions as they instead try to deal with the anxiety without visible compulsion responses. This is to be expected. However, overall, as the individual is treated, not only should their compulsions fall, but so too should their associated anxiety. Concordantly, if a new spike in compulsive behaviour activity were to be detected during the patient treatment, it will mean that the patient has likely relapsed and would be again experiencing a high degree of anxiety. Without the basis of the research proposed herein, in terms of gathering a solid and objective measure of an individual's relative and global anxiety, a mental health professional would need to view large quantities of recorded data from a patient and then assess their anxiety using their own statistical metrics. The proposed system is instead able to produce such a metric from the grouped TMHMs automatically and such data could thus be used in future research to inform a mental health professional.

4. Experimental details

This section explains the how the proposed system was configured, as well as the parameters that were used during the experimental simulations. It includes the following subsections: Section 4.1 details the system configuration and the applications that were used to run the experiments, while Section 4.2 describes the video dataset that was used in the experiments. Finally, Section 4.3 details the performance metrics used in validating the proposed system.

4.1. System configuration

All of our experiments were run within the MATLAB technical computing suite, on a single Desktop computer with a 8-core processor running at 3.5GHz, with 8GB of RAM. The version of MATLAB used was 64-bit R2015a. The SURF algorithm and bag of words model were part of a set of MATLAB functions included in the Computer Vision System Toolbox.

4.2. Video data

The images used during all experiments were Temporal Motion Heat Maps obtained from a set of 40 initial videos. These 40 videos were distinct, simulated examples of OCD compulsions obtained from different camera angles, under various lighting conditions. Each TMHM, produced during our previous research, represents one video and thus one example of compulsive behaviour in a single image that preserves the degree of compulsivity. This representation was considered apt for the purposes of the proposed system and thus the 40 behavioural TMHMs were retained in developing the proposed system.

While the 40 examples of simulated compulsive behaviour used in this proof-of-concept system were performed by the same individual, who does not suffer from OCD, they were nevertheless based on a thorough reading of the OCD literature. Additionally, the compulsive behaviour simulations were further informed by discussions with a mental health professional and then verified by said mental health professional as representing plausible examples of OCD compulsions. These behaviours can be subdivided into five distinct classes, which are as follows:

- **TMHM behaviour set 1 - checking DVD**:
  This behaviour consisted of repeatedly picking up and replacing a DVD on a table. This type of behaviour would typically be considered a *checking* behaviour.
- **TMHM behaviour set 2 - opening and closing a drawer**:
  This behaviour consisted of repeatedly opening and closing a drawer. This is another form of checking behaviour.
- **TMHM behaviour set 3 - lock checking**:
  This behaviour consisted of repeatedly checking if a door was locked by turning the doorknob. This type of compulsion can be common among individuals with security fears, in which a lack of certainty can cause the behaviour to continue excessively.
- **TMHM behaviour set 4 - reading**:
  This behaviour entailed reading a passage from a book before replacing the book and picking it up again, repeatedly. This type of behaviour would typically be considered an example of repeating a routine behaviour.
- **TMHM behaviour set 5 - walking back and forth**:
  This behaviour involved walking back and forth repeatedly over the same location. This type of behaviour can occur when an individual gets stuck in a rumination and can't rid themselves of the negative thought. Behaviours such as these are good demonstrations of how OCD can end up consuming such a large amount of time out of an individual's day.

Table 1
Details of the video groups used in testing the proposed system

Compulsive behaviour	Mean	Sample standard deviation
Checking DVD	22.45	9.40
Opening and closing a drawer	7.84	4.57
Lock checking	14.64	9.72
Reading	34.03	22.59
Walking back and forth	30.84	19.08

Table 2
Details of the video groups used in testing the proposed system

Video group	Duration range (seconds)
Checking DVD	4 – 96
Opening and closing a drawer	2 – 39
Lock checking	4 – 77
Reading	12 – 126
Walking back and forth	4 – 102

Each of the above five behaviour classes contained four compulsive examples of the behaviour and four non-compulsive examples for comparison. Specific details of this video set are listed in Table 1 and Table 2.

4.3. Performance metrics

In determining the accuracy of the behavioural clustering aspect of the proposed system, we performed 4-fold cross validation on the model five times and averaged the result. This was done to avoid biasing the results, as the examples placed in the folds were chosen at random, though this was achieved whilst preserving relative behaviour distributions. From this, a confusion matrix was produced. Furthermore, the overall accuracy was determined by tallying the total number of correct results across all behaviours. The average precision and recall are also noted as additional indicators of the accuracy of the proposed system.

Regarding the performance of the anxiety metrics, naturally an evaluation of the proposed method of ascribing anxiety can only be made by a mental health professional based on its utility to their profession. This is because this research exists at the nexus of computer science and psychiatric evaluation.

5. Experimental results and analysis

Across five runs, shuffling the instances contained in the folds, an average accuracy of 82% was obtained across all instances. The accuracy for each of the five runs can be seen in Table 3. Across

Table 3
Average accuracy for each experimental run

Run number	Accuracy (%)
1	85
2	80
3	85
4	82.5
5	77.5
Average:	82%

Table 4
Confusion Matrix of experimental results for classifying the five behaviours

Classified as (%) ⇒	Behaviour 1	Behaviour 2	Behaviour 3	Behaviour 4	Behaviour 5
Behaviour 1	**85**	10	2.5	2.5	0
Behaviour 2	2.5	**90**	0	0	7.5
Behaviour 3	0	0	**37.5**	0	62.5
Behaviour 4	0	0	0	**97.5**	2.5
Behaviour 5	0	0	0	0	**100**

Table 5
Overall results

Metric	Result
True positive rate (Sensitivity)	82%
True negative rate (Specificity)	95.5%
Positive predictive value	87.27%
Negative predictive value	95.93%

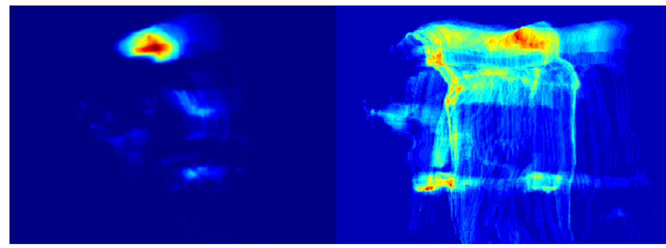

Fig. 6. An example of the variation between two examples of the same behaviour (lock checking). Despite being the same type of behaviour, these two behaviours appear visually to be quite different, giving this type of behaviour a high-class intra-class variation.

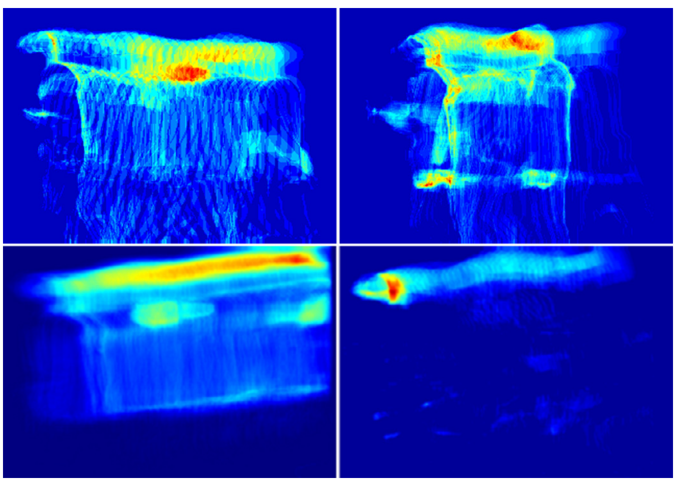

Fig. 5. Two examples of walking behaviour (left) compared to two examples of lock checking behaviour (right). It can been seen that despite the different behaviours, their appearance in this case can be quite similar. This has to do with the movement towards the door in the lock checking behaviour. The lack of bodily definition in some pictures is a result of the way different lighting interacts with certain colours of low contrast clothing. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

the five runs, the accuracy hit a low of 77.5% and reached a high of 85%. These results indicate a decent accuracy on the tested data, especially when we take into account that most of the inaccuracies are the result of strong visual similarities between the chosen compulsive behaviours. If we tease apart from a different perspective, we get the results detailed in Table 4 and Table 5. This confusion matrix considers the accuracy from the perspective of each of the five behaviour classes, in which the rows indicate the actual class of the behaviour and the columns indicate what the behaviour was classified as.

When we consider the deficiencies in overall accuracy from this perspective we can see that they occurred primarily across one specific behaviour class, namely Behaviour Group 3. This behaviour group involved a compulsive door lock checking behaviour, which was notably confused with a behaviour group involving walking back and forth. On closer inspection, the reasons for this become clear, the Lock Checking and Walking behaviours both contain a brief walking movement back and forth along the same axial plane. Both behaviours also look, in many ways, similar when one inspects their TMHMs, as can be seen in Fig. 5. Thus, while these two behaviour examples don't look completely identical, their distinct visual similarities mean that, for certain such behavioural examples, there is likely almost as much intraclass variation as there is interclass variation vis-a-vis these two specific classes. This is further exemplified by Fig. 6 in which a notable difference can be seen between this same behaviour type at a low repetition rate and a higher repetition rate. Specifically, most of the cases of misclassifications occurred when the very low-rep, non-compulsive behaviours were misclassified as walking behaviours. Again, the proposed system is most likely classifying these separate behaviours as the same behaviour because they appear visually quite similar to each other. Remedying this primarily low-repetition issue would likely involve a more lax visual definition of the behaviour and a relaxation of the classification model, but it remains an open question for future research. Regardless, it is not a significant issue, as it only occurs in highly visually similar behaviour classes.

Regardless, the other four behaviour groups performed very well, with behaviour groups 1, 2, 4, and 5, being classified with respective accuracies of 85%, 90% 97.5% and 100%. These behaviours, DVD checking, Drawer Opening, Reading, and Walking back and forth, respectively, tended to involve behavioural examples that were very similar within each class i.e. low intraclass variation, thus demonstrating a more distinct class cohesion. These behaviours were also notably less likely to be confused with other behaviours. It is interesting, however, that certain of these behaviours experienced asymmetrical class confusion, that is, behaviour group 3 was confused with behaviour group 5, but not vice-versa. This is likely the result of the aforementioned higher class cohesion within class 5, meaning that many behaviour examples were probably more similar to each other than to behaviour examples in class 3. Conversely, the same was probably not true for Behaviour group 3, where its results instead indicate that many of its examples were considered to be somewhat more similar to behaviour group 5 than their own class.

As we are primarily concerned with using grouped behaviours to produce a baseline level of anxiety, it is not critical that the grouped behaviours are identical, just similar. Furthermore, not all examples of a given behaviour are going to be visually identical, so some leeway and a somewhat wide distribution of pixel intensities across a behavioural group is to be expected. Finally, if the proposed system were to be implemented in general practice, a confusion of two visually similar behaviours would likely not pose any issue to a mental health professional viewing the anxiety results of the various classes of compulsive behaviours. Regardless, ideally the proposed system would still be able separate most such behaviours and thus we regard the overall accuracy of 82% to indicate quite good accuracy given the circumstances, but with the noted drawbacks.

6. Related work

Ouanane et al. demonstrated the use of SURF descriptors, combined with PCA and an SVM to classify aggressive and non-aggressive behaviours [41]. Aggressive and non-aggressive behaviours are compared by first extracting a set of interest points using known algorithms. SURF is used for feature description in order to make the system robust and efficient. Once the features have been matched across images of different types of behaviour, the dimensionality of the matches is then reduced using PCA. These reduced points are then used to represent aggressive or non-aggressive behaviour classes when an SVM classifier is subsequently applied. The research of Ouanane et al. shares some similarities with ours, such as the use of SURF descriptors and an SVM for modelling different behaviours. However, the input used to generate the SURF descriptors for our system is a set of Temporal Motion Heat Maps, as opposed to the use of raw images by Ouanane et al. Furthermore, we classify behavioural classes based on their behaviour type, which is a multi-class problem, as opposed to the binary problem of aggression vs non-aggression. Finally, this first component of our system is used to inform an anxiety rating in the second component, whereas the system proposed by Ouanane et al. is not concerned with anxiety.

Karaman et al. introduced a wearable video system for human activity indexing in patients with dementia [42]. In order to achieve this, the system uses motion based temporal segmentation, SURF descriptors for image localisation MPEG colour information for determining information about the patient's environment. These elements were then used in a Hierarchical Hidden Markov Model for activity representation. The research of Karaman et al. further exemplifies the recent interest in video-based interventions capable of providing objective information and analyses of the outcomes of patients with mental health issues. While this system uses a some similar techniques to ours, such as SURF descriptors for image description and localisation, it doesn't use any form of motion history image or heat map for this purpose, nor does it look towards any determination of anxiety of the individual. Conversely, our system applies SURF descriptors to our temporal motion heat maps for the purpose of analysing distinct actions for individual images. Our system also provides a method of ascribing an approximate anxiety to behaviours based on the repetition intensity of the behaviours exhibited.

Nievas et al. looked into testing the novel idea of violence detection in video clips [43]. They achieved this using a bag of words model combined with an SVM classifier, using various kernels, and compared STIP and MoSIFT classifiers. When testing their system on two video datasets, they achieved similar performance for all three descriptor-SVM kernel combinations on the first dataset, whereas for the second the MoSIFT + Histogram Intersection Kernel combination was the most accurate. While these results are slightly inconclusive, the authors argued that such results nevertheless demonstrate the viability of motion descriptors and the bag of words model for the purpose of violence detection in videos, due to the relatively high overall accuracy across both datasets. Although the approach of Nievas et al. uses some comparable elements to ours, such as a bag of words model and an SVM, we use SURF descriptors with our TMHMs, rather than MoSIFT or STIP, as SURF has demonstrated good performance across activity classification and is both scale and rotation invariant. Moreover, our system is designed to detect anxiety, whereas the system of Nievas et al. is designed to detect only violence and thus does not implement or compare the Temporal Motion Heat Map elements that our system is based upon.

Sivalingam et al. proposed an interesting system for the purpose of tracking children for early indications of mental disorders [44]. The benefit of such a system is that earlier diagnosis will allow mental health professionals to provide effective treatments sooner rather than later. The authors detail the benefits of automated systems for addressing such tasks and note that mental health professionals often don't have the time or access to perform such tasks. Like the proposed system, the system by Sivalingam et al. is designed to analyse behaviour in individuals with conditions such as OCD or ASD, however, their system is primarily designed for tracking children in a classroom setting, whereas ours is used for analysing the behaviour of, typically adult, individuals in their own home. Moreover, our system is designed to differentiate behaviour clusters characteristic of OCD and to produce associated anxiety levels. Whereas the system by Sivalingam et al. is designed primarily for tracking children across a room, handling occlusions and managing tracking hand-off between multiple different cameras.

Girard et al. proposed a system for the video-based analysis of different facial cues for indicators of depression [45]. They did this as patients were treated, over time, in clinical sessions with a mental health professional. As with our system, the system of Girard et al. looks for characteristics typical of an individual's mental health condition and uses these markers as indicators of times when the individual is experiencing a greater severity of symptoms. In the case of Girard et al., these were, fewer affiliative facial expressions and decreased head motion. In the case of the proposed system, this was the distinction between different compulsive sets of behaviour, as well the designation of an associated anxiety level. Thus, while both systems use video-based behaviour detection and analysis, the primary difference between this research and ours is that we focus on an anxiety disorder, namely OCD, and its compulsive characteristics, whereas Girard et al. focus on depression and its typical manifestations. Nevertheless, Girard et al.'s system takes into account more minutia in expression, notably facial expression, than does the proposed system, which indicates that it could also act, with some modifications as a compliment to our system. This would enhance the proposed system by involving less common facial cues in OCD in addition to the more distinct bodily cues.

Pavlidis et al. pioneered some early work on anxiety detection using thermal imagery in order to detect changes primary in average facial temperature [46]. This was used as a means to detect suspects engaged in illegal and potentially harmful activities. Pavlidis et al. achieved this by mapping various facial areas, such as the forehead, cheeks, nose, chin and neck, in order to calculate the average thermal activities within these areas across time. This thermal activity could then be viewed for changes which may indicate anxiety in the individual being monitored. While a very interesting and original approach, it nevertheless requires specialised thermal camera equipment and is designed, as such, to primarily detect changes in body temperature. Conversely, our system focuses specifically on repetitive behaviours as these are the hallmark of compulsions, and thus serve as an indicator of OCD. Moreover, not all anxiety experiences, even for an individual with OCD, are associated with compulsions. Regardless, if economically feasible, the use of systems such as that of Pavlidis et al. with some modifications applied may yet prove potentially useful for detecting some of the changes associated with anxiety, and may even prove useful for gauging some of the aspects of certain anxiety disorders.

7. Conclusion

In this paper, we have proposed a system that is capable not only of being trained to recognise and group simulated compulsive behaviours into specific clusters, but also of providing a relative anxiety metric which could be used to objectively approximate the severity of an individual's OCD compulsions at a given point in time. While the proposed system represents a proof-of-concept, it nevertheless demonstrates the potential such systems hold to aid mental health professionals in understanding their patients' progress across time, while they're being treated, though the exact implementation of this clinical aspect is beyond the scope of the proposed system. The aforementioned contributions were achieved on the one hand through a combination of Temporal Motion Heat Maps and SURF descriptor classification for training a behavioural classifier and, on the other, through a combination of metrics for approximating an anxiety level. As a result of the aforementioned contributions, we believe that the proposed system represents a new avenue of insight into how such systems could assist mental health professionals in understanding each unique patient's compulsions and anxiety.

The primary benefit of the proposed system is its ability to provide an objective approximation of the anxiety of compulsive behaviours across compulsion type and time. This demonstates the potential of such systems to offer mental health professionals access to a fine-grained understanding of their patient's particular compulsions and how those compulsions evolve across time. Thus, future research building upon this could prove highly useful in assessing how various treatment plans play out, as well as how effective various behavioural interventions, such as Exposure and Response Prevention Therapy, prove to be on the patient and even on their individual compulsions over time. However, the deployment of the proposed system in a specific form appropriate for a mental health professional as the end-user is still a subject for future research.

Despite the value of the proposed work, there are still some notable drawbacks. Namely, the proposed system requires manual labelling in order for the behaviour classification process to function. Naturally, this is only necessary in so far as it teaches the classifier new compulsions that an individual exhibits. However, as most individuals with OCD don't develop new compulsions very often and instead tend to maintain consistent compulsions, we don't consider this to be a serious issue. Furthermore, as the proposed system only seeks to create loose categories of compulsive behaviours, it is not essential that the compulsive behaviour be rigorously defined, so long as behavioural groups are similar enough that their TMHM signatures don't differ significantly and thus potentially throw off the anxiety baselines.

Future work will focus on providing a direct algorithmic approach to analysing and objectively documenting a patient's progress across time as a function of their treatment. We furthermore plan to extend this proof-of-concept system to provide a real clinical implementation, which could be used by mental health professionals in cooperation with their patients.

References

[1] M. Hu, Y. Wang, Z. Zhang, D. Zhang, J.J. Little, Incremental learning for video-based gait recognition with lbp flow, Cybern. IEEE Trans. 43 (1) (2013) 77–89.
[2] J. Kuo, S. Koppel, J.L. Charlton, C.M. Rudin-Brown, Computer vision and driver distraction: Developing a behaviour-flagging protocol for naturalistic driving data, Accident Anal. Prev. 72 (2014) 177–183.
[3] A.S. Won, J.N. Bailenson, J.H. Janssen, Automatic detection of nonverbal behavior predicts learning in dyadic interactions, Affective Comput. IEEE Trans. 5 (2) (2014) 112–125.
[4] A. Galata, N. Johnson, D. Hogg, Learning variable-length Markov models of behavior, Comput. Vis. Image Understand. 81 (3) (2001) 398–413.
[5] R. Mehran, A. Oyama, M. Shah, Abnormal crowd behavior detection using social force model, in: Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on, IEEE, 2009, pp. 935–942.
[6] D. Glowinski, N. Dael, A. Camurri, G. Volpe, M. Mortillaro, K. Scherer, Toward a minimal representation of affective gestures, Affective Comput. IEEE Trans. 2 (2) (2011) 106–118.
[7] R. Cucchiara, C. Grana, A. Prati, R. Vezzani, Probabilistic posture classification for human-behavior analysis, Syst. Man Cybern. Part A 35 (1) (2005) 42–54.
[8] A.B. Albu, B. Widsten, T. Wang, J. Lan, J. Mah, A computer vision-based system for real-time detection of sleep onset in fatigued drivers, in: Intelligent Vehicles Symposium, 2008 IEEE, IEEE, 2008, pp. 25–30.
[9] E.L. Andrade, S. Blunsden, R.B. Fisher, Modelling crowd scenes for event detection, in: Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, 1, IEEE, 2006, pp. 175–178.
[10] G.A. Jones, N. Paragios, C.S. Regazzoni, Video-based surveillance systems: computer vision and distributed processing, Springer Science & Business Media, 2012.
[11] N. Robertson, I. Reid, M. Brady, Behaviour recognition and explanation for video surveillance, in: Crime and Security, 2006. The Institution of Engineering and Technology Conference on, IET, 2006, pp. 458–463.
[12] Z. Liu, S. Sarkar, Improved gait recognition by gait dynamics normalization, IEEE Transactions on Pattern Analysis and Machine Intelligence 28 (6) (2006) 863–876.
[13] J. Preis, M. Kessel, M. Werner, C. Linnhoff-Popien, Gait recognition with kinect, in: 1st International Workshop on Kinect in Pervasive Computing, New Castle, UK, 2012, pp. P1–P4.
[14] C.N. Doukas, I. Maglogiannis, Emergency fall incidents detection in assisted living environments utilizing motion, sound, and visual perceptual components, IEEE Transactions on Information Technology in Biomedicine 15 (2) (2011) 277–289.
[15] S. Fleck, W. Straßer, Smart camera based monitoring system and its application to assisted living, Proceedings of the IEEE 96 (10) (2008) 1698–1714.
[16] A. Majumder, L. Behera, V.K. Subramanian, Emotion recognition from geometric facial features using self-organizing map, Pattern Recognition 47 (3) (2014) 1282–1293.
[17] J. Hoey, A. Von Bertoldi, P. Poupart, A. Mihailidis, Assisting persons with dementia during handwashing using a partially observable Markov decision process, in: Proc. Int. Conf. on Vision Systems, 65, 2007, p. 66.
[18] L. Piccardi, B. Noris, O. Barbey, A. Billard, G. Schiavone, F. Keller, C. von Hofsten, Wearcam: A head mounted wireless camera for monitoring gaze attention and for the diagnosis of developmental disorders in young children, in: RO-MAN 2007-The 16th IEEE International Symposium on Robot and Human Interactive Communication, IEEE, 2007, pp. 594–598.
[19] N. Sharma, T. Gedeon, Objective measures, sensors and computational techniques for stress recognition and classification: A survey, Computer methods and programs in Biomedicine 108 (3) (2012) 1287–1301.
[20] J. Joshi, R. Goecke, G. Parker, M. Breakspear, Can body expressions contribute to automatic depression analysis? in: Automatic Face and Gesture Recognition (FG), 2013 10th IEEE International Conference and Workshops on, IEEE, 2013, pp. 1–7.
[21] Y. Yang, C. Fairbairn, J.F. Cohn, Detecting depression severity from vocal prosody, Affective Computing, IEEE Transactions on 4 (2) (2013) 142–150.
[22] A.A. Chaaraoui, J.R. Padilla-López, F.J. Ferrández-Pastor, M. Nieto-Hidalgo, F. Flórez-Revuelta, A vision-based system for intelligent monitoring: Human behaviour analysis and privacy by context, Sensors 14 (5) (2014) 8895–8925.
[23] J.D. Amor, C.J. James, Behavioral pattern detection from personalized ambient monitoring, in: Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, IEEE, 2010, pp. 5193–5196.
[24] D. Feil-Seifer, M. Mataric, Automated detection and classification of positive vs. negative robot interactions with children with autism using distance-based features, in: Proceedings of the 6th international conference on Human-robot interaction, ACM, 2011, pp. 323–330.
[25] National Institute of Mental Health, Obsessive Compulsive Disorder Among Adults, 2016, (https://www.nimh.nih.gov/health/statistics/prevalence/obsessive-compulsive-disorder-among-adults.shtml).
[26] D. Offer, M. Kaiz, K.I. Howard, E.S. Bennett, The altering of reported experiences, Journal of the American Academy of Child & Adolescent Psychiatry 39 (6) (2000) 735–742.
[27] A.A. Stone, S. Shiffman, J.E. Schwartz, J.E. Broderick, M.R. Hufford, Patient non–compliance with paper diaries, Bmj 324 (7347) (2002) 1193–1194.
[28] T. Xiang, S. Gong, Video behavior profiling for anomaly detection, IEEE Trans. Pattern Anal. Mach. Intell. 30 (5) (2008) 893–908.
[29] S. Han, S. Lee, F. Peña-Mora, Vision-based motion detection for safety behavior analysis in construction, in: 2012 Construction Research Congress (CRC), 2012, pp. 21–23.
[30] S. Han, S. Lee, A vision-based motion capture and recognition framework for behavior-based safety management, Autom. Constr. 35 (2013) 131–141.
[31] A.N. Belbachir, A. Nowakowska, S. Schraml, G. Wiesmann, R. Sablatnig, Event–driven feature analysis in a 4d spatiotemporal representation for ambient assisted living, in: Computer Vision Workshops (ICCV Workshops), 2011 IEEE International Conference on, IEEE, 2011, pp. 1570–1577.
[32] J.A. Levine, I.T. Pavlidis, L. MacBride, Z. Zhu, P. Tsiamyrtzis, Description and clinical studies of a device for the instantaneous detection of office-place stress, Work 34 (3) (2009) 359–364.
[33] Q. Ji, Z. Zhu, P. Lan, Real-time nonintrusive monitoring and prediction of driver fatigue, IEEE Trans. Veh. Technol. 53 (4) (2004) 1052–1068.

[34] F. Mugica, A. Nebot, S. Bagherpour, A. Serrano-Blanco, L. Baladón, Madrim: A major depression remote intelligent monitor, in: Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on, IEEE, 2012, pp. 1–8.
[35] J. Joshi, A. Dhall, R. Goecke, M. Breakspear, G. Parker, Neural-net classification for spatio-temporal descriptor based depression analysis, in: Pattern Recognition (ICPR), 2012 21st International Conference on, IEEE, 2012, pp. 2634–2638.
[36] M.S. Goodwin, S.S. Intille, F. Albinali, W.F. Velicer, Automated detection of stereotypical motor movements, J. Autism Dev. Disorders 41 (6) (2011) 770–782.
[37] Z. Uhríková, C.D. Nugent, D. Craig, V. Hlaváč, Detection of aberrant behaviour in home environments from video sequence, Ann. Telecommun. 65 (9-10) (2010) 571–581.
[38] A. König, C.F. Crispim Junior, A. Derreumaux, G. Bensadoun, P.-D. Petit, F. Bremond, R. David, F. Verhey, P. Aalten, P. Robert, Validation of an automatic video monitoring system for the detection of instrumental activities of daily living in dementia patients, J. Alzheimer's Dis. 44 (2) (2015) 675–685.
[39] H. Bay, A. Ess, T. Tuytelaars, L. Van Gool, Speeded-up robust features (surf), Comput. Vis. Image Understand. 110 (3) (2008) 346–359.
[40] MATLAB - Detect SURF Descriptors, detectSURFFeatures, 2015, (http://www.mathworks.com/help/vision/ref/detectsurffeatures.html).
[41] A. Ouanane, A. Serir, N. Djelal, Recognition of aggressive human behavior based on surf and svm, in: Systems, Signal Processing and their Applications (WoSSPA), 2013 8th International Workshop on, IEEE, 2013, pp. 396–400.
[42] S. Karaman, J. Benois-Pineau, R. Megret, V. Dovgalecs, J.-F. Dartigues, Y. Gaestel, Human daily activities indexing in videos from wearable cameras for monitoring of patients with dementia diseases, in: Pattern Recognition (ICPR), 2010 20th International Conference on, IEEE, 2010, pp. 4113–4116.
[43] E.B. Nievas, O.D. Suarez, G.B. García, R. Sukthankar, Violence detection in video using computer vision techniques, in: International Conference on Computer Analysis of Images and Patterns, Springer, 2011, pp. 332–339.
[44] R. Sivalingam, A. Cherian, J. Fasching, N. Walczak, N. Bird, V. Morellas, B. Murphy, K. Cullen, K. Lim, G. Sapiro, et al., A multi-sensor visual tracking system for behavior monitoring of at-risk children, in: Robotics and Automation (ICRA), 2012 IEEE International Conference on, IEEE, 2012, pp. 1345–1350.
[45] J.M. Girard, J.F. Cohn, M.H. Mahoor, S.M. Mavadati, Z. Hammal, D.P. Rosenwald, Nonverbal social withdrawal in depression: Evidence from manual and automatic analyses, Image Vision Comput. 32 (10) (2014) 641–647.
[46] I. Pavlidis, J. Levine, P. Baukol, Thermal imaging for anxiety detection, in: Computer Vision Beyond the Visible Spectrum: Methods and Applications, 2000. Proceedings. IEEE Workshop on, IEEE, 2000, pp. 104–109.