How Was It? Exploiting Smartphone Sensing to Measure Implicit Audience Responses to Live Performances

Claudio Martella[1*], Ekin Gedik[2*], Laura Cabrera-Quiros[2,4*], Gwenn Englebienne[3], Hayley Hung[2]
[1]VU University Amsterdam, Netherlands
[2]Delft University of Technology, Netherlands
[3]University of Amsterdam, Netherlands
[4]Instituto Tecnológico de Costa Rica, Costa Rica
c.martella@vu.nl, {e.gedik, l.c.cabreraquiros, h.hung}@tudelft.nl, g.englebienne@uva.nl

ABSTRACT

In this paper, we present an approach to understand the response of an audience to a live dance performance by the processing of mobile sensor data. We argue that exploiting sensing capabilities already available in smart phones enables a potentially large scale measurement of an audience's implicit response to a performance. In this work, we leverage both tri-axial accelerometers, worn by ordinary members of the public during a dance performance, to predict responses to a number of survey answers, comprising enjoyment, immersion, willingness to recommend the event to others, and change in mood. We also analyse how behaviour as a result of seeing a dance performance might be reflected in a people's subsequent social behaviour using proximity and acceleration sensing. To our knowledge, this is the first work where pervasive mobile sensing has been used to investigate spontaneous responses to predict the affective evaluation of a live performance. Using a single body worn accelerometer to monitor a set of audience members, we were able to predict whether they enjoyed the event with a balanced classification accuracy of 90%. The collective coordination of the audience's bodily movements also highlighted memorable moments that were reported later by the audience. The effective use of body movements to measure affective responses in such a setting is particularly surprising given that traditionally, physiological signals such as skin conductance or brain-based signals are the more commonly accepted methods to measure implicit affective response. Our experiments open interesting new directions for research on both automated techniques and applications for the implicit tagging of real world events via spontaneous and implicit audience responses during as well as after a performance.

Categories and Subject Descriptors

H.1.2 [**Models and Principles**]: User/Machine Systems—*Human Information Processing*; H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing—*Indexing Methods*; G.3 [**Probability and Statistics**]: Time Series Analysis

*C. Martella, E. Gedik, and L. Cabrera-Quiros contributed equally to this article

MM'15, October 26-30 2015, Brisbane, Australia
© 2015 ACM. ISBN 978-1-4503-3459-4/15/10 ...$15.00.
DOI: http://dx.doi.org/10.1145/2733373.2806276.

Keywords

Human behavior, proximity sensing, wearable sensors, accelerometers, arts, dance, audience response, mobile phones.

1. INTRODUCTION

Art and cultural events such as dance, art exhibitions, and concerts have an inherent value [14], which some studies have shown are correlated with a perceived quality of life [25] as well as self-rated health levels [26]. Driven by the high reward that such information would have for the art and cultural industry, the aim of this paper is to investigate ways of automatically measuring the response to the experience of art and cultural events as a means of enhancement for both consumers and practitioners.

As public subsidies for the arts and culture reduce, we argue that society loses a significant contribution to public life in both intellectual, emotional, and social stimulation. Art and cultural institutions increasingly find that they must be able to quantify the service they provide for monetary investment. The quantification of the effect of such events has the potential to benefit society by providing members of the public with a more targeted enhancement of art and cultural experiences and a way of showing that such experience on their own contribute significantly to the improvement of people's social lives. While this may appear to be a luxury to have in society, numerous studies have shown the benefits of art and cultural events for stimulating the social life of public spaces [42], and health and mental wellbeing [27, 32, 14].

In this paper, we target the task of quantifying people's experience and automatically predicting factors related to watching a modern dance performance. We show experimental results based on two different public modern dance performances. However, unlike prior works that have used biosignals or physiological sensing [36, 13], we hypothesise that behaviour, as sensed by sensors that one would find in smart phones, could also be used to sense affective responses to a performance. Our experiments show that using more pervasive sensors opens up huge possibilities for implicit affective sensing on a large scale in the wild.

By working closely for the last 2 years with Holland Dance (HD)[1], an organization whose role is to promote dance in The Netherlands, we have identified some key challenges to measuring audience response:

The limits of survey responses: Organisers of live performances are always interested to gauge audience opinions about the performance they organise — if they enjoyed a performance or enjoy dance performances in general, they are more likely to recommend it to others, thus sustaining the popularity of the art form. Such responses must also be obtained after the performance and do not therefore

[1]http://www.holland-dance.com

capture the spontaneous response of the performance audience to specific moments. Note that sentiment about a performance can also be assessed via social media but again requires audience members to actively participate in putting forward an opinion publicly [38].

Obtaining implicit measurements on a large scale: To our knowledge, most related work that tries to use implicit responses to visual stimuli such as movies [36, 13] or live performances [41] have tended to rely on physiological or brain activity measurements. While such signals are considered fairly reliable, equipment that is able to sense this data is still not particularly pervasive. As social norms would dictate, one tries to stay as quiet and still as possible when sitting and watching a live dance performance, so measuring implicit and measurable responses from pervasive sensors are likely to be even more challenging.

Obtaining detailed audience responses on a large scale: Even when survey responses are available for a performance, typical Likert scale questions cannot provide detailed insights into what moments of a performance could have triggered someone to dislike or like it. One way to circumvent this problem involves using free text answers, which can provide richer but still incomplete information about someone's experience that need to be manually processed. Going further, interviews can also be used, but provide more unstructured responses for those few who are willing to spend more time reflecting on their experience. They are, therefore, limited to an even smaller subset of an entire audience.

Quantifying the impact of a Performance on our Social Lives To our knowledge, most works focus on measurements obtained solely during the performance, measuring direct responses to it. However, the value of a performance can stretch beyond this period to its affect afterwards. For instance, it could affect ones mood, serve as a topic of stimulating discussion over drinks, leading to positive feelings about the entire performance and socialising experience. In an ideal case, its effects should last beyond the performance itself, perhaps even providing lasting memories that are recalled collectively by friends. This is perhaps the most difficult challenge but, if answered, even in part, would provide a broader metric to quantify the value of arts and cultural events.

More concretely, we make the following novel contributions; we show that (i) even when people are sitting and watching a live dance performance, they spontaneously react it via body movements that can be captured from a standard acceleration sensor, (ii) moments of common spontaneous bodily reaction correspond to memorable events in the performance as reported by survey responses relating to the performance, (iii) their reactions can be used to predict their enjoyment of the performance, whether they felt immersed in the experience, would recommend it to others, or thought dance performance changed their mood positively, (iv) and finally by considering the social context that surrounds the activity of going to a live dance performance, we also provide initial results using acceleration and proximity sensors, that suggest that a change in the mood of a person as a result of watching a live dance performance is reflected in their general body behaviour while mingling.

2. RELATED WORK

Traditional methods to investigate the response of an audience to a live dance performance make use of self-reports, such as surveys and interviews [6, 30]. Digital technologies can overcome the limitations of surveys and interviews, by giving more direct and fine-grained insights into the response of an audience. For example, the explosion in popularity of the social media, e.g. Twitter, and mobile computing has broadened the borders of a live performance, as fans comment and post information and opinions live to the online community [4]. Practitioners are interested in the activity of their audience through the social media to understand both their response and to leverage their activities as marketers of their performances [21, 20]. For example, some theatres, including Broadway, have experimented with so-called "tweet-seats" reserved for customers who promised to tweet about the performance live [1].

Other rather less pervasive technologies can also overcome the granularity issues of surveys using sensors. For example, work in neuroaesthetics use fMRI scanning to relate viewer responses to the aesthetics of the performance [5, 9, 7]. Moreover, tracking of eye gaze from video has been used when trying to distinguish novice from expert observers of dance [37]. Finally, physiological sensing such as galvanic skin response (GSR) sensors, have been investigated to measure the arousal of individuals watching a video of a dance performance, and its relationship with the individuals' self-reports [19]. Similarly, GSRs have been used also to measure the response to other types of live performance, such as comedy [41] and movies in a cinema [13].

These attempts show an increasing interest in quantifying the experience of arts and cultural events, such as live dance performances. Unlike these approaches, we advocate the use of pervasive sensors which are readily available in smartphones, which enable less obtrusive measurements and on a massive scale, compared to those obtained via physiological sensing. In particular, in this work, we focus on using acceleration and proximity sensors to measure people's reactions to live performance, which have been used thus far to measure very different phenomena.

Specifically, most work that consider accelerometers and people have addressed the problem of activity recognition of daily activities such as walking, running, sitting, climbing the stairs [18], daily household activities including eating or drinking, vacuuming or scrubbing, lying down [2], or identifying modes of transportation taken [31]. There is a trend moving towards the detection of medically relevant events, such as fall detection [11, 44], but all of these approaches focus resolutely on physical activities where the behaviour can be represented directly by quite specific movements of the body. It is possible to classify these types of activities with excellent performance, yet these activities are very different to analysing the response to a live performance. Few works do exist where less specific body movements have been classified. For example, Matic et al. also used acceleration to detect speaking status by strapping an accelerometer to the chest so that vibrations directly caused by speaking could be detected [24], Hung et al. [17] used body movements to predict socially relevant actions with a device hung loosely round the neck or for detecting conversations[16]. Such works highlight the potential of measuring spontaneous bodily responses to external stimuli using more pervasive sensing.

Apart from focusing on different activities and tasks, the above-mentioned works measure behaviour in environments that are far less challenging than a theatre where the audience sits in silence, and where the link between activity and behaviour is not as direct. The most similar work to our own was presented by Englebienne and Hung [12] who found that they were able to identify audience members as professors and non-professors from their behaviour while attending an inaugural lecture. Although they were sitting, the small movements made in reaction to parts of the lecture demonstrated implicit responses of interest to particular moments and content delivered during the lecture. However, they did not analyse whether reactions from the audience to the lecture correlated with enjoyment of the lecture, for example. Another closely related work where the audience response was measured was presented by Bao et al. [3] who investigated how users watching movies on abuffer tablet could have their implicit responses sensed by a wide variety of modalities from the tablet itself including the video, audio, tablet interactions, and accelerometer. In this case, movements from the tablet whilst the user was holding it were used to gauge responses. Using a multimodal approach they were able to predict the rating of users to movies they watched on the tablet. However, in this case, the user sat alone to watch the movies and was not inhibited by the social norms usually adhered to in an auditorium.

Proximity sensors have been used to study the interactions between individuals with approaches more similar to complex network analysis. Cattuto et al. [8] have used wearable sensors to analyse social interactions in crowded social settings, by means of proximity data collected through RFIDs. Martella et al. [23] used data collected through a series of wearable proximity sensors to identify the different communities attending a multi-disciplinary ICT conference. Roggen et al. [33] and Wirz et al. [43] proposed the usage of wearable sensors to discover spatio-temporal relationships between a number of individuals in the context of crowd dynamics. While these studies show that social relationship between individuals can be captured by means of spatio-temporal information, none of these works focus on the measurement of spatio-temporal relationship information in the context of arts and cultural events.

3. CASE STUDY 1: DIRECT RESPONSES TO A PERFORMANCE

To inspect whether it is possible to predict responses to a performance by using data collected with wearable sensors, we have conducted an experiment in an actual dance performance. We start this section by explaining the characteristics of the performance and the resulting dataset we obtained. Then, we describe the features we used in the data analysis and classification experiments. The next two sections present the qualitative analysis and classification experiments we did based on questionnaire responses. The qualitative analysis aims to show that the movements of the audience are informative of their reactions to the performance. We do this by showing that the captured activity of the participants and common responses correspond to the salient events happening in the performance. The classification experiments present our methodology for automatically predicting a participant's evaluation of the event. The performance and further analysis presented show that the proposed method is indeed promising and worth investigating further.

3.1 Data Collection

The sensor set-up: We organised a data collection experiment during a dance performance. This event consisted of almost an hour and a half of performance without intermission. It mainly consisted of dancing but also included monologues by the performers in Italian, while the music was mainly based on live cello arrangements but also included pre-recorded songs. Using triaxial acceleration and IR cameras (for additional data verification), we recorded 41 participants watching the performance. The accelerometers were located in a custom-made device hung around every participant's neck. These devices recorded at 20Hz and were synchronised to a global time obtained by communicating through a wireless network. However, due to hardware malfunctions, only 32 accelerometers recorded data. In addition, the performance was recorded using a GoPro Hero +3 to analyse salient moments (i.e. favourite moments that were reported by the participants).

Survey responses: To evaluate the experience, a questionnaire was filled in by all 41 participants after the performance. Each questionnaire had 12 questions, where each group of three questions aimed to measure one aspect of the experience. The four aspects were "enjoyment", "recommendation (to a friend)", "immersion" and "mood changes". Each participant evaluated these aspects of the performance on a ten-point Likert scale, where one means "I completely disagree" and ten means "I completely agree". For measuring enjoyment, we adapted and selected questions presented in [35]. For the task of immersion, we selected involvement questions from the Igroup Presence Questionnaire [34]. For recommendation we used items from O'Brien's questionnaire [28]. Each of these questions were carefully chosen to measure each task and slightly adapted to match our scenario. We formed the questions regarding mood by ourselves. Given that the majority of the audience members were Dutch, we used a back-translation procedure to ensure that each questionnaire item was accurately describing the original English wording. This involved finding three different Dutch speakers to translate the questions from English to Dutch, then from Dutch to English and then from English to Dutch again ensuring that the finally chosen words best matched the original English. The complete set of questions asked in this questionnaire in both English and Dutch are listed in the Appendix 9. From the total number of participants, 32 responded with the Dutch questionnaire and 9 to the English one.

Of the 32 participants with working accelerations, 25 reported a favourite moment of the performance. Two moments were particularly memorable: the *motorcycle sequence* was declared as favourite by 32% of the participants, and the *bolero finale*, favourite of 52% of the subjects. Note that in some cases that participants declared more than one favorite moment.

3.2 Feature Extraction

We used the variance of the accelerometer readings, which is expected to act as a proxy for the physical activity level of the participants. Our assumption was that both subtle as well as more expansive movements of the the participants is related to the experience of the event. We expect participants with different evaluations of the event to have different movement patterns throughout the event, especially during salient parts of the performance. We calculate the variance in a sliding window of 2 seconds with 1 second shift, which corresponds to 40 samples for each window with a shift of 20 samples. This window size is carefully selected to capture the subtle and short variations in motion while still preserving a fine time scale and is empirically proven to perform well.

We extract features from an interval of $\sim$79 minutes, starting just before the first piece, when all participants are seated and ending when the final piece of the performance finishes. With this we obtain 4705 different variance values for each axis. Before calculating the variance along each axis, each axis is normalised by computing the z-score to remove interpersonal differences. We also calculate the variance of the magnitude, resulting in 4 different variance values for each interval. For the qualitative analysis, we only use the variance of the magnitude. For classification, we treat the variance values of each axis as well as the magnitude as our features, resulting in a 18820 dimensional feature vector for each participant. This feature choice restricts the representation to be temporally dependent. Therefore we may not capture cases where two participants liked (or disliked) the event during different parts of the performance. With another formulation of the problem, like using bag-of-words or a multiple instance learning approach, temporal dependence can be avoided. Although such cases are quite probable, in this experiment we assume that participants with similar evaluations of the performance tend to respond similarly during salient parts of it.

3.3 Data Analysis

The variance in magnitude signals from all the participants were compared against each other to create a pairwise co-occurrence measurement over time using Mutual Information (MI). These signals were calculated over a sliding window (size of 60 samples) shifted

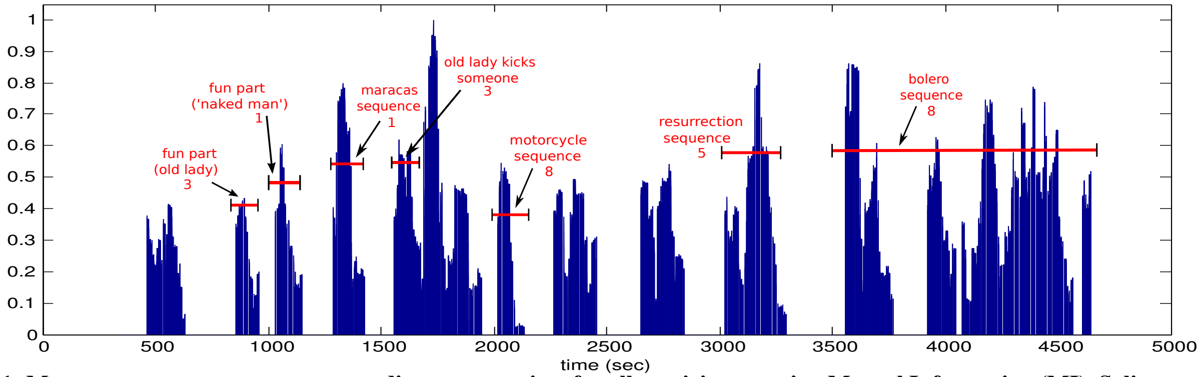


Figure 1: Mean co-occurrence measurement distance over time for all participants using Mutual Information (MI). Salient moments are highlighted in red with number of appearance

by one sample, resulting in a vector of co-occurrence over time between two participants. The mean mutual information at each time interval was calculated, allowing us to evaluate the collective response of the participants to the performance over time. We hypothesised that salient moments would correspond to a high MI among all participants. These moments were chosen using a Otsu threshold [15] on the values of the computed signal. Figure 1 shows these salient moments captured by points where the mutual information goes beyond the threshold (blue), as well as the frequency of the reported favorite moments (red) that appeared in the free text survey responses. The time stamps were generated by manually identifying the time period(s) where the reported favourite moments occured. Notice that the two moments declared as favourites for the majority of participants (*motorcycle* and *bolero finale*) are captured . This shows that memorable moments for people during these events can be captured by their coordinated movements, as they share the experience.

Furthermore, the role of music during the performance is also interesting and we want to understand its effect, if any, on the collective behaviour of the participants. To do so, we looked at the sound intensity of the performance as obtained from the video and annotated song changes. Figure 2 shows the sound intensity of the performance (green) compared to the normalized co-occurrence measurement for MI (blue). The performance's songs are also highlighted in this figure in red. Here, it can be seen that although the music had a correlation with the response of the public in a performance in certain sequences, other moments of high mean MI are also correlated with acts with no music. This suggests that the music may not have been the main factor stimulating coordination between our participants. For this reason, we decided that the song changes would not be useful for the following classification experiments and have focused solely on acceleration data instead.

With this preliminary qualitative information, the next subsection describes our classification experiments using the four questionnaire tasks mentioned in Section 3.3.

3.4 Classifying Experience

Labelling samples

We use a standard pattern recognition approach to automatically predict the responses of participants. We start by labelling our participants according to the questionnaire answers they gave. For each group of three questions, we obtained one numerical value by averaging and rounding the three answers. This way, we obtain four different labels for our participants, where each label corresponds to the one of the tasks. We divided the participants into two classes

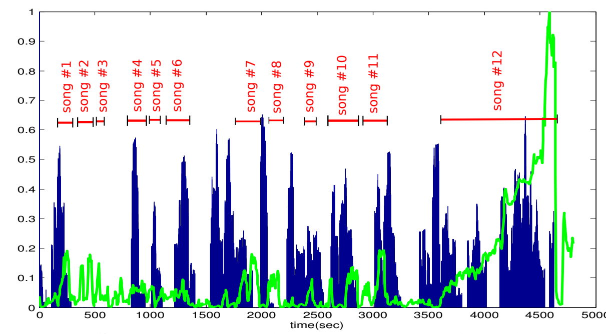


Figure 2: Sound intensity of the performance (green) compared against the normalized co-occurrence measurement calculated by MI (blue).

for each task corresponding to a "positive" and "negative" report on their experience of the performance. Participants whose averaged answer to any group of questions was below 5 was placed in the negative class for that task, meaning this participant, either did not enjoy the event, would not recommend the performance, did not feel immersed throughout the performance or did not think the performance uplifted their mood. The positive class thus contains participants who gave positive responses to the questions.

In this study, we mainly focus on enjoyment and recommendation, since these two are questions with clearer indications, but still provide results for immersion and the mood changes as we believe they can help in obtaining a general understanding of the performance's effects on the participants. After defining the classes for each task, our class distributions were not always balanced. For "enjoyment" and "recommendation", the majority of participants (26) gave positive answers. 22 participants thought "the performance affected their mood positively". Therefore, for these tasks, the problem becomes challenging for the smaller negative classes. These imbalance in class distributions also affected our choices of performance measure; in addition to accuracy, we also provide the balanced accuracy so each classes contributes equally to the final measure. The distribution for the "immersion" task is more balanced with 17 participants in the positive class.

Methodology

To emphasise the connection between the information contained in the motion data and the participants' experience of the event, in our classification experiments we focus on a simple set of features and a well-understood classifier. More precisely, our features are the variance of acceleration along each axis and of acceleration magnitude, extracted with the aforementioned setup. We selected a Linear SVM as our classifier. Since the number samples is limited, we chose to use a model with few parameters . For evaluating the performance of our method, we used leave-one-out cross validation, training with 31 samples and with the remaining one is used for testing. The hyperparameters of the SVM are also selected using cross validation on the training set.

As stated in Section 3.2, representing the features using the whole performance would require classification on a 18820-dimensional feature vector for each participant. Since we do not expect all intervals to be equally informative and to avoid the curse of dimensionality, we decided to use a filtering approach which selects the informative intervals before feature extraction.

To do so, we selected a Dynamic Time Warping (DTW) distance computed over a window (sizes were set from 20 to 60 samples) with a shift size of 1 sample . Similar to the mean MI co-occurrence vector used in Section 3.3. Our assumption here is that if we select the intervals where the average DTW distance between each pair is significantly higher than the rest, we should end up with time intervals that are more discriminative than the rest. In an ideal scenario, intra-class distances should stay relatively stable throughout the event, so the parts where the average DTW distance between pairs is high coorespond to intervals where the intra-class distances are maximised. We hypothesise that using this metric provides better discrimination between classes compared to using mutual information where moments of high mutual information could also correspond to moments where the mean DTW is low and both classes would be almost indistinguishable . Empirical results using a threshold on the mean MI supported this claim, with performance scores significantly lower than the proposed method for the majority of tasks.

Using the same setup explained in Section 3.3, we detect the intervals where pairwise DTW distances are significantly higher than the rest and extract variance features for each person from these intervals only. The number of remaining intervals after filtering depends on the window size selection. In our experiments, where we have used windows of 20 to 60 samples, the number of selected intervals ranged from 86 to 830. Finally, after interval selection and feature extraction, we perform further dimensionality reduction by applying principal component analysis (PCA) to the feature vectors. We keep the principal components which preserve the 99 percent variance of features and use them for training and testing our model. This resulted in 12 and 19 resulting feature dimensions.

Results

The performance obtained by the proposed method, with different window size selections when using the thresholded DTW distance to pre-filter the salient intervals, are presented in Table 1. This table also includes the performance scores obtained without interval selection. The statistically significant results between those using interval selection and those using whole event are specified with a sign '*'.

Method \ BAcc | Acc(%)	Enjoyment	Recommend	Immerse	Mood
DTW IS(20 Sample)	60 | 66*	61 | 78	63 | 63	60 | 63*
DTW IS(40 Sample)	90 | 94**	70 | 81	53 | 53	55 | 66*
DTW IS(60 Sample)	78 | 84**	71 | 84	49 | 50	48 | 63*
Whole Event	36 | 38	74 | 78	51 | 50	38 | 38

($* \rightarrow p<0.05$) ($** \rightarrow p<0.01$)

Table 1: Performance scores obtained with different methods

The performance obtained without interval selection, which are reported in the final row of Table1, are unsatisfactory in general. Any task other than predicting recommendation has an accuracy lower than or equal to the proposed method, regardless of the window size.

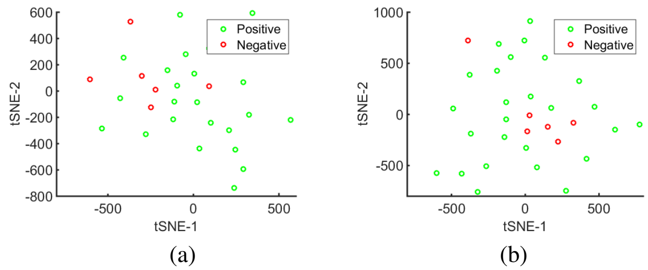


Figure 3: Non-linear embedding of feature vectors for the enjoyment class for (a) the whole event and (b) using interval selection with a 40 sample window.

We should note that we did also apply PCA to the feature vectors for the non-filtered method. However, these scores showed that the extracted principal components were still affected by variance features extracted from many non-informative intervals, supporting our claim of interval selection is necessary.

The performance obtained with different window sizes have some interesting implications. For the tasks of immersion and mood, we obtained the highest performance with a window of 20 samples, corresponding to one second. With the increasing window size, our performance for these tasks dropped below random. We can see that this same window size is not optimal for the enjoyment and recommendation tasks. This could suggest that some tasks are shorter in time scale than others. However, we need more data to draw solid conclusions about such implications. It should be also noted that for the mood task, our proposed method always provides a significantly better result than the whole interval method, regardless of the window size.

Another interesting implication can be seen in the performance scores for "recommendation". For this task, the highest performance score is obtained when the whole event was used. Supporting this idea, the performance for this task increases as the window size increases. We should note that using a wider window may not always guarantee that more intervals are selected for classification. None of the results presented for the recommendation task showed significant improvement over the whole event baseline. However, the recommendation accuracy is very high for the whole event, suggesting that the behaviour of people who are interested in recommending dance performance to others could be distinguished from those who did not tend to recommend dance performances to others, and that this was independent of any particular moments during a performance.

Finally, we can see that for enjoyment, one of our most important tasks, we obtain relatively high performance, with the highest score of 94% accuracy and 90% balanced accuracy. With a window size of 40 samples, only one sample from each class is misclassified, resulting in high precision and recall scores for both classes. This result is significantly better than the whole interval method ($p>0.01$). To further investigate how the interval selection affects the distribution of samples, we visualise the filtered set of feature vectors of each participant using t-SNE [39], a non-linear embedding technique. The resultant two-dimensional embedding shown in Figure 3 illustrates that interval selection allows data points in the negative class to be clustered together in the feature space.

We also experimented with computing DTW distances on the raw accelerometer magnitude signal, instead of the variance. This experiment resulted in performance scores that were worse than random for "enjoyment", "immersion" and "mood". For "recommendation", we obtained a balanced accuracy score of 85 percent. This result is quite interesting, since the highest performance we obtained for "recommendation" in Table 1 was the case where the whole event was used. However, in general it can be said that this experiment empirically supported our claim that the variance in acceleration is a valid feature for our experiments, both as a feature and for the interval selection using the thresholded DTW distance.

3.5 Further Analysis of Salient Moments

This section aims to further explain the salient moments of the performance, now relating these with the classes identified in Section 3.4. For space reasons, we focus on the enjoyment task as it has given us the best performance. The pairwise similarity measurements from the previous qualitative analysis were separated into two groups for each task: the ones who completely agreed with the statement and everyone else. For each group, the same unified similarity measurement as described in Section 3.3 was calculated and the salient moments were obtained using the Otsu threshold level. Since the goal is to asses the similarity of people within the same class, pairs of different classes are left out.

Figure 4 shows the measurements of mean MI over time for both classes in the enjoyment task (where the negative class was plotted below the positive class). Notice how the two moments considered as favorites for the majority of participants (*motorcycle sequence* and *bolero finale*) reappears for the mean MI in the group that enjoyed the performance but not for those who disliked it. Actually, there is almost no overlap between the salient moments for the classes 1 and 2. This reaffirms that specific acts or sequences in a performance (or movie) can have a significant impact in the final assessment of enjoyment.

Furthermore, Figure 5 shows the mean DTW distance for members within the 'Enjoy' class (blue), the 'Not Enjoy' class (green) and all pairs in opposing classes (red). The 'Not Enjoy' class resulted in a higher overall DTW distance, over the complete performance, compared to the 'Enjoy' class. This might indicate a lack of synchrony among people who dislike the performance, which echoes findings by Wang and Cesar with Galvanic Skin Response measures to an audience's reaction to a live performance [40].

4. CASE STUDY 2: IMPACT OF A PERFORMANCE ON SOCIAL BEHAVIOUR

Section 3 provided interesting insights into how the response to a dance performance can be measured with pervasive sensing. However, while working with HD, we came across a different perspective on the problem. Can we quantify the influence of a performance on the audience even after the performance is over?

There is clearly a context that surrounds the event itself — typically, people will attend a performance with friends and/or family, may come for a drink beforehand and stay for drink afterwards. We hypothesised that people's social behaviour (as measured through proximity and acceleration) could also be affected by watching a dance performance. As a business model HD was already co-organising networking events around dance performances together with two local networking organizations. The idea was that the dance performance could be an occasion to enhance the networking event, and the co-located networking event would encourage more people to watch dance.

To investigate this hypothesis, we decided to investigate whether we could measure differences in how people socialized during the event. Hence, we measured mingling behavior during two networking sessions, one right before a dance performance and another right after it. With HD and regional networking groups, we co-organised a networking event with 48 volunteers. The same sensing devices described in Section 3 were used. An example snapshot of the mingling data is shown in Figure 6.

Figure 6: Snapshots of the instrumented mingling room.

Although a networking event is not exactly the same as the more casual ways that people might attend dance performances socially, we believe this initial investigation provides a feasibility study for larger scale less controlled studies in the future.

Similar to previous work using proximity sensors to analyse social behaviour in conferences [23], musea [8], and work-places [29], we used proximity sensors as proxies for face-to-face social interactions, together with accelerometer data as described above. Two months after the experiment, we published sensor analysis results for the volunteers and asked them to answer a survey about their experience of the dance performance and the networking event.

4.1 Setup

Each device we used during the previous study is also equipped with a wireless radio, which we used to broadcast the device's unique identifier (ID) every second up to a distance of some 2-3 meters. The reception of such broadcast by the devices nearby is considered a proximity detection. Since the device lies on the front of the torso, the radio transmission is shielded by the body, hence restricting the proximity sensing mostly towards the front of the individual. The device logs each detection on the on-board storage along with their timestamps. We used an energy-efficient MAC protocol [10] to allow the devices to communicate their IDs and detect each other's proximity. Because of the unreliability of the wireless medium, we used a density-based filtering technique to increase the sensitivity of the signal for detecting face-to-face proximity [22].

To measure whether people were interacting or not, we processed the proximity detections collected by the devices as follows. For each pair of individuals, we computed the intervals where pairs were continuously facing each other, formally $[t_i, t_j]$ where t_i is the timestamp of the first detection and t_j is the timestamp of the last detection of the interval. Because pairs can be close for multiple non-overlapping time intervals during the same measurement, we computed multiple intervals for the same pair. Here, we refer to an interval of proximity between any pair as an *interaction*. For our experiments we considered only intervals of proximity longer than 60s to indicate interactions.

4.2 Results

Since some attendees did not attend the entire event, only the data from 35 of the participants was available for analysis. We first hypothesized a difference in the length or the number of interactions between the two sessions. For example, one could imagine that individuals would interact in longer conversations, or with more people. In Figure 7(a) we present the distribution of the length of the interactions for the two sessions (from here on referred to as round 1 and round 2) across all the individuals. In both rounds shorter interactions are predominant. Note that as drinks were served at the bar, during both rounds often individuals left a conversation to fill their glass and went back right afterwards to the same conversation, which would be measured as two distinct interactions. No significant mean difference was seen between the distribution in interaction length for the two rounds. In Figure 7(b) we present the distribution of the number of distinct interactions for round 1 and round 2 across all the individuals.

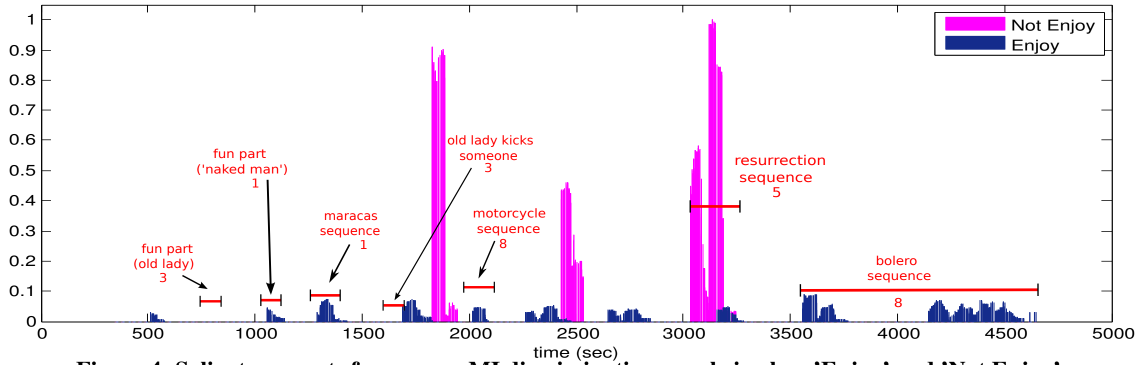


Figure 4: Salient moments from mean MI discriminating people in class 'Enjoy' and 'Not Enjoy'

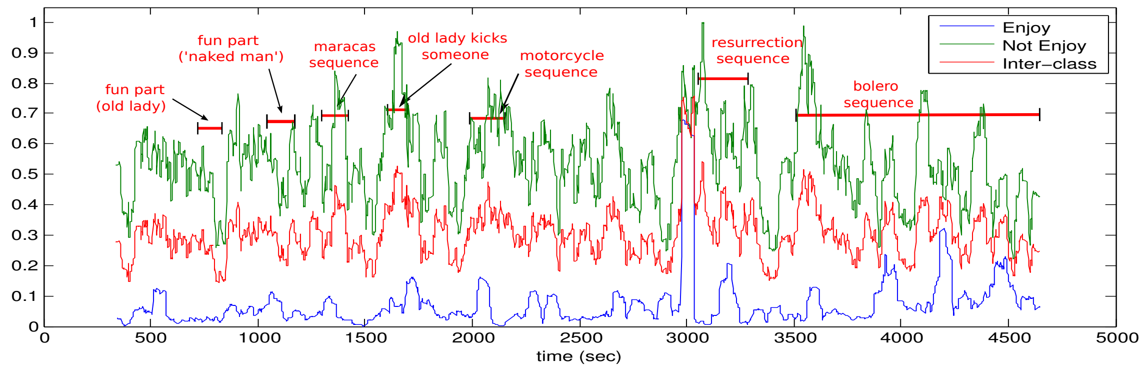
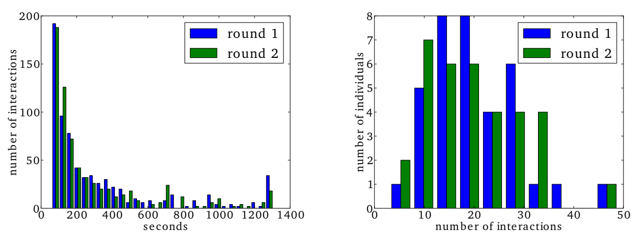
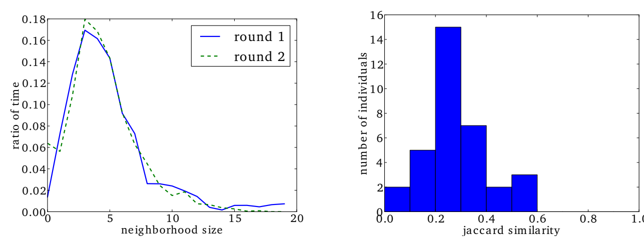


Figure 5: Similarity measurement using DTW for each class in the enjoyment task.



(a) interaction lengths (b) interaction numbers

Figure 7: (a) Distribution of the lengths of the interactions during the two rounds. (b) Distribution of the number of interactions for round 1 and round 2 across all the individuals.



(a) conversational group sizes (b) change in partners

Figure 8: (a) Relative amount of time sensors detected a certain number of other sensors (at a specific moment in time). (b) Distribution of the jaccard similarity across the individuals.

A second difference we hypothesized was in the size of conversational groups. For example, people could be engaged in conversations involving more people, or conversely more one-to-one conversations, perhaps to discuss the content of the performance. We define a *neighborhood* as the set of nodes a sensor a detects at a given moment in time, i.e. the individuals in physical proximity of the individual wearing sensor a. In Figure 8(a) we present the distribution of neighborhood size with respect to the amount of time they were observed together, expressed as a ratio over the round duration. In other words, it represents the amount of time individuals have spent in proximity to another n individuals. The results show a peak around four individuals, a reasonable group size for a conversation. Similar to the interaction lengths, the two distributions look very similar.

The third hypothesis regarded changes in conversational partners. For example, people could be interacting with the same individuals as before the performance, or be stimulated to engage with others. To this end, for each individual we computed the *Jaccard similarity* between the set of participants an individual has interacted with during the two rounds. Given two sets of IDs R_1 and R_2, the Jaccard similarity function is defined as $J(R_1, R_2) = \frac{|R_1 \cap R_2|}{|R_1 \cup R_2|}$ and computes a value in the interval $[0, 1]$. Figure 8(b) presents the distribution of the Jaccard similarity across all individuals between round 1 and round 2. The results show that although the mingling pattern of the individuals did not change between the two rounds, they did interact with different individuals. In particular, they changed at least 50% of their interaction partners between round 1 and round 2 (mean 0.278 and standard deviation 0.121).

4.2.1 Acceleration

The image emerging from the pure proximity measurement is that of an ordinary mingling event. Overall, these results indicate that the volunteers, as a group, applied a consistent pattern in their mingling behavior during the two rounds, a pattern that they used, however, to target different conversational partners between the two rounds. The measurement pictures a socializing context, but it is difficult to reach conclusions about the impact of the performance. For this reason, we focused on the acceleration data as well.

Similar to the direct approach, we used variance in the acceleration magnitude as the main feature. We then correlated the participant's self-reported behavior with our findings. Correlation between the answers to question "Do you think the performance had an effect on your mood? Yes|No" and the difference between the acceleration magnitude variance in round 1 and 2 is computed. Since not all participants filled in the post event survey and some accelerometers failed due to a firmware bug, we were only able to use the accelerometer data from 14 participants.

The variance values are extracted using the whole intervals for round 1 and 2. A statistically significant ($p = 0.02$) positive correlation value of 0.60 wass obtained as a result. This correlation supports our hypothesis that the mood change can be linked to implicit behaviour as measured by acceleration, though we would like to verify this with a larger dataset later. In conclusion, the results suggest that while individuals acted similarly as a group in terms of networking behaviour captured by the proximity sensors, the quality of those interactions seemed different between the two sessions, as captured by the accelerometers.

5. DISCUSSION

Our experiments show that using sensors commonly available in smart phones, we are able to leverage group dynamics in the audience to predict an audience's experience of a dance performance and that their implicit responses to the performance can automatically highlight salient moments in it.

In addition, to our knowledge, it is the first time that the context surrounding a performance has also been analysed. Our results suggest that measureable changes in social behaviour before and after watching a dance performance are correlated with an audience's perceptions of how a performance can affect their willingness to socialise as well as their mood. Given that the problem of automatically analysing and quantifying the experience of audience members of live performances is a difficult and multi-faceted domain to conduct experiments, we believe that our first steps show much potential for further devising automated methods to enrich live artistic performances via implicit responses.

5.1 Opportunities

Our experiments shed light on the huge and under-explored potential of linking implicit responses from pervasive sensing to augment digital signals that already connect with live performance such as verbal expressions of sentiment via social media. Given the potential of acceleration and proximity to be measured pervasively, we identify a number of key areas in which this information, when coupled with multimedia systems could be of real societal benefit.

Enriching the experience of a live performance Live performances tend already to have much multi-media data associated with them such as advertisements of the event on the Internet, video or audio recordings from smart phones, the associated social media, critical reviews of a performance in news or blog posts, attendees and their associated social media profiles, to name but a few. However, in all cases, online responses to an event requires an active and declarative response by audience members. One could easily imagine an audience member's experience of an event could be further enriched by seeing whether their responses to the performance matched that of others in the audience.

Live performance recommendation By using the acceleration signal generated by the sensor data, we were able to capture spontaneous responses. Our experiments show that the temporal characteristics of the variation in acceleration were similar for people who enjoyed the performance and sufficiently different for those that didn't. An immediate question following this would be whether such response signatures could be used for recommendation via collaborative filtering – if person A who responded similarly at particular moments of a live performance to a person B also like other performances that person B enjoyed?

Benefitting performers One could easily imagine that the salient moments in the performance that were identified automatically in our experiments could be used to verify or highlight key moments of audience response. While our analysis was performed afterwards offline, one could also imagine that such information could be provided in real time to the performers while they are performing. Moreover, the implicit responses from multiple audience members show that even if they do not report certain salient events in the performance to be memorable, their implicit responses can still provide supporting evidence for more sparse explicit survey responses.

Benefitting organisers of live performances For many, going to a live performance involves both the experience of the performance as well as the social event surrounding going to the performance. The national dance organisation that we worked with is aware of this and needs quantitative proof of its benefit. Importantly, from both events that we organised, volunteers paid to take part so they saw inherent value in it. Both experiments that we carried out show that dance performances can have an effect on how people behave and that these responses reflect positive experiences to the performance. This suggests that further studies should be carried out to investigate what triggers people to recommmend dance performance in general and using pervasive sensing provides a realistic means of doing this.

5.2 Open Challenges

Based on our experiences, it seems that ordinary members of the public are interested in being measured about their responses to a dance performance and having their sensed responses given back to them. However, working with real events with members of the public in dynamic and uncontrolled environments leads to a number of open challenges.

From controlled to uncontrolled large scale measurement First, to obtain large scale measures, we would need to use people's personal smart phones to record their behaviour. Like in our experiments, better responses can probably be gained from hanging their phones around their neck rather than keeping it in a pocket or handbag. Aside from requiring a special pouch to hang the phone around the neck (which could be easily made at low cost and large scale), we are not able to prevent, for example, tampering with the phone during the performance. Moreover, to investigate the role that live performances could have in improving people's social lives, we must also be able to measure their behaviour before and after the performance, which requires further collaboration by audience members unless some other method of incentivisation is provided.

Handling low numbers of explicit responses Implicit responses to a live performance can only be better understood when coupled with survey responses. However, as reported by HD, people tend not to answer surveys about a performance voluntarily unless they have extreme views about it. There are two ways of considering the problem. One perspective is to address how to more easily obtain even light declarative forms of sentiment about a performance (e.g. 'like' vs. 'dislike') – perhaps socialised incentives could lead to more willingness to report such information. Another perspective would be to consider label propagation techniques to estimate the reaction of the larger unlabelled data.

Subjectivity of responses Perhaps one of the biggest challenges remains in the fact that even if a set of people enjoyed a performance, it is highly probable that they enjoyed different parts. There is no guarantee that responses to the dance performance will be the same for everyone, particular if performances are unstructured or more abstract. Fortunately, the performance that we analysed was less abstract, having quite specific components that could be easily referred to. It remains an open question as to how more abstract performances would be responded to and whether they would as easily analysed.

6. CONCLUSIONS

We have made a first investigation of how a seated audience's perception of a performance can be perceived and measured from their body movements using an accelerometer that exists typically in smart phones. We analysed whether subtle and complex concepts such as "enjoyment", "immersion", an in improvement in mood as results of the performance, and whether participants would "recommend" dance in general, would be reflected in body motion using a simple accelerometer hung around the neck. Using the variance of the acceleration, we were already able to predict with 90% accuracy whether somebody enjoys the performance, which performed significantly above a random baseline.

Importantly, joint coordination in the variance in acceleration helps to distinguish salient from non-salient moments, which lead to significant improvements over using each person's body movements from the entire performance period. As well as the obvious usefulness to the entertainment industry of such direct measurements of an audience's reaction, we have also made a first attempt to measure the role that a live performance can have on the social behaviour that precedes and follows it. Our experiments shows huge promise in enabling us to measure the implicit responses of people while watching a live performance without the need for more traditional sensing approaches using physiological or brain signals. However, and perhaps more importantly, our experiments demonstrate the potential of quantifying the experience of 'a cultural night out', highlighting the relevance of the social context in moderating an individual's enjoyment of an event.

7. ACKNOWLEDGEMENTS

For the case study presented in Section 3, we would like to thank the Distributed & Interactive Systems group (`http://www.dis.cwi.nl`) at CWI for their involvement in designing, preparing, and executing the event. Thanks to Jesse Krijthe and Wouter Kouw for assisting with the back translations of the questionnaires, and Julian Kooij for assisting with the translation of the Dutch free text survey responses. We would also like to thank the staff at the Lucent Theatre for their logistic support. Thanks to the board of the networking groups Young the Hague, JongOCW, and Teater aan het Spui for their assistance in the logistics the case study in Section 4. Thanks to Djana Eminovic from Holland Dance for her assistance in the preparation and execution of both studies described in Sections 3 and 4. We thank all the anonymous volunteers who took part in both studies. This publication was partially supported by the Dutch national program COMMIT, the European Commission under grant agreement number 601033 - MOnarCH, the Costa Rican Institute of Technology, and is affiliated with the Delft Data Science Consortium.

8. REFERENCES

[1] USA Today. Providence theater experiments with 'tweet seats', 2013. `http://www.usatoday.com/story/tech/2013/01/27/theater-tweet-seats/1868693/`.

[2] L. Bao and S. Intille. Activity recognition from user-annotated acceleration data. *Pervasive Computing*, page 1Ű17, 2004.

[3] X. Bao, S. Fan, A. Varshavsky, K. Li, and R. Roy Choudhury. Your Reactions Suggest You Liked the Movie: Automatic Content Rating via Reaction Sensing. In *Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing*, UbiComp '13, page 197Ű206, New York, NY, USA, 2013. ACM.

[4] L. Bennett. Patterns of listening through social media: online fan engagement with the live music experience. *Social Semiotics*, 22(5):545Ű557, 2012.

[5] B. Bläsing, B. Calvo-Merino, E. S. Cross, C. Jola, J. Honisch, and C. J. Stevens. Neurocognitive control in dance perception and performance. *Acta psychologica*, 139(2):300Ű308, 2012.

[6] A. S. Brown and J. L. Novak. *Assessing the intrinsic impacts of a live performance*. WolfBrown San Francisco, CA, 2007.

[7] B. Calvo-Merino, C. Jola, D. E. Glaser, and P. Haggard. Towards a sensorimotor aesthetics of performing art. *Consciousness and cognition*, 17(3):911Ű922, 2008.

[8] C. Cattuto, W. Van den Broeck, A. Barrat, V. Colizza, J. Pinton, and A. Vespignani. Dynamics of Person-to-Person Interactions from Distributed RFID Sensor Networks. *PLOS ONE*, 5(7):e11596, 07 2010.

[9] E. S. Cross, L. Kirsch, L. F. Ticini, and S. Schütz-Bosbach. The impact of aesthetic evaluation and physical ability on dance perception. *Frontiers in human neuroscience*, 5, 2011.

[10] M. Dobson, S. Voulgaris, and M. van Steen. Merging ultra-low duty cycle networks. In *Proceedings of the 41st International Conference on Dependable Systems & Networks (DSN 2011)*, 2011.

[11] C. Doukas, I. Maglogiannis, P. Tragas, D. Liapis, and G. Yovanof. Patient fall detection using support vector machines. *Artificial Intelligence and Innovations 2007: from Theory to Applications*, page 147Ű156, 2007.

[12] G. Englebienne and H. Hung. Mining for motivation: using a single wearable accelerometer to detect people's interests. In *Proceedings of the 2nd ACM international workshop on Interactive multimedia on mobile and portable devices*, IMMPD '12, page 23Ű26, New York, NY, USA, 2012. ACM.

[13] J. Fleureau, P. Guillotel, and I. Orlac. Affective Benchmarking of Movies Based on the Physiological Responses of a Real Audience. In *Proceedings of the 2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, ACII '13, page 73Ű78, Washington, DC, USA, 2013. IEEE Computer Society.

[14] D. Fujiwara, L. Kudrna, and P. Dolan. Quantifying and Valuing the Wellbeing Impacts of Culture and Sport. Technical report, UK Department of Culture, Media and Sport, Apr. 2014.

[15] R. Gonzalez and R. Woods. *Digital Image Processing*. Prentice Hall, 2008.

[16] H. Hung, G. Englebienne, and L. Cabrera Quiros. Detecting conversing groups with a single worn accelerometer. In *Proceedings of the 16th International Conference on Multimodal Interaction*, page 84Ű91. ACM, 2014.

[17] H. Hung, G. Englebienne, and J. Kools. Classifying Social Actions with a Single Accelerometer. In *Ubicomp*, page 207Ű210, 2013.

[18] J. R. Kwapisz, G. M. Weiss, and S. A. Moore. Activity recognition using cell phone accelerometers. *ACM SIGKDD Explorations Newsletter*, 12(2):74Ű82, 2011.

[19] C. Latulipe, E. A. Carroll, and D. Lottridge. Love, hate, arousal and engagement: exploring audience responses to performing arts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, page 1845Ű1854. ACM, 2011.

[20] A. Leask, A. Hassanien, and P. C. Rothschild. Social media use in sports and entertainment venues. *International Journal of Event and Festival Management*, 2(2):139Ű150, 2011.

[21] M. R. Lockstone, L. Olga Junek, S. Hudson, and R. Hudson. Engaging with consumers using social media: a case study of music festivals. *International Journal of Event and Festival Management*, 4(3):206Ű223, 2013.

[22] C. Martella, M. Dobson, A. van Halteren, and M. van Steen. From Proximity Sensing to Spatio-Temporal Social Graphs. In *Pervasive Computing and Communications (PerCom), 2014 IEEE International Conference on*, 2014.

[23] C. Martella, A. van Halteren, M. van Steen, C. Conrado, and J. Li. Crowd Textures as Proximity Graphs. In *IEEE Communications Magazine*, 2014.

[24] A. Matic, V. Osmani, A. Maxhuni, and O. Mayora. Multi-modal mobile sensing of social interactions. In *Pervasive computing technologies for healthcare (PervasiveHealth), 2012 6th international conference on*, page 105Ű114. IEEE, 2012.

[25] A. C. Michalos and P. M. Kahlke. Arts and the perceived quality of life in British Columbia. *Social indicators research*, 96(1):1Ű39, 2010.

[26] T. Nenonen, R. Kaikkonen, J. Murto, and M.-L. Luoma. Cultural services and activities: The association with self-rated health and quality of life. *Arts & Health*, 0(0):1Ű19, 0.

[27] C. L. Nightingale, C. Rodriguez, and G. Carnaby. The Impact of Music Interventions on Anxiety for Adult Cancer Patients A Meta-Analysis and Systematic Review. *Integrative cancer therapies*, 12(5):393Ű403, 2013.

[28] H. L. O'Brien and E. G. Toms. The development and evaluation of a survey to measure user engagement. *Journal of the American Society for Information Science and Technology*, 61(1):50Ű69, 2010.

[29] A. Pentland, T. Choudhury, N. Eagle, and P. Singh. Human dynamics: computation for organizations. *Pattern Recognition Letters*, 26(4):503Ű511, 2005.

[30] M. Reason and D. Reynolds. Kinesthesia, empathy, and related pleasures: An inquiry into audience experiences of watching dance. *Dance Research Journal*, 42(02):49Ű75, 2010.

[31] S. Reddy, M. Mun, J. Burke, D. Estrin, M. Hansen, and M. Srivastava. Using mobile phones to determine transportation modes. *ACM Transactions on Sensor Networks (TOSN)*, 6(2):13, 2010.

[32] M. Ritter and K. G. Low. Effects of dance/movement therapy: A meta-analysis. *The Arts in Psychotherapy*, 23(3):249Ű260, 1996.

[33] D. Roggen, M. Wirz, D. Helbing, and G. Tröster. Recognition of Crowd Behavior from Mobile Sensors with Pattern Analysis and Graph Clustering Methods. *Networks and Heterogeneous Media*, 6, 2011.

[34] T. Schubert, F. Friedmann, and H. Regenbrecht. The experience of presence: Factor analytic insights. *Presence*, 10(3):266Ű281, 2001.

[35] E. W. See-To, S. Papagiannidis, and V. Cho. User experience on mobile video appreciation: How to engross users and to enhance their enjoyment in watching mobile video clips. *Technological Forecasting and Social Change*, 79(8):1484Ű1494, 2012.

[36] M. Soleymani, J. Lichtenauer, T. Pun, and M. Pantic. A multimodal database for affect recognition and implicit tagging. *Affective Computing, IEEE Transactions on*, 3(1):42Ű55, 2012.

[37] C. Stevens, H. Winskel, C. Howell, L. Vidal, J. Milne-Home, and C. Latimer. Direct and indirect methods for measuring audience reactions to contemporary dance. 2009.

[38] M. Thelwall, K. Buckley, and G. Paltoglou. Sentiment in Twitter Events. *JOURNAL OF THE AMERICAN SOCIETY FOR INFORMATION SCIENCE AND TECHNOLOGY*, 62(2):406Ű418, 2011.

[39] L. Van der Maaten and G. Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(11), 2008.

[40] C. Wang and P. Cesar. Do we react in the same manner?: comparing GSR patterns across scenarios. In *Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational*, 2014.

[41] C. Wang, E. Geelhoed, P. Stenton, and P. S. Cesar Garcia. Sensing A Live Audience. In *Proceedings of ACM CHI Conference on Human Factors in Computing Systems 2014 (CHI 0)*, page 1909Ű1912. ACM, 2014.

[42] W. H. Whyte. *The Social Life Of Small Urban Spaces*. Project for Public Spaces Inc, Jan. 1980.

[43] M. Wirz, D. Roggen, and G. Tröster. A Methodology towards the Detection of Collective Behavior Patterns by Means of Body-Worn Sensors. In *Workshop at the 8th International Conference on Pervasive Computing (Pervasive 2010)*, 2010.

[44] T. Zhang, J. Wang, P. Liu, and J. Hou. Fall detection by embedding an accelerometer in cellphone and using KFD algorithm. *International Journal of Computer Science and Network Security*, 2006.

9. APPENDIX

Post-event Questionnaire (English)

1 . The dance performance was interesting.(*)
2 . The dance performance was exciting.(*)
3 . The dance performance was enjoyable.(*)
4 . I lost track of the world while I was watching the dance performance.(**)
5 . I still paid attention to my surroundings while I was watching the dance performance.(**)
6 . I was completely captivated by the dance performance.(**)
7 . I will definitely want to come to another dance performance again.(***)
8 . I will recommend dance performances to my friends.(***)
9 . Dance performance was worthwhile.(***)
10 . This dance performance uplifted my mood.(****)
11 . This dance performance energized me.(****)
12 . This dance performance made me feel more cheerful.(****)
13 . It was natural for me to wear the sensors during the performance
14 . Did you came with friends or family?
15 . Did you have a favorite moment? If yes, please describe it.

Post-event Questionnaire (Back-translated Dutch)

1. De voorstelling was interessant.(*)
2. De voorstelling was opwindend.(*)
3. De voorstelling was aangenaam.(*)
4. Ik vergat de wereld om me heen gedurende de voorstelling.(**)
5. Ik had gedurende de voorstelling aandacht voor mijn omgeving.(**)
6. Ik was volledig in de ban van de voorstelling.(**)
7. Ik kom zeker terug voor een andere dansvoorstelling.(***)
8. Ik zal dansvoorstelling aan mijn vrienden aanraden.(***)
9. Dansvoorstellingen zijn de moeite waard.(***)
10. Deze dansvoorstelling heeft me opgebeurd.(****)
11. Deze dansvoorstelling heeft me een energetisch gemaakt.(****)
12. Deze dansvoorstelling eeft me blij gemaakt.(****)
13. De sensoren voelden gedurende de voorstelling niet onnatturlijk ann
14. Bent u met wrienden of familie gekomen?
15. Had u een favoriet moment? Zo ja, gelieve dit te omschrijven:

(*)Enjoyment [35], (**)Immersion [34], (***)Recommendation [28], (****)Mood.