
Towards an Adaptive Cultural Heritage Experience Using Physiological Computing

Alex J. Karran
School of Natural Sciences and
Psychology
Liverpool John Moores
University
a.j.karran@ljmu.ac.uk

Stephen H. Fairclough
School of Natural Sciences and
Psychology
Liverpool John Moores
University
S.Fairclough@ljmu.ac.uk

Kiel Gilleade
School of Natural Sciences and
Psychology
Liverpool John Moores
University
gilleade@gmail.com

Abstract

The contemporary heritage institution visitor model is built around passive receivership where content is consumed but not influenced by the visitor. This paper presents work in progress towards an adaptive interface designed to respond to the level of interest of the visitor, in order to deliver a personalised experience within cultural heritage institutions. A subject-dependent experimental approach was taken to record and classify physiological signals using mobile physiological sensors and a machine learning algorithm. The results show a high classification rate using this approach, informing future work for the development of a real-time physiological computing component for use within an adaptive cultural heritage experience.

Author Keywords

Human Computer Interaction; Cultural Heritage; Affective Computing; Machine Learning; Adaptive Systems; Physiological Computing;

ACM Classification Keywords

H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous; H.5.1. Multimedia Information Systems: Artificial, augmented, and virtual realities. H.5.2 User Interfaces: Theory and methods.

Introduction

The research presented here, motivated in part and funded by the ARTSENSE project [1], represents a work in progress towards an adaptive interface designed to respond to the level of interest of the viewer [2] in order to deliver a personalised experience within cultural heritage institutions using technology such as augmented reality systems. The approach we take is the first step in creating a real-time physiological computing component for use within an adaptive cultural heritage experience. The study we present utilised genuine material derived from a cultural heritage institution to elicit psychophysiological responses in an ambulatory setting.

Motivation

In recent years cultural heritage institutions have started to adopt mobile technologies such as smart phones and tablet computers to increase information provision and retain audiences for exhibits or installations [3]. Cultural heritage installations can be considered to be a distinct form of HCI that is closely related to tourism, navigation and education applications.

The contemporary heritage institution visitor model is built around passive receivership where content is consumed but not influenced by the visitor [4]. This passivity is reflected in how mobile technologies are currently used in heritage installations, by offering static content delivery using methods such as quick response codes (QR) and media enhanced guided tours. QR codes involve encoding information about an exhibit into a matrix barcode that is scanned by the device, and then displays the information to the user. Media enhanced guided tours provide navigation paths

through heritage installations, serving audio or multimedia content about exhibits via displays or augmented reality overlays [5]. These technologies can act as standalone experiences or as tools to augment the traditional museum-provided human expert tour. However, these technologies are opaque to the visitor and ignore, potentially, the most important source of data about the cultural heritage experience: the visitors themselves.

Personalising Cultural Heritage

To personalise the cultural heritage experience we propose to use mobile physiological sensors and an interest recognition framework to collect, record and classify indices of psychophysiological variance. The level of interest for the visitor that emerges from this classification can subsequently be used to inform choices of interactivity for users at system level. This choice can take the form of a heritage content “recommender”, which offers content navigation decisions to users based on current or previous levels of interest in presented material.

Classifying psychophysiological states using pattern recognition algorithms is standard practice in the laboratory. This multidisciplinary research area, concerned with affective or physiological computing applications, [6] has been used with varying degrees of success to determine emotional states [7, 8], cognitive workload [9] and physiological activation [10]. The research reported here is concerned solely with the application of the support vector machine (SVM) pattern recognition algorithm to classify two states, representing high and low interest, in an ambulatory setting focused on real-time deployment.

Study: A Virtual Heritage installation

Selection and creation of the classification engine are necessary first steps toward developing a real-time application for use within cultural heritage institutions. To this end, we created a virtual heritage installation that replicated in part, a late 18th century Valencia kitchen mosaic (installed with the Museo Nacional de Artes Decorativas). This allowed participants in the study to stand in a natural fashion, while simultaneously viewing the mosaic and listening to audio narrative about elements of the representation.

The study was designed with a threefold purpose:

1. To measure and classify psychophysiological reactivity in response to cultural heritage content presented as visual and audio stimuli
2. To define the psychophysiological variance as a two condition level of interest (high and low)
3. To evaluate the performance of the SVM classification algorithm for real-time application and the precision of the classifier, when compared to subjective response data

Experimental Task

Participants were asked to view a visual representation of the mosaic and to listen to an audio narration salient to the highlighted parts of the representation, then to answer post-hoc questions relating to interest level per auditory segment.

Methodology

10 subjects 2 male 8 female, aged 19-75 participated in the study. Physiological responses from the autonomic system was measured during experimental sessions, using the Electrocardiogram (ECG) and Skin

Conductance Level (SCL) channels of the Mind Media Nexus X Mk II in lead 2 configuration for ECG and second and forth fingers of the non-dominant hand for SCR. Four channels of electroencephalographic (EEG) data were recorded using the Enobio wireless 4-channel sensor. A Biosemi EEG cap was fitted and inion/nasion aligned to ensure sensor placement (Figure 1). Electroconductive gel was added to sites FP1, FP2, F2 and F3 and electrodes were attached.

Procedure

After receiving instruction about the experimental procedure, participants were asked to complete a consent form in accordance with the Liverpool John Moores Ethical Committees lease of ethical approval, and then fitted with a wearable pouch to hold the nexus sensor hardware at the hip. Electrodes for ECG were placed on the torso. The biosemi sensor cap was fitted and electrodes attached. Participants were asked to stand in a relaxed position approximately 2 meters in front of a 2*3 meter projection screen. This was followed by the audio-visual presentation of the Valencia kitchen. The presentation of the kitchen stimulus was linear and timed to progress through the narrative, giving four stories consisting of 3 factual elements. On completion of the presentation each participant was asked to rate which two stories were perceived to be the most interesting out of the four that were presented.

Analysis

Prior to commencing classification analysis using the physiological data, features were derived* from measures of heart rate, skin conductance and EEG. This resulted in a total of 9 features for each of the 14 stimulus events. These features were further

*Feature Derivatives

- **Activation** : mean heart rate, inter-beat interval and mean skin conductance level
- **Cognition** : Where the ratio : \mathcal{X} is expressed as lognormal β (power) divided by α (power) at sites $fp1$ $fp2$ $f3$ $f4$

$$: x = \ln \left(\frac{y_{\beta}^i}{y_{\alpha}^i} \right)$$

- **Motivation** : Where the ratio : \mathcal{X} is expressed as lognormal of α (power) subtracting right from left hemispheric activity at sites ($fp1,fp2$) and ($f3,f4$)

$$: x = \ln(z_{\alpha}^i - y_{\alpha}^i)$$

- **Normalisation** : standard z score:

$$z = \frac{X_i - \mu}{\sigma}$$

Where X_i is the value to be scored μ is the columnar mean of the dataset and σ the columnar standard deviation.

subdivided into components of a three dimensional interest model. The interest model is comprised of activation, cognition and motivation, such that each feature set created a unique classifier feature vector for each element.

This approach has a number of advantages. Each feature vector is identified as a separate element of the model, feature sets can be combined as a fusion of features, and the effect of each feature set or fusion of features on classifier class recall can be evaluated. Fusion refers to the combination of feature data, into a vector that represents either single or multiple dimensions of the interest model. Table 1 displays the feature sets and subsequent class recall accuracy of the classifier for each fusion of features. Feature sets are denoted by: A (activation); C (cognition); and M (motivation) with r representing raw values.

Each participant's data was analysed separately to determine the recall accuracy of the SVM classifier for individual participant responses. The SVM classifier is a supervised pattern recognition algorithm, requiring an n dimensional vector (observation) and an associated label (class) for training. This training set is then used as the basis for classifying new instances of data into its respective class. We used the SVM implementation by *prtools* [10] within *matlab* 2012Rb. Each feature set was tested using k -fold cross-validation. In k -fold cross-validation, $k-1$ folds are used for training and the last fold is used for evaluation. This process is repeated k times, leaving one different fold for evaluation each time. Furthermore, in order to test the capacity of the classifier to generalise across all participants, the feature data was combined into one dataset and classified twice: firstly, using 5 fold cross-validation and

secondly, by splitting the dataset in two parts selected at random, one for training and one for testing.

Results

Table 1 summarises the results obtained from the subject-dependent classification of the feature data. The feature sets (activation, cognition and motivation) were classified alone and in combination, to determine which permutation of features provided the best class recall accuracy over all participants. Initial testing revealed that no benefit was gained by normalising the data for individual participant classifications. The data table indicates that the fusion of raw activation and cognition features afforded mean class recall accuracy across all participants of 80%, with a minimum 66% and maximum 100% spread over all participants.

Feature(s)	P1	P2	P3	P4	P5	P6	P7	P8	P9	P10	Mean Recall
A^r	0.44	0.72	0.80	0.83	0.90	0.80	0.63	0.89	0.67	0.90	76%
A^r, C^r	1.00	0.70	0.66	0.90	0.77	0.87	0.60	0.80	0.70	1.00	80%
A^r, M^r	0.66	0.62	0.66	1.00	0.90	0.78	0.45	0.88	0.66	1.00	76%
A^r, C^r, M^r	0.81	0.62	0.70	0.90	0.80	0.67	0.60	0.80	0.70	1.00	76%
C^r, M^r	1.00	0.66	0.66	0.50	0.66	0.77	0.80	0.77	0.81	1.00	76%
C^r	0.87	0.90	0.72	0.55	0.50	0.44	0.80	0.70	0.87	1.00	74%
M^r	0.75	0.55	0.60	0.54	0.83	0.60	0.50	0.87	1.00	0.80	70%

Table 1 Subject dependent classifier recall accuracy, with mean recall accuracies. (*Five-fold cross-validation*)

This significant classification rate offers strong evidence that the combination of activation and cognition features affords an effective method from which to ascertain a user's level of interest in a cultural heritage setting. However, examining the individual recall rates in isolation shows that, for some participants, this combination of features resulted in lower class recall accuracies, highlighting the influence of individual differences in physiological responses towards the heritage material.

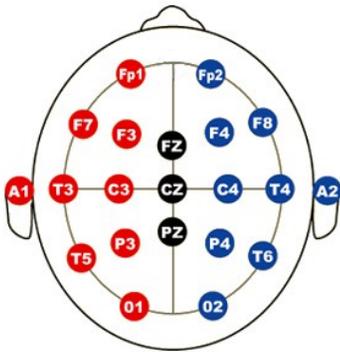


Figure 1 The International 10-20 (EEG) System

When comparing the classifier recall accuracies from other feature sets, it can be seen that the features of activation alone are only 3% less accurate overall than those of the combined activation and cognitive feature sets, with a maximum of 77% mean recall accuracy. However, although this result was promising, there can be seen a greater (negative) intra-subject classification variation in recall accuracies, when compared with the combined feature set. Combining the features of activation with motivation, or cognition with motivation provided no clear benefit to recall accuracies with a maximum recall accuracy of 76% respectively for raw feature data.

Findings from the literature and previously completed work indicate that problems exist with classifier generalisation accuracy due to magnitude differences in psychophysiological responses between individuals. In generalisation tests the classifier reported a steep drop in accuracy in line with these findings. However, combining all three feature sets resulted in a predictive accuracy of 65%, still 15% above that of chance. It was these finding that informed the development of the current subject-dependent classifier approach.

Discussion and Conclusions

The results of this study provide strong evidence, that the combination of activation and cognition features coupled with a subject-dependent classification approach and the SVM classifier can reliably infer the “knowledge emotion” interest within a cultural heritage context. It is interesting to note however, that these same results also highlight the possibility that an approach based on an ensemble of classifiers for each dimension of the interest model may provide even greater recall accuracies compared to a combined

approach. Furthermore, another approach we seek to pursue involves the use of a real-time decision engine consisting of gated threshold logic for each feature set, which would map magnitude variance as it occurs onto the model of interest.

We set an initial accuracy floor of 75% below which interactive systems using the bio-sensing component would become unusable. This figure was determined arbitrarily, thus requiring further research. The results show that, overall, the approach taken can exceed this figure. The next task is to integrate the bio-sensing component into an interactive system and evaluate its performance using receiver operator characteristic techniques and real-time user feedback, to determine the level of acceptable accuracy for users of the system and create a model of adaptivity. This is currently work in progress and a real-time interactive heritage application is in development. We envision many possible applications of this approach within the context of cultural heritage, such as automated or semi-automated recommendation of cultural heritage content informed by real-time psychophysiological assessment (a digital curator) or “interest” profiling involving implicit tagging of heritage material to build up heat maps that use interest as a basis to inform future interactions.

Furthermore, by modifying the psychophysiological measures under investigation the reach of the approach can be extended into areas outside of the cultural heritage context, such as targeted marketing and media or even e-health. We understand that the current state of the art for ambulatory sensor technologies is still somewhat intrusive to individuals. However, early usage statistics from users wearing the sensors appear favorable, possibly due to the extreme

novelty of the technology. Moreover, improvements in ambulatory sensor technology are accelerating, with “tattoo”, nano-scale and microwave emission sensor technologies in development, making the intrusiveness and social stigma of wearing bulky technologies less of an-issue.

Acknowledgements

The ARTSENSE project is funded through the cooperation programme of the Seventh Framework Programme (FP7) of the European Commission.

References

- [1] ARTSeNSE - <http://www.artsense.eu/>
- [2] Silvia, P. J. (2005). What is interesting? Exploring the appraisal structure of interest *Emotion*, 5(1), *Emotion* 89-102.
- [3] Kray C., Baus J., A survey of mobile guides. Workshop HCI in mobile guides at *Mobile HCI*, Udine, Italy, September 2003.
- [4] Serio, M. *Politiche dell'educazione al patrimonio artistico*. Firenze: Giunti, 2004.
- [5] DAMALA, A. (2009) Interaction Design and Evaluation of Mobile Guides for the Museum Visit: A Case Study in Multimedia and Mobile Augmented Reality. Centre d'Etude et de Recherche en *Informatique du CNAM*. Paris, Conservatoire National des Arts et Métiers
- [6] Fairclough, SH. Fundamentals of physiological computing *Interacting with computers* 21 (1), 133-145.
- [7] Picard, R. W. Vyzas, E. & Healey, J. (2001). Toward machine emotional intelligence: analysis of affective physiological state. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(10), 1175-1191.
- [8] Frantzidis, C. A. Bratsas, C. Papadelis, C.L. Konstantinidis, E. Pappas, C. Bamidis, P.D., Toward Emotion Aware Computing: An Integrated Approach Using Multichannel Neurophysiological Recordings and Affective Visual Stimuli *IEEE Transactions on Information Technology in Biomedicine*, vol.14, no.3, pp.589-597, May 2010.
- [9] Wu, D. Courtney, C. G. Lance, B. J. Narayanan, S. S. Dawson, M. E. Oie, K. S. (2010). Optimal Arousal Identification and Classification for Affective Computing Using Physiological Signals: Virtual Reality Stroop Task. *IEEE Transactions on Affective Computing*, 1(2)
- [10] Laine, T.I.; Bauer, K.W.; Lanning, J.W.; Russell, C.A.; Wilson, G.F.; , Selection of input features across subjects for classifying crewmember workload using artificial neural networks. *IEEE Transactions on Systems, Man and Cybernetics, Part A: Systems and Humans*, vol.32, no.6, pp. 691- 704, Nov 2002