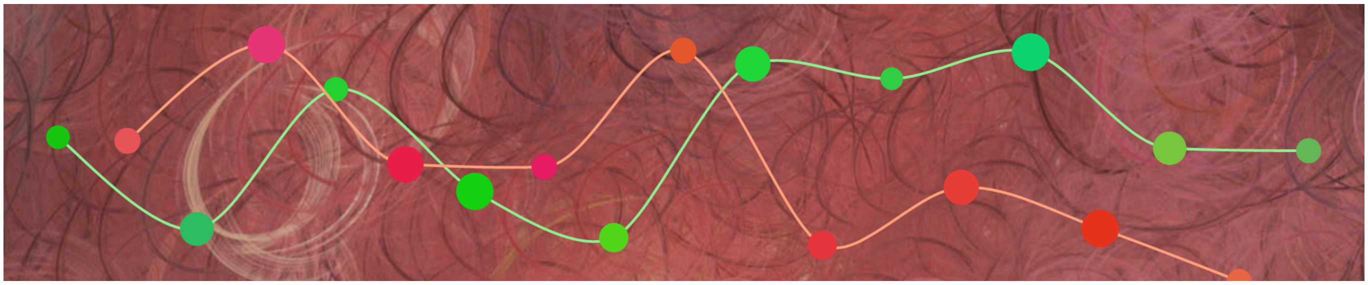


Proceedings of EmoVis 2016

ACM IUI 2016 Workshop on Emotion and Visualization

Sonoma, CA, USA
March 10, 2016

Edited by

Andreas Kerren, Daniel Cernea, and Margit Pohl

Linköping University Electronic Press (LiU E-Press)
Linköping, Sweden

Series: Linköping Electronic Conference Proceedings, No. 103
ISBN: 978-91-7685-817-2
ISSN: 1650-3686 (print)
eISSN: 1650-3740 (online)
URL: http://www.ep.liu.se/ecp/contents.asp?issue=103

Preface

This year, we are happy to announce the proceedings for the first Workshop on Emotion and Visualization (EmoVis 2016) that takes place as part of the ACM Intelligent User Interfaces (IUI 2016) conference in Sonoma, CA, USA. Planned as a bi-annual event, the Workshop on Emotion and Visualization welcomes researchers, practitioners and experts from a variety of scientific domains, including visualization, human-computer interaction, artificial intelligence, cognitive psychology, and multimedia.

EmoVis 2016 acts as a forum where people with diverse backgrounds can present design principles and introduce novel techniques for affect measurement and visualization. The papers accepted at this year's workshop focus on topics like emotion measurement through wearable technologies, sound and emotion, real-time emotion visualization, emotion visualization in different contexts, as well as the challenges faced when detecting and visualizing emotions. All papers in this proceedings book have been peer-reviewed by at least three reviewers from the international program committee consisting of 12 experts listed below.

Many have contributed to make the workshop an enjoyable and enlightening experience. We would like to express our gratitude to the invited speaker, Michelle X. Zhou, and the international program committee for their commitment and reviewing efforts. Without all those great people, this workshop would not have been possible. Finally, we thank the ACM IUI 2016 organization committee for the acceptance of our workshop including local support during the conference and our sponsor for providing financial aid.

Welcome to EmoVis 2016. We hope that you enjoy the workshop and use this event to share your experiences and be inspired.

Andreas Kerren, Daniel Cernea, and Margit Pohl

Conference Co-chairs

Andreas Kerren, Linnaeus University, Sweden
Daniel Cernea, AGT International, Germany
Margit Pohl, Technical University of Vienna, Austria

International Program Committee

Erik Anderson, Electrical Geodesics, Inc., USA
Remco Chang, Tufts University, USA
Christopher Collins, University of Ontario Institute of Technology, Canada
Mary Czerwinski, Microsoft Research, USA
Geoffrey Ellis, University of Konstanz, Germany
Camilla Forsell, Linköping University, Sweden
Liang Gou, IBM Almaden Research Center, USA
Tera Greensmith, AT&T Mobility Advanced Analytics, USA
Kwan-Liu Ma, University of California at Davis, USA
Helen Purchase, University of Glasgow, UK
Jonathan C. Roberts, Bangor University, UK
Wesley Willett, University of Calgary, Canada

Workshop Sponsor

Table of Contents

Invited Talk

Design Studies

Position Statements

Invited Talk

Interactive Visual Analysis of Human Emotions from Text

Michelle X. Zhou
Juji, Inc.
Saratoga, CA, USA

ABSTRACT

People's write-ups, such as online reviews and personal micro-blogs, often reflect their emotions, ranging from just-in-the-moment sentiment to long-lasting mood. In this talk, I will first give an overview on modeling human emotions encapsulated in people's write-ups. I will then sample two visual analytic systems that use very different methods to automatically extract and visualize human emotions from text for two very different purposes. The first is an interactive visual analytic system that automatically summarizes human sentiment captured in online reviews and leverages the power of a crowd to rectify the imperfections in machine sentiment analysis. The second is a timeline-based visual analytic tool that extracts and visualizes a person's moods over time based on the person's tweets. Finally, I will discuss the challenges of inferring human emotional DNA from text in general and potential research directions.

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linköping University Electronic Press.

Biography: Dr. Michelle Zhou is the Co-Founder and CEO of Juji, Inc., a high-tech startup that develops the next-generation of interactive intelligent systems that can deeply understand users and guide their behavior based on their psychological characteristics. Prior to starting Juji, Michelle led the User Systems and Experience Research (USER) group at IBM Research – Almaden and then the IBM Watson Group. Michelle's expertise is in the interdisciplinary areas of intelligent user interaction (IUI), information visualization, and visual analytics. She has published over 80 peer-reviewed, refereed articles and filed about 40 patents in above areas. Michelle has also served on multiple journal editorial boards and numerous technical committees and is currently the Editor-in-Chief of ACM Transactions on Interactive Intelligent Systems. She received a Ph.D. in Computer Science from Columbia University in 1999 and was named an ACM Distinguished Scientist in 2009.

Design Studies

Visualizing the Emotional Journey of a Museum

Shen DU, Edouard SHU, Feifei TONG, Yinghao GE, Lu LI, Jingbo QIU
Ecole Centrale de Lyon
36 Avenue Guy de Collongue,
69134 cully, France
first.name@ec-lyon.fr

Philippe GUILLOTEL, Julien FLEUREAU, Fabien DANIEAU
Technicolor
975, avenue des Champs Blancs, 35576
Cesson-Sevigne, France
first.name@technicolor.com

Daniel MULLER
Ecole Centrale de Lyon
36 Avenue Guy de Collongue,
69134 cully, France
daniel.muller@ec-lyon.fr

ABSTRACT
Wearable devices and new types of sensors make it possible to capture people behavior, activity and, potentially, cognitive state in their daily life. Today those devices are mainly used for well-being applications, by recording and displaying people's activity. Some work have been published going a step further by inferring from the recorded signals the emotional state of individuals or group of people. However, the information provided and the way it is presented are still in their infancy, with time lined graphs showing calories, heart-rate, steps, temperature, and sometimes affective intensity.
In this paper we present an experiment done during the visit of different people in a museum of arts to capture the emotional impact of the exposed paintings. We also propose an associated visualization of their emotional journey. The emotion is here measured as the affective response to the paintings observation, and the processing algorithm is based on an existing technique adapted to the particular case of different observation durations. The visualization is based on a 3D map of the museum with different colors associated to the different paintings to get the emotional heat-map of the museum (more precisely the arousal dimension). The validation has been done in the museum of arts at Lyon, France, with 46 visitors, for a total of 27 paintings, exposed in three different rooms.

Author Keywords
Emotion; Visualization; Physiological responses; Data processing; Museum; Art;

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linköping University Electronic Press.

INTRODUCTION
Evaluating people experience means being able to evaluate engagement, emotions, pleasure, etc. Notions that by definition are difficult to express and measure. Traditionally self-assessment, helped with questions by professionals, is used and further analyzed by psychologists. For instance the Self-Assessment Manikin (SAM) [2] is widely used to measure the pleasure, arousal, and dominance associated with a person affective reaction. It is a non-verbal pictorial assessment technique to ease interpretation and coherence between different subjects. However, there are several bias and limitations that make such methods unusable for real applications where the subjects are experiencing different events in various locations and live conditions.
Recently, wearable devices to measure physiological responses have been made available, with corresponding signal processing and algorithms to infer affective reaction from those signals. The main advantage of this measure is its unconscious and objective nature. Three main signals have been used for emotion detection: the facial expressions [11], the body gestures [13], and some biological reactions from the autonomic nervous system, such as skin conductivity, heart rate, skin temperature [3] or pupillary response and eye blinking. Such methods have been experimented for various entertainment applications: music [5], movies and video [14], [4], [17], or advertisement [10]. However, there is not so much trials done for paintings. The most relevant is probably the eMotion mapping museum experience [16], [15]. Probably because art appreciation addresses deep cognitive information such as religion, culture, education, history, etc.
Besides how to capture emotional reactions, the second issue raised is how to represent them to reflect time (when it happened), space (where was the subject), intensity (how strong the emotion was), type (what was the valence of the emotion), and variability (how does it compare between individuals and different experiences). Interesting attempts have been proposed for well-being applications, especially with smartphones, such as the Moodscope [9], an online platform aimed at tracking ones mood throughout a certain period of time. In general the user is prompted to rank his feelings towards different emotional states, ideally on a daily basis. A simple 2D graph curve indexed by days is used to represent the mood score. MoodJam (`http://moodjam.com/`) adds patterns of colors and words to describe people mood and share it with others. Other similar apps have been proposed linked or not

to social networks, but the representation is still very simple, using only simple colored patterns or scales, representative of people's mood. It also requires the user to manually set his own feelings. Besides the bias of this interpretation, and the difficulty to associate words and colors to an emotion, it may not necessarily represents the truth. More complex representations have been proposed for music experiences by Krcadinac [6] and Kaushal Agrawal in Data Visualization Mood of the Artist (`http://www.kaushalagrawal.com/moodoftheartist.php`). The book Emotional Cartography by Christian Nold [12], provides an unique collection of essays discussing the desire to map emotions, by visualizing intimate biometric data and emotional experiences. It is a very relevant work providing *"a tangible vision of places as a dense multiplicity of personal sensations"*.

In this paper we want to answer two main questions: how to measure the quality of a museum experience? and how to visualize the resulting affective experience? For that we propose a new way to visualize the emotional experience, based on physiological responses induce by pieces of art (paintings in that case). In particular we provide a capture and playback system of the journey, highlighting the emotion intensity (arousal) for each painting. Different analytic visualization methods are also proposed to evaluate individuals, groups, rooms, and paintings specifically. The first part details the affective detection method and how values are extracted from the visitors' body responses. The second part describes the different representations used. Finally the third part provides the results of the experiments done with real visitors of the museum of arts of Lyon, France.

AFFECTIVE DETECTION

The affective responses have been captured using the "galvanic skin response" (GSR), also known as "skin conductance" or "electrodermal activity" (EDA). When an emotion is experienced, the autonomic nervous system activity causes a change in the skin's conductivity as a result of the activity of the sweat glands. This link between the GSR and the people emotional state was shown in various publications [8, 1, 7]. We have re-used this property and developed a dedicated algorithm to infer the emotional state from the GSR by means of the affective responses. This algorithm is detailed in this section.

Acquisition

First the affective responses have been captured using the BodyMedia Armband™wearable sensor depicted in Figure 1 (`http://www.bodymedia.com/armband.html/`). The data rate was set to 32Hz, thus sample rate is $T = 1/32$. To keep only the relevant signal parts, i.e. the one when the visitor is watching a painting, we observed the visitors and annotated the start and end time when he was watching each painting. It allows us to remove noise due to motion between two paintings and to measure the observation duration for each painting. This observation duration is necessarily different for each visitor and each painting. By doing that way, all the observations can be synchronized for each painting and each visitor. Then, the raw data from the sensors are extracted for each observation time slot, the remaining data are removed.

Figure 1. Wearable sensor used. BodyMedia Armband™.

Those raw signals are then processed according to the next paragraph.

Processing

To compute the affective highlights from the raw GSR signal, the algorithm reported by Fleureau et al. [4] was used. However, this algorithm was developed for movie assessment, for which the observation duration is the same for all the observers. In our case we had to adapt the algorithm to this variability of duration, and also to the constraint that we want to compare the different exposition rooms against each other. The number of paintings per room being different, this second variability has also to be taken into account. It leads to the algorithm depicted in Figure 2, with the following successive steps:

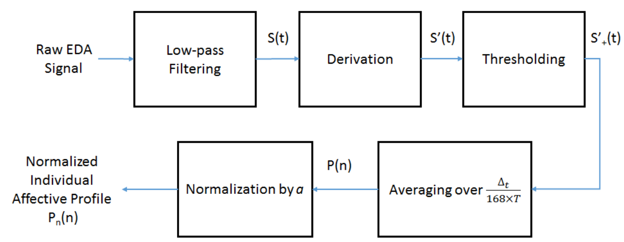


Figure 2. Description of the different steps of the algorithm to compute an Individual Affective Profile from the raw EDA.

- **Low-pass filter** : to remove noise and unnecessary information
- **Derivation** : only the GSR variations are relevant for the affective responses ([8])
- **Truncation with positive data** : to detect the phasic changes consisting in a fast increase until a maximum, followed by a slow return to the initial value
- **Elimination of unnecessary values** : to remove data out of the observation time
- **Averaging** : to handle the different observation duration between visitors
- **Normalization** : to be able to compare between paintings and observers
- **Histogram** : as a visualization of the result

The low pass filter is a FIR filter with a 2Hz cutoff frequency. The filtered values are then derivated and truncated to positive values $x(n)$ in order to highlight the relevant phasic changes according to [7]. After removing data out off the observation time interval, the signal is temporally filtered and subsampled using a variable time window based on the visitor observation duration $\Delta_t(x)$.

Averaging
Since all observers have different ways to observe the paintings, the integration interval time is adapted by fixing the number of observation samples N per room and observer. A reasonable number of observations per painting appeared to be ~ 20, leading to a total number between 140 and 240 per room. The minimum common multiplier for the number of paints per rooms being $7 \times 4 \times 3 \times 2 = 168$, $N = 168 \times T$ provides a sufficient resolution and accuracy. In addition, since the minimum observation time per individual and painting is around 4-5 s (based on our experiments), at least 6 to 8 raw samples are integrated to get an observation. Similarly, the average observation time is around 45-60 s, leading to 72 to 96 integrated raw samples. Therefore it provides a good filtering of the values and also a good accuracy regarding the slow phasic changes of the GSR (i.e. $\sim 2s$). Finally, it gives the same number of observations per viewer and per room, making comparisons easier. It also intrinsically takes into account the observation duration difference between paintings, as an emotional parameter. It can be assumed that the longer one watches a painting, the higher the impact is, compared to the other paintings.
It leads to affective measures $p(n)$ per individual for one room defined as:

$$p(n) = \frac{1}{N_x} \sum_{i=1}^{N_x} x(i+n) \tag{1}$$

where $N_x = \Delta_t(x)/N$ is the integration step, for a total observation duration Δ_t for observer x in a room.

Normalizing
Then, the individual affective profile $p_n(n)$ is obtained after normalization by the individual affective intensity a (computed to get the area under the curve equals to one). We assume there that $p(n)$ is analogous to the probability of an affective response at time n. It also removes the user-dependent part related to the amplitude of the GSR derivative which may vary from one subject to another.
The mean affective response $\overline{p_n}$ of an individual is given by the average value of the $p_n(n)$ values:

$$\overline{p_n} = \frac{1}{N} \sum_{i=1}^{N} p_n(i) \tag{2}$$

This whole process is illustrated in Figure 3 and Figure 4, for room "Religion" and 7 paintings. The raw signals $x(n)$ is illustrated in Figure 3, and the corresponding normalized individual affective profile $p_n(n)$ in Figure 4.

VISUALIZATION
The visualization is twofold, first a 3D model of the museum rooms is designed to map the emotional profiles (affective highlights) with the museum layout. Second the emotional profiles are represented by a color histogram associated to the paintings, and a wall color representing the average emotional intensity of the considered painting.

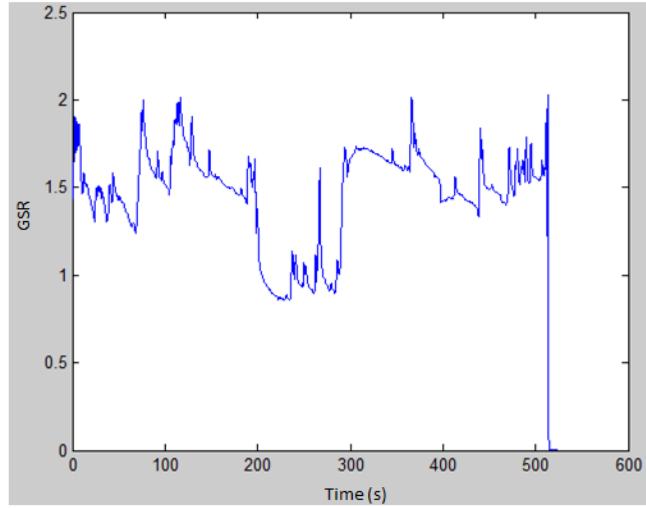


Figure 3. Captured individual raw GSR signal during the visit of one visitor for the 7 paintings of room "Religion".

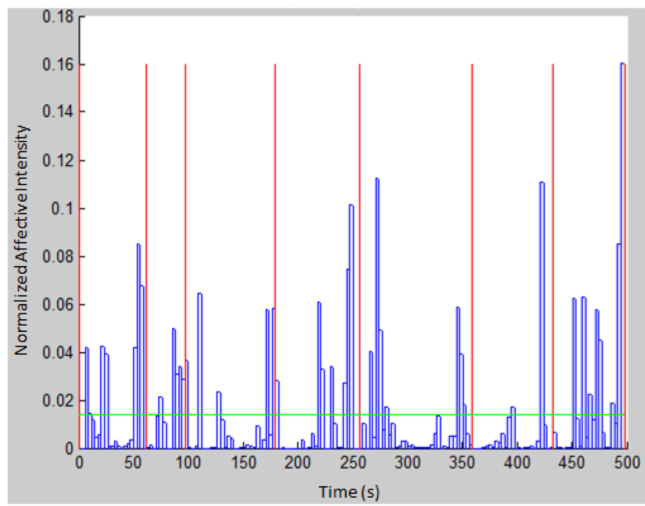


Figure 4. Individual affective response for the 7 paintings of room "Religion". Red lines are the separation between the different paintings, the green line is the average response $\overline{p_n}$ for this visitor.

Museum 3D model
The museum rooms have been synthesized using a 3D design software (Trimble SketchUp, `https://www.sketchup.com/fr`). An point of view is provided in Figure 5. All the paintings are exposed as in the real museum. All the proportions are kept consistent with the originals.

Paintings emotional profiles
To represent the emotional profiles we have used an histogram with colors corresponding to a heat-map, where blue means low emotional intensity, and increasing intensity levels for the others colors, up to the red. Each bar corresponds to one visitor.

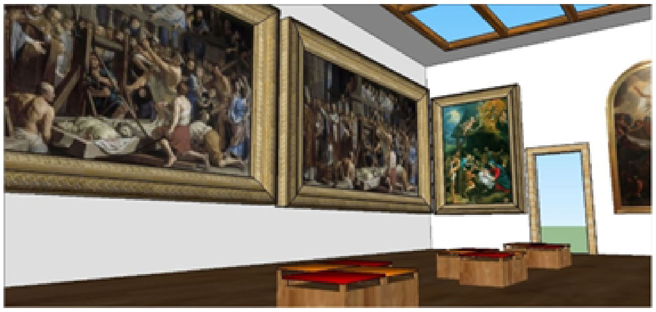


Figure 5. Example of the 3D model for room "Impressionists".

Museum heat-maps

The same color convention is used to color the wall and ground of the museum so as to directly identify the higher emotional areas. However the color is selected as the average intensity level over all observers (see Figure 11).

EXPERIMENTS

The previous algorithm has been used to process the data with a total of 46 visitors for 3 rooms: "Religion", "Impressionists" (aka Monet), and "Abstractionists" (aka Picasso). The capture has been done during several days at the Lyon museum of arts, taking volunteers at the entrance and following them to index the different start/end time at each painting. Each visitor has been tracked only in one room. The total number of paintings and observers per room is the following:

Room	Religion	Impressionists	Abstractionists
# Paintings	7	12	8
# Visitors	21	17	8

Table 1. Number of paintings and visitors per room.

Once captured, all the raw GSR signals have been uploaded and processed to compute the following different results and analytics.

Per individuals

The first analytic measure computed is the emotional profile of an individual in one room, i.e. how each participant responded in front of each painting. It is based on the average emotional response per painting computed as described in Section *"Processing"*. The corresponding values are given in Figure 6 and Figure 7 for two different visitors (visitors 7 and 17 respectively) when watching the 7 paintings of the room "Religion". Each bar corresponds to one painting and the amplitude is the affective intensity level. One can note the significant difference between the two observers. Visitor 7 is mostly sensitive to the last three paintings (painting 5 to 7), while visitor 17 is more responsive to paintings 3, 4 and 6. It means that each of them is impacted differently by the paintings depending of his history, education, feelings, belief, etc. However, thanks to these individual emotional profiles, we can compare the differences between rooms and paintings over a population, and even between observers.

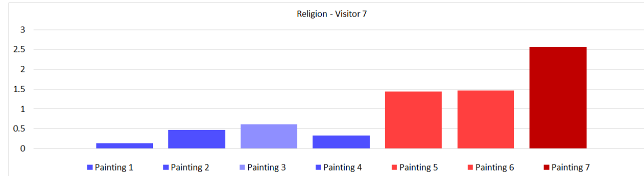


Figure 6. Resulting individual affective response of visitor 7, for the 7 paintings in room "Religion".

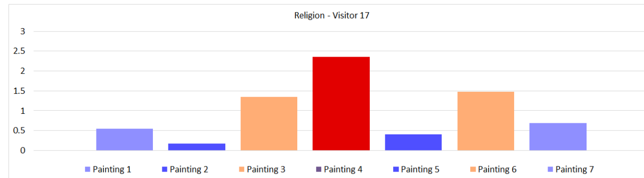


Figure 7. Resulting individual affective response of visitor 17, for the 7 paintings in room "Religion".

Per paintings

In addition to the individuals' responses, it would be interesting to compute the paintings profiles, i.e. what is the impact of a painting on visitors. It is computed as the average responses of the different observers for each painting. Figure 8 shows an example for one painting and 21 visitors. Once again the individuals variability is clearly visible, with visitors 4, 6, 7, 8, 10, 15, 21 as the most responsive, and visitors 2, 3, 5, 11, 12, 20 the less. Then, with these profiles it is possible to compare the different paintings. Figure 9 is another painting with the responses of the same 21 visitors. The second painting (painting 6) is clearly more impacting visitors

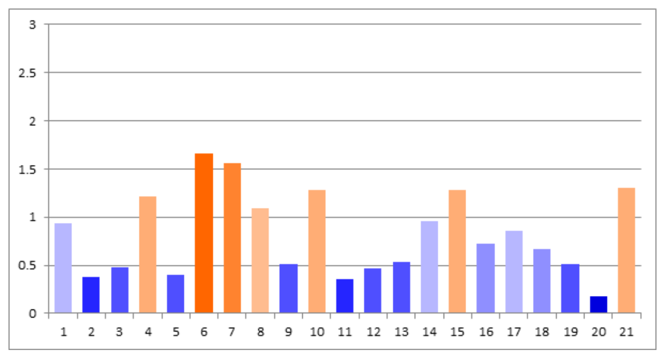


Figure 8. Resulting affective profile of painting 2, room "Religion".

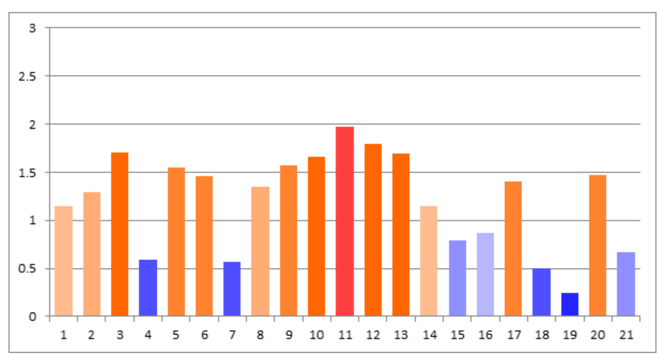


Figure 9. Resulting affective profile of painting 6, room "Religion".

than the first one (painting 2), with average intensity levels higher.

To have an idea of the corresponding paintings content the reader should refer to Figure 10 where the corresponding pictures are illustrated.

From these profiles, it is then possible to define a ranking for each painting. The 3 most higher mean affective responses $\overline{p_n}$ per painting and visitor are selected. Then the number of occurrences of each painting in this top 3 is computed to serve as ranking. This method has been preferred to the direct mean affective responses $\overline{p_n}$ to be able to take into account the high variability between observers. We assume that the three highest responses are the most significant whatever the differences. This ranking is given in Tables 2, 3, 4 for each room respectively under line *# Occurrences*. $\overline{p_n}$ is also included for comparison. *# Occurrences* is thus the number of visitors with this painting in their top 3 higher mean affective responses. When comparing Figure 8 and 9, we observe that painting 2 is lower than painting 6 on average. This is also reflected by the ranking where painting 2 is ranked 5 times in the top 3, while painting 6 is ranked 12 times in top 3 (the higher, the better). This is also confirmed by the $\overline{p_n}$ values with respectively 0.83 and 1.15. This is not always the case (see for instance painting 8 in room "Abstractionists") because top 3 may be more representative of *"interesting"* paintings that the direct mean affective arousal.

Room Religion							
# Painting	1	2	3	4	5	6	7
# Occurrences	4	5	11	5	6	12	7
$\overline{p_n}$	0.92	0.83	1.10	0.99	0.97	1.15	1.03

Table 2. Ranking for room "Religion" by means of the number of occurrences in top 3 highest responses and average affective intensity for all visitors.

Room Impressionists												
# Painting	1	2	3	4	5	6	7	8	9	10	11	12
# Occurrences	5	4	5	4	3	1	2	6	3	5	3	7
$\overline{p_n}$	0.99	0.87	0.88	0.99	0.81	0.89	0.90	1.22	0.812	1.26	1.04	1.33

Table 3. Ranking for room "Impressionists" by means of the number of occurrences in top 3 highest responses and average affective intensity for all visitors.

Room Abstractionists.								
# Painting	1	2	3	4	5	6	7	8
# Occurrences	3	1	2	2	2	2	1	1
$\overline{p_n}$	1.32	0.81	0.78	0.97	0.89	1.07	0.64	1.52

Table 4. Ranking for room "Abstractionists" by means of the number of occurrences in top 3 highest responses and average affective intensity for all visitors.

Once the *affective profile* of each painting computed it can be displayed on the 3D graphic model of the museum. We decided to add it bellow the paintings as described in Figure 10. On the same pictures, we also changed the color of the wall where the painting is displayed to reflect the mean $\overline{p_n}$ value over all visitors for each painting. The heat-map correspondence is used (blue is low, orange/yellow is medium, red is high). Therefore, when looking to a painting, the visitor directly access the average affective intensity and to a more detailed affective profile.

Figure 10. Emotional profiles visualization for painting 2 (left) and 6 (right), room "Religion".

Per rooms

Another way to represent the collected data is to compare the different rooms to each other. The first representation format used is the direct display of the visitors affective responses for each painting, and each room. It gives Figure 12, Figure 13 and Figure 14 for respectively room "Religion", room "Impressionists" and room "Abstractionists".

	Variance
Room Religion	0.11
Room Impressionists	0.18
Room Abstractionists	0.30

Table 5. Variance of the affective values for all paintings and visitors in each room.

It allows a direct and complete comparison. For instance, room "Impressionists" is the one with the highest responses, and room "Abstractionists" the room with the highest variability. This is confirmed with the variance computed on all the paintings per room in Table 5.

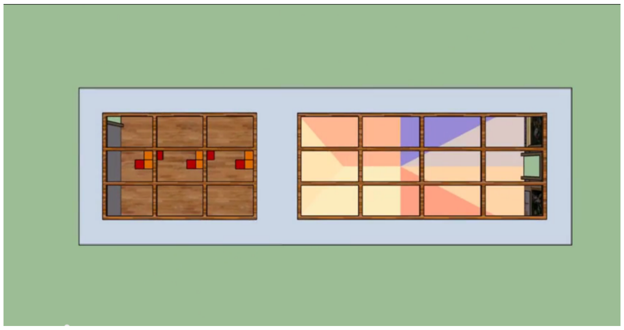
The final representation depicted in Figure 11, uses the previous color mapping on the 3D model and extend it to the floor. Then, a global view of the museum rooms is easy, with the paintings "emotion profiles" depicted bellow the paintings, and the wall and ground colorized according to the painting mean emotional value. In addition, navigation is possible within the 3 rooms of the museum, to explore the paintings, and have an overview of the paintings with higher emotion intensity, lower intensity or higher variability. It would be particularly useful for the visitors to prepare their visit.

DISCUSSION

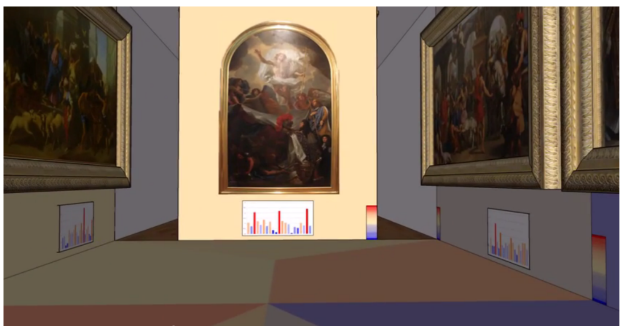
Based on the experiments and various discussions with the visitors, the results on the different rooms can be described as follows.

Room "Religion"

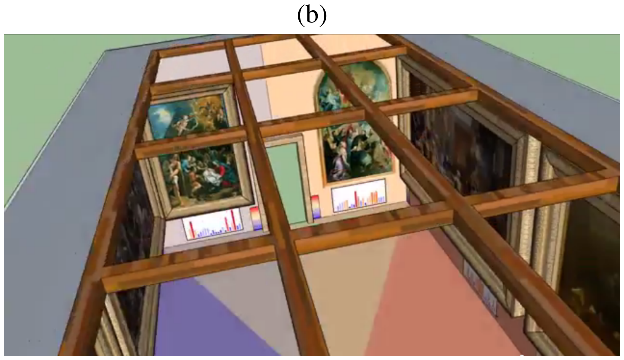
This room shows 7 religious paintings about the Christianity. Their size is larger than the other rooms, as well as their colorfulness. From our observations, it leads visitors to spend more time watching those paintings. In addition, a large part of the visitors are not Christians but rather from Asian culture which could also explain the rather low average responses. This is increased by the fact that all paintings are very similar in color and atmosphere, factors that usually may stimulate different reactions. Therefore, besides the richness of these magnificent paintings, the visitors responses are not very strong nor very weak, and relatively homogeneous between paintings as depicted in Figure 12, and confirmed by the variance in Table 5.



(a)



(b)



(c)

Figure 11. Screenshots of the final visualisation in the 3D model of the museum. (a) top view showing the ground color in front off each painting with the color representing the average intensity level. (b) inside front view of one painting with the affective profile bellow the painting and both the wall and ground color representing the average intensity level. (c) intermediate viewing angle.

Figure 15. Example of paintings. Left: room "Impressionists" *"Stormy Sea in Étretat"* by Claude Monet, 1883. Right: room "Religion" *"Le Repas chez Simon Le Pharisien"* by Jean-Baptiste Jouvenet, 1706.

An example of the paintings of this room is given in Figure 15, with the *"Le Repas chez Simon Le Pharisien"*,1706, by French painter Jean-Baptiste Jouvenet. It corresponds to painting number 6 in Figure 12.

Room "Impressionists"

Despite the highest number of paintings (12), this room is actually the smallest one. The different paintings are very small and similar to each other. Claude Monet is one of the exposed artist, and the paintings depict mainly landscapes and portraits. The mean observation duration per painting is lower in this room than in the two other rooms. It can be explained by the small size, the close distribution and relatively simple content. However we can extract three different behaviors, i) paintings 1 to 5 are similar, consisting of landscapes with bright colors, generally peaceful, ii) paintings 6 to 8 are dark, showing sad people, and iii) the remaining paintings with dense pictures colorful and textured. It may explain the high responses in paintings 7 and 8, and the relatively higher mean responses for paintings 10 to 12 in Figure 13. These different types of content also explain the higher variability shown in Table 5.
An example of the paintings of this room is given in Figure 15, with the *"Stormy Sea in Étretat"* an 1883 painting by founder of French impressionist Claude Monet. It corresponds to painting number 4 in Figure 13.

Room "Abstractionists"

This room exhibits abstractionist artworks that really encourage the imagination of visitors, even those without artistic expertise. Pablo Picasso is one of the exposed painter, and is the one that exhibits the higher responses (paintings 1 and 8). The content and style of the 8 paintings in this room are very different to each other. It explains the higher variability in the resulting affective values in Table 5. Not surprising, some paintings are difficult to understand (for instance paintings 3 to 5), since abstract work often requires background knowledge on the artist and his work. Anyway, in this room, there is a lot of elements that lead to a strong emotion (see Figure 14). However, it should be pointed out that as this room was a temporary exhibition, we have not been able to capture as many visitors as for the other two rooms.

CONCLUSION

We conducted an innovative experiment to visualize the journey of a person visiting a museum of arts. Based on the observer physiological responses (the GSR here) we computed the individual and average affective responses and provide the *"emotion profile"* of a painting, as well as the *"emotional map"* of the museum. The main advantage of this representation compared to state of the art, is that we do not try to interpret people reactions, which is very complicated and challenging. It implies psychology, art, education, culture, etc. Instead, we prefer to rely on people's nervous system response, which provides unconscious and objective responses. We observed a very high variability between visitors, but taking a population is more interesting and allows to compute the museum map. In addition, the proposed representation is a first step towards more advanced solutions and possible new ways to visit a museum (more interactive and emotionally selective). Further analysis with art experts and psychologists would be helpful to go further. In addition, comparisons of different museums would also be an interesting extension of this work. We also expect that this new way to represent a museum content or experience could be useful to define the museum strategy of placement, communicate on the expositions, as well as be an additional helpful information for the potential visitors.

ACKNOWLEDGMENTS

We thank the Lyon Museum of Fine Arts "Musée des Beaux-Arts" for their contribution to this work (`http://www.mba-lyon.fr/`), as well as the Ecole Centrale de Lyon (`www.ec-lyon.fr/`) for their support during this study.

REFERENCES

1. Boucsein, W. *Electrodermal Activity (Second Edition)*. Springer, 2012.
2. Bradley, M. M., and Lang, P. J. Measuring emotion: The self-assessment manikin and the semantic differential. *Journal of Behavior Therapy and Experimental Psychiatry 25* (1994), 49–59.
3. Fleureau, J., Guillotel, P., and Huynh-Thu, Q. Physiological-based affect event detector for entertainment video applications. *IEEE Trans. on Affective Computing 99*, PrePrints (2012).
4. Fleureau, J., Guillotel, P., and Orlac, I. Affective benchmarking of movies based on the physiological responses of a real audience. In *Proc. Affective Computing and Intelligent Interaction (ACII 2013)*, IEEE (2013), 73–78.
5. Koelstra, S., Mühl, C., Soleymani, M., Lee, J.-S., Yazdani, A., Ebrahimi, T., Pun, T., Nijholt, A., and Patras, I. Deap: A database for emotion analysis ;using physiological signals. *IEEE Trans. on Affective Computing 3*, 1 (2012), 18–31.
6. Krcadinac, U., Jovanovic, J., Devedzic, V., and Pasquier, P. Textual affect communication and evocation using abstract generative visuals. *IEEE Trans. on Human-Machine Systems to appear* (2015).
7. Lajante, M., Droulers, O., Dondaine, T., and Amarantini, D. Opening the black box of electrodermal activity in consumer neuroscience research. *Journal of Neuroscience, Psychology, and Economics 5*, 4 (2012), 238–249.
8. Lang, P. The emotion probe. *American Psychologist 50*, 5 (1995), 372–385.
9. LiKamWa, R., Liu, Y., Lane, N. D., and Zhong, L. Moodscope: Building a mood sensor from smartphone usage patterns. In *Proc. 11th International Conference on Mobile Systems, Applications, and Services (MobiSys'13*, ACM (2013), 389–402.
10. McDuff, D., El Kaliouby, R., Cohn, J. F., and Picard, R. Predicting ad liking and purchase intent: Large-scale analysis of facial responses to ads. *IEEE Trans. on Affective Computing 11*, 4 (2012).
11. Pantic, M., and Rothkrantz, L. Automatic Analysis of Facial Expressions: The State of the Art. *IEEE PAMI 22*, 12 (2000), 1424–1445.
12. Raqs Media Collective, van de Drift, M., Davis, S. B., van Kranenburg, R., Hope, S., and Stafford, T. *Emotional Cartography - Technologies of the Self*. Nold, Christian, 2009.
13. Saha, S., Datta, S., Konar, A., and Janarthanan, R. A study on emotion recognition from body gestures using kinect sensor. In *Proc. International Conference on Communications and Signal Processing (ICCSP 2014)*, IEEE (2014), 56–60.
14. Soleymani, M., Pantic, M., and Pun, T. Multi-modal emotion recognition in response to videos. *IEEE Trans. on Affective Computing 3*, 2 (2011), 211–223.
15. Tröndle, M., and Tschacher, W. The physiology of phenomenology: The effects of artworks. *Journal of Empirical Studies of the Arts 30*, 1 (2012), 79–117.
16. Tschacher, W., Kirchberg, V., van den Berg, K., Greenwood, S., Wintzerith, S., and Tröndle, M. Physiological correlates of aesthetic perception of artworks in a museum. *Psychology of Aesthetics, Creativity, and the Arts 6*, 1 (2012), 96–103.
17. Wang, S., and Ji, Q. Video affective content analysis: A survey of state-of-the-art methods. *IEEE Trans. on Affective Computing 6*, 4 (2015), 410–430.

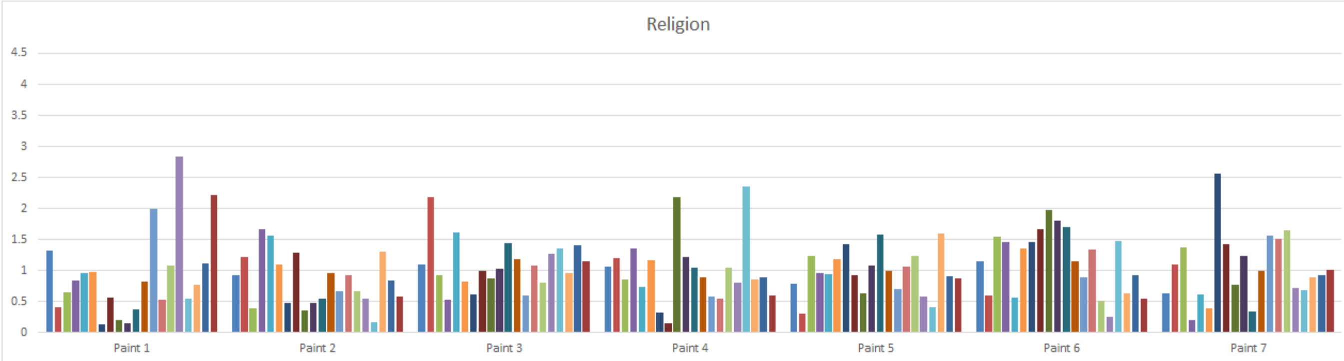


Figure 12. Emotional profiles of all 7 paintings in room "Religion" for the 21 visitors.

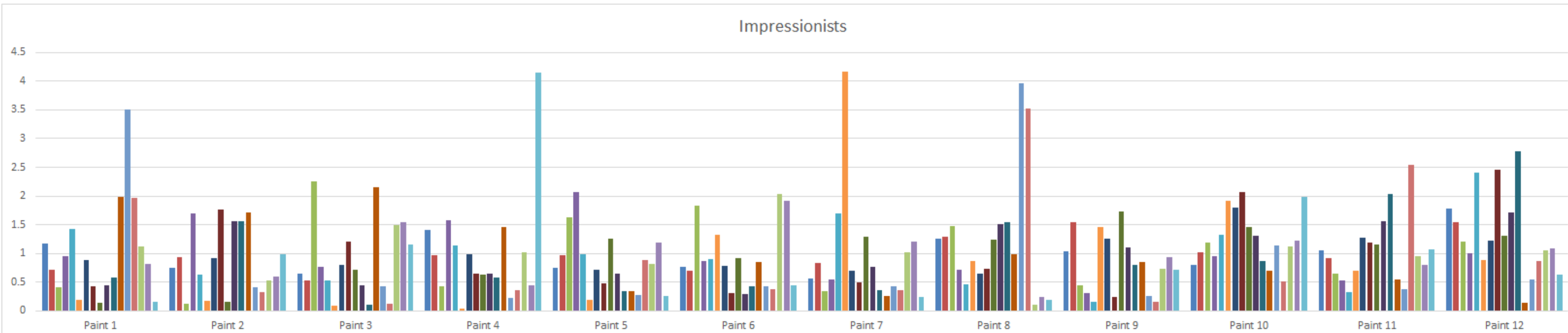


Figure 13. Emotional profiles of all 12 paintings in room "Impressionists" for the 17 visitors.

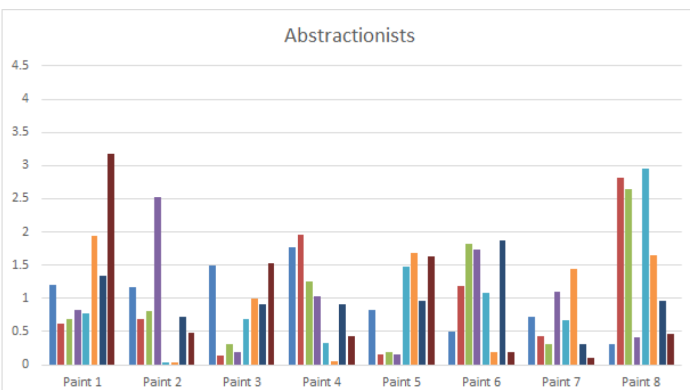


Figure 14. Emotional profiles of all 8 paintings in room "Abstractionists" for the 8 visitors.

Representation of the affective profiles for all visitors, all paintings and the three rooms respectively. Each bar corresponds to one visitor affective response, and the same color is used for one visitor between each painting in a room.

Visualizing Excitement of Individuals and Groups

Kostiantyn Kucher
Linnaeus University
Växjo, Sweden
kostiantyn.kucher@lnu.se

Daniel Cernea
AGT International
Darmstadt, Germany
dcernea@agtinternational.com

Andreas Kerren
Linnaeus University
Växjo, Sweden
andreas.kerren@lnu.se

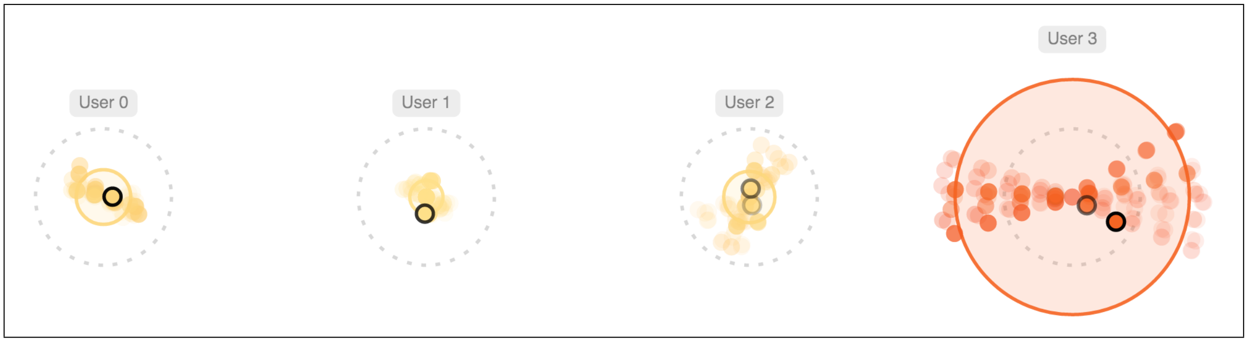


Figure 1: Overview of our visualization applied to a synthetic data set: each circular glyph corresponds to a single user, its concentric rings encode the baseline and current excitement values, and the nested dots provide the detailed information about time-varying measurement values. Here, the oscillating trail encoding is used for the nested dots.

ABSTRACT

Excitement or arousal is one of the main emotional dimensions that affects our lives on a daily basis. We win a tennis match, watch a great movie, get into an argument with a colleague—all of these are instances when most of us experience excitement, yet we do not pay much attention to it. Today, there are few systems that capture our excitement levels and even fewer that actually promote awareness of our most exciting moments. In this paper, we propose a visualization concept for representing individual and group-level excitement for emotional self-awareness and group-level awareness. The data used for the visualization is obtained from smart wristbands worn by each of the users. The visualization uses animated glyphs to generate a real-time representation for each individual's excitement levels. We introduce two types of encodings for these glyphs: one focusing on capturing both the current excitement and the excitement history, as well as another focusing only on real-time values and previous peaks. The excitement levels are computed based on measurements of the user's galvanic skin response and accelerometer data from the wristbands, allowing for a classification of the excitement levels into experienced (excitement without physical manifestation) and manifested excitement. A dynamic clustering of the individual glyphs supports the scalability of our visualization, while at the same time offering an overview of the group-level excitement and its distribution. The results of a preliminary evaluation suggest that the visualization allows users to intuitively and accurately perceive both individual and group-level excitement.

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linköping University Electronic Press.

Author Keywords

Excitement visualization; emotion visualization; group excitement; personal visualization; galvanic skin response.

ACM Classification Keywords

H.5.2 Information Interfaces and Presentation (e.g. HCI): User Interfaces; K.8.1 Personal Computing: Application Packages

INTRODUCTION

In every day of our lives, irrespective of the environment and activities we are involved in, we experience a multitude of different emotions that both affect and even guide our decision making. It comes then as no surprise that emotional theory and the measurement of human emotion has slipped further into focus in research and industry as a means to increase emotional awareness and devise emotion-adaptive computing systems.

In this context, user excitement or arousal is one of the main attributes used to capture emotional variation in various dimensional models [24, 28] from emotion theory. Throughout our daily lives, we experience varying levels of excitement in a multitude of settings: calmness during our commute to work, stress during a business meeting, excitement while watching sports, etc. Still, there are few solutions that try to capture and represent user excitement in the context of emotional self-awareness and group-level awareness. Imagine if you could employ a visualization system in order to explore your most stressful moments of the day, in order to improve your habits and remove unwanted stimuli. Similarly, in terms of group-level awareness, consider as an example a visualization that shows your and your friends' real-time excitement levels while playing a multi-player strategy game.

In this paper, we introduce a novel concept for visualizing individual and group-level excitement in real-time based on the information streamed from smart wristbands worn by each member of the group. The data is gathered from a set of Microsoft Band 2 wearables [25] that periodically measure the galvanic skin response (GSR) and accelerometer (ACC) values for each user. We process this data and compute the levels of experienced excitement (excitement that is detected through the GSR measurements but is not manifested physically) and manifested excitement (high GSR levels coupled with sudden motion of the arms, e.g., cheering at a soccer game).

The visualization is based on animated glyphs that encode each person's excitement levels and allow them to perceive the overall excitement of the group. Figure 1 presents the excitement levels for four individuals. We propose two visual encodings for the glyph: on the one hand, focusing on the representation of the current and previous levels of excitement for each individual, while on the other hand, with scalability in mind, employing animation to better convey the current level of excitement for each person. A dynamic layout enables glyphs with similar excitement levels to be grouped in order to better perceive the distribution of the excitement levels in large groups.

In the following sections, we will initially discuss related research in terms of emotional theory, emotion measurement, and personal visualization. We continue by highlighting the data acquisition and processing for the excitement levels of the individuals. Next, we detail the visualization design and its implementation, followed by a preliminary evaluation of our concept. We conclude by presenting our final remarks and future steps.

RELATED WORK

Nowadays, there is a set of different techniques that are being used for estimating user emotions, ranging from classification of facial expressions and voice tonality to neurological activity and other physiological measurements. While these techniques all have their various estimation capabilities in the context of emotional theory, when focusing on the dimension of arousal some of the most accurate predictors are based on biosignals such as heart rate, respiration, electrodermal activity, body temperature, pupil size variation, all affected by the autonomic nervous system (ANS) [31].

Throughout the decades, electrodermal activity (EDA) or galvanic skin response (GSR) has shown promising results in terms of correlation with user excitement levels [26, 27, 32]. In this technique, two electrode measure the skin conductivity of a user. The conductivity is affected by the moisture that is generated by the eccrine sweat glands, which can be found in our hands and feet, and are mostly responsive to changes in emotional arousal. The activation of these sweat glands is not under the conscious control of a person and can be tied especially to a heightened level of arousal [35] and cognitive workload [30].

Recently, smart wristbands were introduced on the market featuring GSR sensors (e.g., Microsoft Band 2 [25], Empatica E4 [13]), enabling a new dimension of personal data collection and interpretation. Equipped with additional sensors like accelerometers, these wearables have the potential to bring emotion estimation to the real world by considering the context and the activity of the users when processing their physiological signals [17].

While many techniques based on physiological measures have already been used to estimate excitement levels in individuals, there are few systems that focused on raising awareness and visualizing these emotional states. Part of the area of personal visualization [18, 33], emotion visualization is one of the approaches used to inspect and interpret information about user and group affective states.

McDuff et al. [23] introduced AffectAura, a visualization for estimated emotional states for a single user over time using data from multiple sensors. Their work uses an aggregated visual representation for representing valence, arousal and engagement levels for a longer time span of one day. In contrast, our approach is focused on the real-time visualization of a single attribute (excitement) for short time spans, and supports measurements from multiple users simultaneously.

In terms of real-time visualization of emotions, Cernea et al. [7] introduced Emotion-prints, a visualization system for touch-enabled interfaces, where the current level of a user's valence and arousal (estimated through EEG measurements) is represented in the shape of a animated halo around the virtual objects that the user touches. Saari et al. [29] introduced a mobile emotion visualization system devised for improving group performance and awareness. One step further, the GAT visualization [5] enabled the representation of consistent affective experiences at a group level in real-time, thus enabling emotional awareness in teams and improving collaboration. However, emotions have been visualized through a wide range of representations, ranging from expression-inspired approaches—avatars and icons [16, 19, 22]—and to abstract interface integrations [6, 16].

In the context of group psychology, emotional awareness has been linked to decision making [15] and communication [8, 12], while emotional engagement (including excitement) has been shown to be closely linked to gamification [21].

Visualization of real-time measurements collected from multiple wristband devices presents additional challenges and can be interpreted as *streaming data visualization*. One of the earliest papers on this topic is a work by Wong et al. [36] who defined the corresponding visualization challenges: the data may arrive as an unpredictable and unbounded stream without any clear patterns. The authors proposed several techniques to visualize streaming data with scatterplots using variations of multidimensional scaling (MDS) [4]. Alsakran et al. [2] introduced *STREAMIT*, a scalable visualization system that uses a force-direct dynamic system to lay out visual items representing text documents. Huron et al. [20] discussed *Visual Sedimentation*, a powerful metaphor that uses animation to represent aging and aggregation of incoming data entries. In contrast to these papers, our work does not focus specifically on the streaming nature of data (i.e., it is equally applicable to non-streaming data sets), but takes this aspect into account when providing a personal visualization.

Liu et al. [22] described an emotion recognition and visualization technique based on EEG measurements. They supported processing of real-time data at 128 Hz with a sliding window algorithm and used a 3D avatar for encoding of emotions for a single user. In the case of our approach, data measurements are made at a much lower frequency and the supported visual encodings are much more compact. Furthermore, our work supports visualization of excitement for multiple users simultaneously.

Finally, Cottam et al. [10] proposed a taxonomy for streaming data visualization techniques based on dynamics of spatial and retinal variables. They introduced design guidelines based on the comparison task, stability of the displayed items set, and effects of time-varying data attributes. In our case, a single design option based on their guidelines is not sufficient since we want to provide a visualization of streaming data items grouped by user—i.e., there are two representation levels—with support for fixed and dynamic layouts at the inter-user level and two different visual encodings at the intra-user level.

DATA

We have used Microsoft Band 2 devices [25] to record galvanic skin response (GSR) and accelerometer (ACC) values for several people engaged in common activities. The activities involved three to four individuals and included: watching a movie, an opera, and going to a museum. During these activities, each person was wearing a device that would provide GSR and ACC measurements at approximately 0.2 and 9–11 Hz, respectively. This data from each Band 2 was streamed in real-time to an Android mobile phone through a bluetooth connection, from where it was forwarded and stored in a secured database. Currently, our visualization system can access only text-based exports from this database, with future version being able to obtain information directly from the server.

For our purposes, ACC values were sampled to synchronize them with GSR values. For the GSR values, a normalization was applied based on the extracted normal and stressed states, as described in [3]. These states have been extracted either offline, or based on a calibration stage where the intervals for the normal and stressed states were estimated.

Furthermore, the GSR values can fluctuate due to user physical activity or—as perceived in our measurements—a tight fit of the devices on the wrist leading to excessive sweating. To compensate for these events, the ACC values from the mobile device were compared with the ones from the Band 2s. If the user was exerting physical activity as detected both by the mobile device and the wristband, an increased level of excitement was not reported. Moreover, as slow constant sweating could increase the GSR values over time, a sliding window was applied that enables the detection of local increases in the GSR values, thus capturing acute states of excitement.

The resulting data sets comprised timestamped recordings of two variables, normalized in range [0; 1] for a correspond-

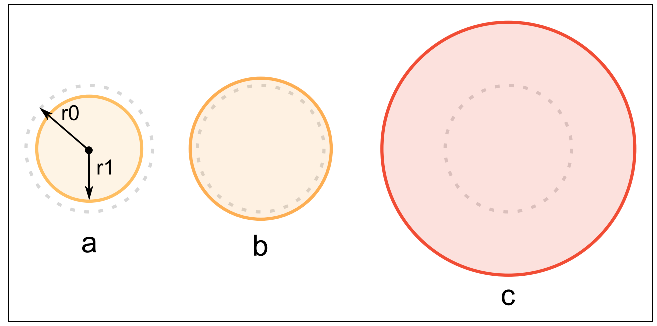


Figure 2: Design of a glyph: (a) a case with rather small GSR and ACC values and highlighted radii; (b) a case with a large GSR and small ACC value; (c) a case with both large GSR and ACC values.

ing user, that are expected to be emitted every 5 seconds for sessions ranging from 20 minutes to several hours. However, there is a degree of uncertainty caused by measurement delays and errors (e.g., the wearables can lose the contact with the user's arm during a measurement session) as well as potential networking delays when transferring data to the visualization front-end (which is less of a concern for pre-recorded data). These issues were relevant factors for our visualization design, as the visualization should support streaming data in order to enable real-time measurements.

Additionally, several synthetic data sets were created for development and demonstration purposes that use a random walk model for each user's GSR and ACC values. Such data sets are a viable option to test the approach with a larger number of users (for instance, 10–15), as the real devices can be unavailable in such quantity.

VISUALIZATION TASKS AND DESIGN

We have had the following visualization tasks in mind for our work:

- monitoring of current excitement values for each user in a group,
- temporal overview of excitement values (for relatively short periods of time) for each user, and
- clustering based on excitement value similarity between users.

Our overall visualization design has two levels: the intra-user level for visualizing measurements for an individual user, and the inter-user level for visualizing measurements for a group of users simultaneously.

We represent the data for each user as a circular glyph, thus enabling a scalable visualization that can capture the excitement levels for small to medium sized groups of people. Each glyph is presented with two groups of elements: concentric circles (rings) and trails. Concentric circles displayed in Figure 2(a) encode the baseline GSR value boundary (dashed circle with radius *r0*) and the combination of GSR and ACC values (shaded circle with radius *r1*). The motivation for this encoding was the assumption that GSR values correspond to

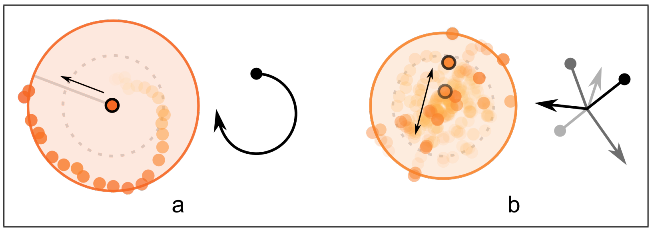


Figure 3: Design of trails nested in a glyph: (a) a clock-style trail with angle increasing over time; (b) an oscillating trail with angle changing randomly. In both encodings, the target distance from the glyph center encodes the combined GSR + ACC values.

experienced excitement levels, while ACC values correspond to *manifested* excitement levels (e.g., the behaviour of a person clapping hands while cheering for a soccer team would be easily observable this way). ACC values are on average larger than the corresponding GSR values in our recorded data sets. Therefore, we use the GSR value as the base, but if it reaches a certain threshold (set to 0.8 by default), the ACC value is added to it to produce the combined excitement level. For example, the radius of the shaded circle that barely exceeds the baseline value would signal that a user is excited based mostly on GSR (Figure 2(b)), but if ACC values are also large, the radius of the shaded circle (and the glyph in general) would exceed the baseline level significantly (Figure 2(c)), suggesting a manifested excitement.

To provide a temporal overview for the combined GSR + ACC values for each user, our technique supports two visual encodings based on animation: *clock-style* and *oscillating* trails which are rendered for each incoming data entry. The clock-style encoding demonstrated in Figure 3(a) uses a clock metaphor and assigns increasing angle values for items corresponding to data entries (to complete the metaphor, a clock hand is rendered as well). With this representation, a data entry is mapped to a single visual item (dot), its timestamp is mapped to the angle value, and its combined excitement value is mapped to the distance from glyph center. A new item is created for any incoming entry, highlighted with a thick stroke, and its transition from glyph center to the target position is animated to help the visualization users focus on this latest entry. Opacity of the visual items decays over time, and with the default settings, items fade out completely in a given time interval, which is also the default setting for a single "clock" rotation period. This time interval can be set in the configuration of the visualization to better suit the scenario in which it is employed. As such, the clock-style encoding is designed to provide a rather precise and uncluttered overview of the values for a limited time period and facilitates comparison between several glyphs.

The second visual encoding demonstrated in Figure 3(b) follows an oscillation metaphor suggestive of the user's current excitement level. For each incoming data entry, a number of dots is created in succession, the latest created dot is highlighted with a thick stroke, and the overall effect of the animation resembles a trajectory rendering. The excitement value is used to limit the distance from the glyph center for these trails, thus resulting in an oscillatory motion inside the boundary of the excitement circle. In order to better capture the user's excitement levels, the oscillation speed can increase with the level of excitement, resulting in faster motion of the dots for higher levels of excitement (i.e., larger excitement circles). This correlation of the excitement levels and oscillation speed has also shown initial positive results in terms of perceiving the group-level excitement in larger groups (12 or more individuals). The opacity of the dots quickly decays over time, but the items do not fade out completely with the default settings. Furthermore, the movement angle changes randomly when passing through the glyph center, in order to ensure an even distribution of the dot trails over the entire surface of the representation. This way, the visualization users can perceive the maximum excitement values over a time period rather than the excitement history of the user.

The excitement values (i.e., the combined GSR + ACC values) are used for the color coding of both general glyph elements and trail elements. We support a number of color maps from ColorBrewer [9] for three sequential data classes, and use the corresponding colors for the excitement values of 0.0, 1.0, and 2.0 (since we expected both GSR and ACC values to be in range [0; 1]). The colors for 0.0 and 1.0 are then interpolated using the HCL model for an actual excitement value (values 1.0 and 2.0 are used if the value exceeds 1.0). For instance, with the default settings the values 0.0, 1.0, and 2.0 are mapped to yellow, orange, and red, respectively. The colors are used by the outer shaded glyph circles as well as the trail dots.

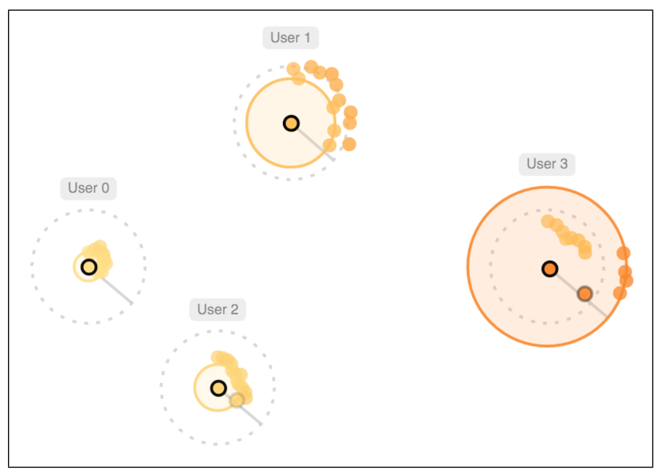


Figure 4: The displayed force-based layout for glyphs is based on similarity of latest excitement values between users.

The final aspect of our technique is the positioning of the glyphs. We support two approaches that could both be described as derived data-driven placement strategies following the Ward's taxonomy [34]. By default, glyphs are placed on a uniform grid based on the corresponding user ID values (see Figure 6 for an example). This approach may be useful when monitoring a specific glyph or continuously comparing values for several glyphs. The users can also enable a dynamic force-based layout [14] that is updated on every incoming data entry. In this case, the expected distances between glyphs are set to be proportional to differences between the latest corresponding GSR + ACC values. This approach facilitates monitoring the group dynamics and identification of clusters. Figure 4 demonstrates how users can be grouped, User 3 being overexcited while Users 0 and 2 experience low excitement.

IMPLEMENTATION DETAILS

Our prototype is implemented as a static web page with JavaScript code that uses D3 [11] for layout and rendering. The design and implementation of the prototype support streaming data. Currently, we provide a choice of several recorded or synthetic data sets with multiple users. The implementation "replays" the selected data set and emits JSON entries for single measurements based on the recorded timestamps.

The actual visualization code processes incoming entries in a streaming fashion. It is designed to handle data for user IDs not encountered previously, to update the global layout, and to interrupt current animations at the intra-user level. Figure 5 provides an overview of the update algorithm: a new glyph for a user is rendered if necessary, the layout is updated, and finally, the new data item is rendered with regard to the currently selected visual encoding for trails.

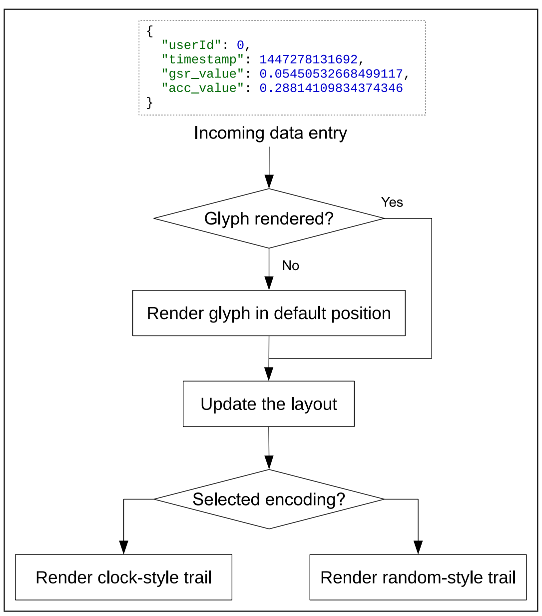


Figure 5: Diagram of the update algorithm for incoming data entries.

Figure 6 demonstrates the interface of our visualization prototype. Users are provided with options to change the visual encoding (in this case, the intra-user visualizations are reset) as well as to modify multiple visualization parameters—e.g., visual encodings and layout of the glyphs, both visual encodings for trails, etc. Moreover, the user can change the color map, the visibility of the glyph boundary elements, or adjust

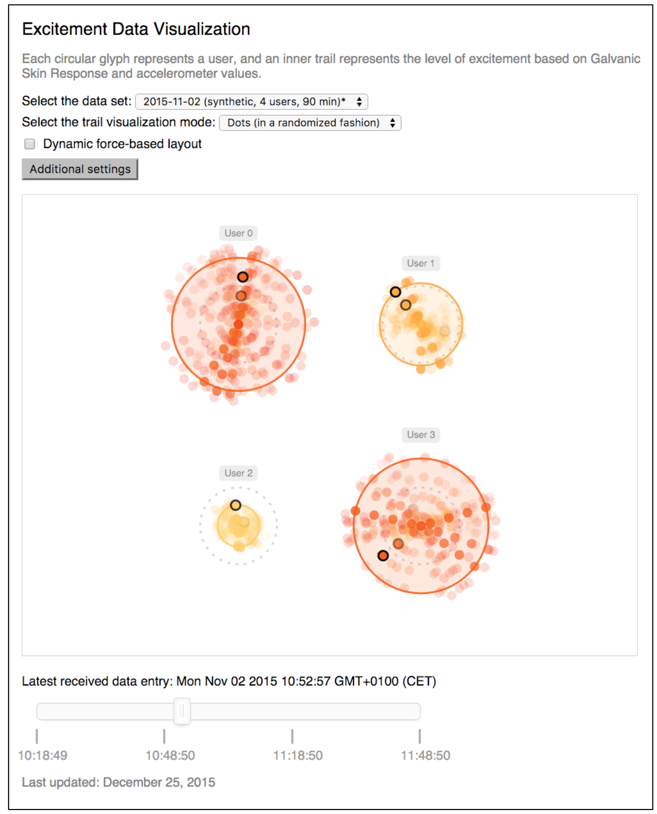


Figure 6: Screenshot of the prototype interface taken for a synthetic data set.

the animation durations and opacity values for trails. To facilitate exploration, the visualization also supports zoom & pan.

We also included a temporal slider to support navigation for our pre-recorded data sets. This choice was made mostly for the prototype testing purposes—as stated above, our general goal is to support actual streaming data sources which would make temporal navigation limited with regard to the current point in time.

VALIDATION AND DISCUSSION

A preliminary evaluation has been executed in order to validate our concept. The question we wanted to answer was whether the excitement visualization design we proposed was a suitable visual representation that could satisfy all three requirements discussed above. A group of 8 participants (5 male, 3 female, with ages between 25 and 33) were given the opportunity to visually inspect the excitement levels for up to 4 different users at the same time. Based on their experience with the visualization, the participants were asked to fill out a questionnaire.

The questionnaire included both closed-ended and open-ended questions. The closed-ended questions were scaled questions distributed on a 5-point Likert scale, where 1 represented "disagree" and 5 represented "agree".

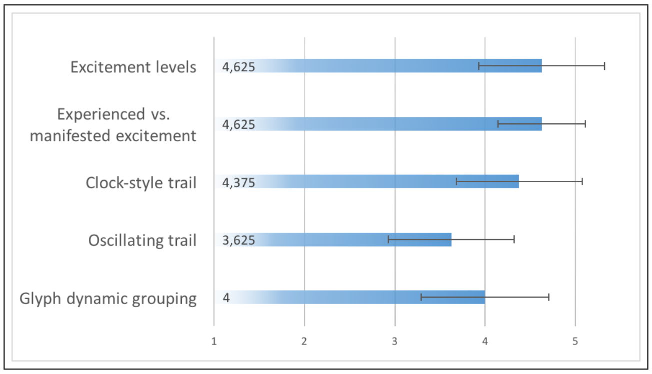


Figure 7: Averages and standard deviations for 5-point Likert scale answers provided by the questionnaire participants.

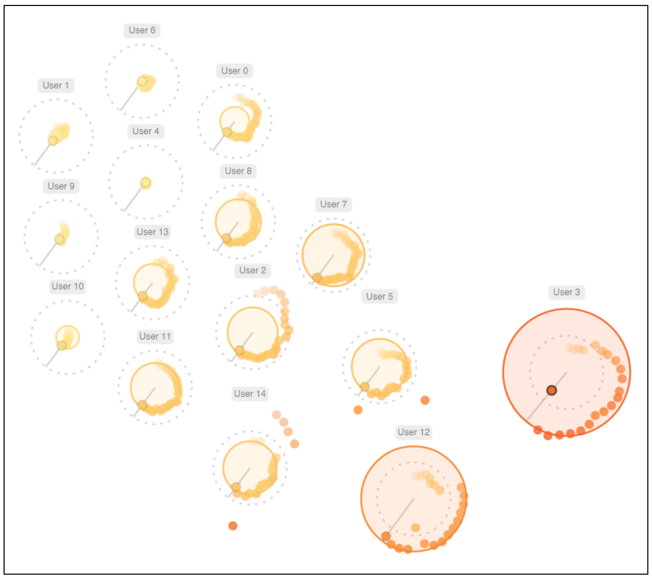


Figure 8: Visualization of a synthetic data set with 15 users with the dynamic force-based layout.

The results of the questionnaire suggest an overall positive and accurate perception of the proposed excitement visualization (see Figure 7). On average, participants rated their ability to differentiate between the various excitement levels in the visualization at 4.6. Given a concrete example (similar to Figure 4), they were asked how many people of the four individuals in the visualization were excited and manifesting it, all but one of the participants gave the correct answer, further supporting the idea that our representation is intuitive and easy to perceive. Moreover, participants felt confident about differentiating between experienced and manifested excitement (average of 4.6).

When asked about the perception of the history of excitement for each individual, the clock-style trail offered a clearer perception of the excitement history (with an average of 4.4) compared to the oscillating trail (average of 3.6). However, as previously stated, the oscillating trails are not meant to encode the excitement history per se, but rather capture the maximum excitement level experienced in a given time period. This has been also reflected in the comments, as one of the participants stated that the oscillating trail "could surface peak events". In terms of group-level excitement, the questionnaire participants found that the grouping of the circles usually helped them to accurately convey the overall excitement level in the group (average of 4.0).

In terms of open-ended questions, the participants had positive remarks related to the visual design, its simplicity and the animation elements that further support the correct perception of the excitement levels. When looking at the issues, participants suggested that the visualization can be difficult to read at times, especially as the grey circle of each individual is not always clearly visible. Additionally, multiple participants suggested extending the visualization to also incorporate positive and negative emotions (emotional valence) in the representation.

We have also inspected the excitement data we have gathered during the recording sessions (at the movies, opera, and museum) in order to investigate our visualization's capability to correctly convey moments of peak excitement. A comparison was executed between the peak excitement moments in the visualization and the self-reported logs of the participants. In multiple instances, we could observe a correlation between the self-reported and visualized excitement levels, in the case of both homogeneous group excitement (all or most members of the groups got excited) and individual excitement (only one member of the group got excited—Figure 9). During both the movie and opera activities, the participants had similar levels of base excitement. However, at certain captivating or emotional moments, the excitement levels increased synchronously. At the same time, the data obtained from the museum had fewer extracted peaks, partly due to the additional noise generated by the users walking, but also due to the fewer high excitement peaks—fact confirmed post-task by the participants as well.

While we had a rather small number of simultaneous users wearing smart wristbands in the recording sessions, the proposed visualization does scale to larger numbers of users. A visualization of a synthetic data set with 15 users in Figure 8 demonstrates that the grouping of users by excitement is present and quite easy to perceive, even though the amount of detailed information about each individual user may be large and harder to follow as the glyph changes its position with the dynamic layout. The relatively large number of users may also make the force-based layout unstable, which we plan to address in our future work.

We have also analyzed the aspects of our visualization technique with regard to the taxonomy by Cottam et al. for streaming data visualization [10]. At the inter-user level, we expect new entries (i.e., glyphs) to be added dynamically, which corresponds to the *Create* category with regard to spatial dimension (this applies to both supported layouts). Incoming data entries affect the attributes of concentric circles for a known range of values, which corresponds to the *Known Scale* category with regard to retinal dimension. According to Cottam et al., this combination of categories can be

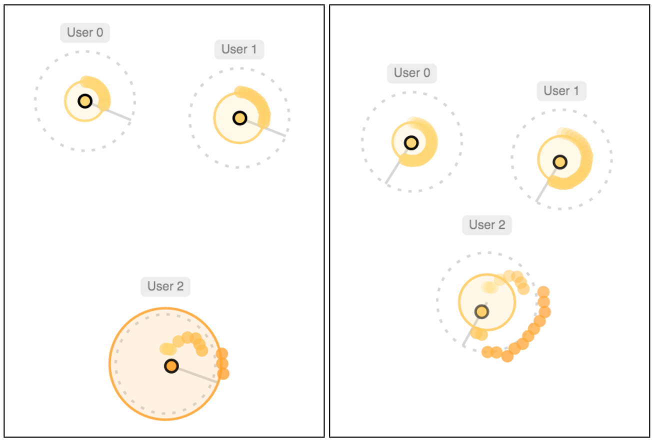


Figure 9: Two screenshots of the excitement visualization of the museum dataset. User 2, a hobby mechanic, noticed a classical car he only knew from images. This, combined with him waving to the other two ("Come and see this!") increased his level of excitement (left). After some time, his level of excitement decreased again to normal (right). As visible in the visualization, the other two individuals were rather unimpressed by the vehicle.

classified as transitional with regard to the elements' identity preservation. Such techniques usually favor comparison over short time spans, and this statement holds for our approach if only glyph layout and concentric circle elements are considered. However, our visual encodings for trails provide an overview for longer time spans than the elements mentioned above (even though the clock-style and oscillating trails on their own would be classified as *Create & Delete* × *Known Scale* and *Create* × *Known Scale*, respectively).

CONCLUSIONS AND FUTURE WORK

In this paper, we have introduced a novel visualization for representing real-time and historical excitement levels for individuals and groups. Two different visual encodings are proposed for representing individual excitement levels through animated glyphs: one focusing on capturing real-time and previous excitement levels, and one more suited for capturing individual excitement peaks and the distribution of excitement in larger groups. The perception of the group-level excitement is aided by a dynamic layout that clusters glyphs with similar excitement levels. An initial evaluation of our visualization offered promising results as well as a couple of potential improvements in terms of visual complexity.

We see several opportunities for the future work:

- The current visual encodings are limited with regard to a temporal overview for longer periods of time (such as 30 minutes or longer). This could be alleviated by introducing an additional time-varying representation [1] either attached to user-related glyphs, or a completely separate one (e.g., a simple line plot). In the latter case, an additional representation based on aggregated values could also be used to provide an overview for the whole group of users.
- Finally, the dynamic layout for user-related glyphs could be improved by either modifying the current implementation (e.g., the initial positions for the dynamic layout could be based on MDS results [4]), or changing the force-based model to use an explicit clustering algorithm.

ACKNOWLEDGMENTS

This work was partly funded by the framework grant "The Digitized Society—Past, Present, and Future" from the Swedish Research Council (Vetenskapsrådet) [grant number 2012-5659].

REFERENCES

1. Aigner, W., Miksch, S., Schumann, H., and Tominski, C. *Visualization of time-oriented data*. Springer Science & Business Media, 2011.
2. Alsakran, J., Chen, Y., Zhao, Y., Yang, J., and Luo, D. STREAMIT: Dynamic visualization and interactive exploration of text streams. In *Proceedings of the IEEE Pacific Visualization Symposium (PacificVis)*, IEEE (2011), 131–138.
3. Bakker, J., Pechenizkiy, M., and Sidorova, N. What's your current stress level? Detection of stress patterns from GSR sensor data. In *Proceedings of the 2011 IEEE 11th International Conference on Data Mining Workshops (ICDMW '11)*, IEEE (2011), 573–580.
4. Borg, I., and Groenen, P. J. F. *Modern multidimensional scaling: Theory and applications*. Springer Science & Business Media, 2005.
5. Cernea, D., Ebert, A., and Kerren, A. Visualizing group affective tone in collaborative scenarios. In *Poster Abstracts of the Eurographics Conference on Visualization (EuroVis '14)* (2014).
6. Cernea, D., Weber, C., Ebert, A., and Kerren, A. Emotion Scents – a method of representing user emotions on GUI widgets. In *Proceedings of the SPIE 2013 Conference on Visualization and Data Analysis (VDA '13)*, vol. 8654, IS&T/SPIE (2013).
7. Cernea, D., Weber, C., Ebert, A., and Kerren, A. Emotion-prints: Interaction-driven emotion visualization on multi-touch interfaces. In *Proceedings of the SPIE 2015 Conference on Visualization and Data Analysis (VDA '15)*, vol. 9397, IS&T/SPIE (2015).
8. Charlton, B. Evolution and the cognitive neuroscience of awareness, consciousness and language. In *Psychiatry and the Human Condition*. Radcliffe Medical Press, 2000.
9. ColorBrewer 2.0—color advice for cartography. `http://colorbrewer2.org/`. Accessed January 28, 2016.
10. Cottam, J. A., Lumsdaine, A., and Weaver, C. Watch this: A taxonomy for dynamic data visualization. In *Proceedings of the IEEE Conference on Visual Analytics Science and Technology (VAST)* (2012), 193–202.

11. D3—data-driven documents. `http://d3js.org/`. Accessed January 28, 2016.
12. Druskat, V. U., and Wolff, S. B. Building the emotional intelligence of groups. *Harvard Business Review 79*, 3 (2001), 80–90.
13. Empatica E4 Wristband. `https://www.empatica.com/e4-wristband`. Accessed January 28, 2016.
14. Fruchterman, T. M. J., and Reingold, E. M. Graph drawing by force-directed placement. *Software: Practice and Experience 21*, 11 (1991), 1129–1164.
15. García, O., Favela, J., Licea, G., and Machorro, R. Extending a collaborative architecture to support emotional awareness. In *Proceedings of the Workshop on Emotion Based Agent Architectures (EBAA '99)* (1999), 46–52.
16. García, O., Favela, J., and Machorro, R. Emotional awareness in collaborative systems. In *Proceedings of the International Workshop on Groupware (CRIWG '99)*, IEEE (1999), 296–303.
17. Healey, J. A. Affect detection in the real world: Recording and processing physiological signals. In *Proceedings of the 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops (ACII)* (2009).
18. Huang, D., Tory, M., Aseniero, B. A., Bartram, L., Bateman, S., Carpendale, S., Tang, A., and Woodbury, R. Personal visualization and personal visual analytics. *IEEE Transactions on Visualization and Computer Graphics 21*, 3 (2015), 420–433.
19. Huisman, G., van Hout, M., van Dijk, E., van der Geest, T., and Heylen, D. LEMtool: measuring emotions in visual interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13)*, ACM (2013), 351–360.
20. Huron, S., Vuillemot, R., and Fekete, J.-D. Visual sedimentation. *IEEE Transactions on Visualization and Computer Graphics 19*, 12 (2013), 2446–2455.
21. Kapp, K. M. *The Gamification of Learning and Instruction: Game-based Methods and Strategies for Training and Education*. Pfeiffer, 2012.
22. Liu, Y., Sourina, O., and Nguyen, M. K. Real-time EEG-based human emotion recognition and visualization. In *Proceedings of the International Conference on Cyberworlds (CW)*, IEEE (2010), 262–269.
23. McDuff, D., Karlson, A., Kapoor, A., Roseway, A., and Czerwinski, M. AffectAura: An intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '12)*, ACM (2012), 849–858.
24. Mehrabian, A. *Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies*. Social Environmental and Developmental Studies. Oelgeschlager, Gunn & Hain, 1980.
25. Microsoft Band. `https://www.microsoft.com/microsoft-band/en-us`. Accessed January 28, 2016.
26. Nakasone, A., Prendinger, H., and Ishizuka, M. Emotion recognition from electromyography and skin conductance. In *Proceedings of the 5th International Workshop on Biosignal Interpretation (BSI-05)* (2005), 219–222.
27. Poh, M.-Z., Swenson, N. C., and Picard, R. W. A wearable sensor for unobtrusive, long-term assessment of electrodermal activity. *IEEE Transactions on Biomedical Engineering 57*, 5 (2010), 1243–1252.
28. Russell, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology 39*, 6 (1980), 1161.
29. Saari, T., Kallinen, K., Salminen, M., Ravaja, N., and Yanev, K. A mobile system and application for facilitating emotional awareness in knowledge work teams. In *Proceedings of the 41st Annual Hawaii International Conference on System Sciences (HICSS '08)* (2008), 44–53.
30. Shi, Y., Ruiz, N., Taib, R., Choi, E., and Chen, F. Galvanic skin response (GSR) as an index of cognitive load. In *CHI '07 Extended Abstracts on Human Factors in Computing Systems (CHI EA '07)*, ACM (2007), 2651–2656.
31. Valenza, G., and Scilingo, E. P. *Autonomic Nervous System Dynamics for Mood and Emotional-State Recognition*. Springer, 2014.
32. Villarejo, M. V., Zapirain, B. G., and Zorrilla, A. M. A stress sensor based on galvanic skin response (GSR) controlled by ZigBee. *Sensors 12*, 5 (2012), 6075–6101.
33. Wang, S., Tanahashi, Y., Leaf, N., and Ma, K.-L. Design and effects of personal visualizations. *IEEE Computer Graphics and Applications 35*, 4 (2015), 82–93.
34. Ward, M. O. A taxonomy of glyph placement strategies for multidimensional data visualization. *Information Visualization 1*, 3-4 (2002), 194–210.
35. Westerink, J. H. D. M., van den Broek, E. L., Schut, M. H., van Herk, J., and Tuinenbreijer, K. Computing emotion awareness through galvanic skin response and facial electromyography. In *Probing Experience*, J. H. D. M. Westerink, M. Ouwerkerk, T. J. M. Overbeek, W. F. Pasveer, and B. de Ruyter, Eds., vol. 8 of *Philips Research*. Springer, 2008, 149–162.
36. Wong, P. C., Foote, H., Adams, D., Cowley, W., and Thomas, J. Dynamic visualization of transient data streams. In *Proceedings of the IEEE Symposium on Information Visualization (INFOVIS '03)*, IEEE (2003), 97–104.

SYM: Toward a New Tool in User's Mood Determination

Willy Yvart
DeVisu (UVHC)
TCTS (UMONS)
Rue Michel Rondet,
59135 Wallers, FRANCE
willy.yvart@neuf.fr

Charles-Alexandre Delestage
DeVisu (UVHC)
Rue Michel Rondet,
59135 Wallers, FRANCE
ca.delestage@gmx.fr

Sylvie Leleu-Merviel
DeVisu (UVHC)
Rue Michel Rondet,
59135 Wallers, FRANCE
sylvie.merviel@univ-valenciennes.fr

ABSTRACT

Even though the emotional state is increasingly taken into account in scientific studies aimed at determining user experience of user acceptance, there are still only a few normalized tools. In this article, we decided to focus on mood determination as we consider this affective state to be more pervasive and more understandable by the person who is experiencing it. Thus, we propose a prototypical tool called SYM (Spot Your Mood) as a new tool in user mood determination to be used in many different situations.

Author Keywords

mood; determination; affective state; theory; emotion

ACM Classification Keywords

Algorithms; Experimentation; Human Factors; Measurement; Theory

INTRODUCTION

Since Aristotle and the Greek philosophers, the question of the affective state has stayed central and very questioned in most areas of science. With the emergence of psychology and other cognitive sciences, the subject seems to be even more unavoidable. The constructivist paradigm, by putting the person in the very center of the sense-making process, collaterally brings the question of the influence of his affective state to the table [1].

Even though the subject has been treated over hundreds of years, it is quite noticeable that it still suffers from a lack of consensus and sometimes a lack of recognition in most technical areas. Actually, there is a lot of confusion; most terms and concepts are ill-defined and they often cannot escape from the epistemic posture of authors.

In this article, we decided to work on the mood concept that is indisputably less "fashionable" that the emotion concept. We put forward that mood is, however, a central parameter of the whole configuration of mental activity. Having said that, the question of knowing about the mood of the person or the population we are studying in an experimentation cannot remain unanswered, it has to be studied as a potential parameter of the experimental situation in order not to become a bias. Thus, after disambiguating as best as possible the concept of mood, we propose a protocol and a prototypical tool intended to solve that question.

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linkoöping University Electronic Press.

DETERMINATION OF AN AFFECTIVE STATE

Why mood?

When taking into account the emotional state (as opposed to traits we are not talking about here) in determining its effect on reasoning or decision making, we first need to disambiguate which affective concept fits the most with user experience considerations.

Empirically, when confronted with some experience, listening to music, visiting a museum and so on, a person induces a change in his actual affective state. Nonetheless, the qualification of this affective state is still quite fuzzy. Generally, in the literature, we can find three main concepts that are sometimes mixed up: "affect", "mood" and "emotion".

First, "affect" seems to be a portmanteau word, a concept embodying "mood", "emotion", "feelings", etc. There is, at the same time, confusion with the word "affect" in itself. As used in the expression "affective state", it refers to a state, an immediate emotionally "colored" cognitive state, but, the word can also be used to qualify feelings in the most sensitive signification. Then, as "affect" appears to be a supra-concept covering different ill-defined infra-concepts, we have to rely on another.

The concept of emotion, even if it is the most commonly encountered in the literature, also seems to be the fuzziest. In fact, more than 30 years ago, Kleinginna and Kleinginna were able to point out more than one hundred different definitions [2]. By researching for this communication we noticed over two hundred definitions that are not compatible: 38% of these consider emotions as being a physiological phenomenon while 62% consider them as being a psychological one. This lack of consensus brings about most of the misunderstandings we can notice in the literature. Nevertheless, there is some agreement; in most definitions "emotion" refers to a response to some internal or external stimulation of short-term duration and quite strong effects on the behavior. Thus, emotions are seen as being intentional (associated to an object or a stimulus) and transient. By looking again in the literature (the articles in which we found the definitions, which numbered over 200) we could see that a third of the statements we found considered that emotions are inaccessible to the conscious mind, versus two thirds which advocated the opposite. To us, emotion seems to be, most of the time, a "*qualia-like*" phenomenon being analyzed as a part of the information available in order to copy with the situation. The flow of data of a physiological nature (heart rate, blood pressure, etc.) or psychological nature (feelings) is melted into the rest of the sensitive and cognitive information during the experience. That implies that emotions are too volatile, too transient and not stable on a macroscopic time scale. According to Sloboda [3], emotions are "rather unmemorable on average";even though this statement looks polemical, it is still possible to agree to with it despite some reservation. Strong reciprocal links between emotions and memory have been shown from the recording to the reading of the information, and these cannot be denied [24-26]. However, and in accordance with Sloboda's statement it is not the emotional experience in itself that is recorded or recalled. We can remember about contextual cues, ongoing actions (like in the Ziegarnik's effect [27]) and about the consequences of the experience. We do record/recall the initial setup (initial mood, environment, etc.), the perturbing element, and the final setup (consequences, final mood,..) with some kind of an emotional label on it which is like the very source of feelings such as nostalgia for example [28]. Body feelings, sensations, *etc.* are, to a large extent, lost as it is for any "*qualia-like*" phenomenon.

So, in an attempt, like ours, to try to characterize the affective experience of a person, they constitute a cul-de-sac.

As a consequence, our epistemology has to be built on the last commonly encountered concept: the "mood". The concept is sometimes considered as being old fashioned, especially in its French translation "*humeur*". Even if we could notice up to 88 different statements in the literature, chronologically from Weld [4] to Juslin and Sloboda [5], there is more agreement between authors.

Hence, the mood stands for a state, at a given time, in a given situation, that is linked to the framework of the whole organism by the attention given to its sense-making and cognitive processes [6], [7]. Thus, mood is seen as the psycho-physiological substratum of any situation; whenever we are in a certain situation - in other words, all the time - we are in a certain mood [8] (very similar to Heidegger's concept of "*stimmung*" [9]) even though we are only able to be conscious of that state if we focus our attention on it.

Mood can be seen as a relatively stable and pervasive setup that can be maintained from minutes to days. Without a perturbing element, a mood can be sustained or can slowly decay to a more "stable" one. We can draw a quick analogy with a homeostatic state in the chemical sciences. In the attempt of Lane and Terry to give a theoretical basis to the concept of mood, they put forward that it is "*a set of feelings, ephemeral in nature, varying in intensity and duration, and usually involving more than one emotion*" [8]. In this context, emotions can then be seen as the perturbing elements we were talking about. This posture is compatible with the fact that emotions are quasi-inaccessible in themselves to the cognitive process. In our opinion, it is the mood and the difference between a pre-stimulation and a post-stimulation mood that are accessible to the cognitive process.

How to measure/represent mood?

That being said, there is still a central question to answer: how to measure mood and how to represent it in a sufficiently efficient and communicable way? Following on from what we just said, we can put as a working hypothesis that we are trying to characterize an affective experience in a certain situation by looking at the mood and at its evolution during this situation. Although emotions are not accessible, we still can notice the "dents" they produce on the mood.

So, we need to follow the evolution of an internal parameter of the visitor / listener / spectator. Nevertheless, to follow an evolution, we need to characterize differences. In fact, we need to find information about the quasi-homeostatic affective state at the beginning of the experience we want to study and, *a minima*, the final one. Thus, the difference between the two "measures" would be analogous to a kind of integrative view of the affective experience.

But, that does not answer the "how to" question. At that time, we have to choose between two epistemic points of view which are opposite to each other. On one hand, many psychologists and scientists put forward that it is possible to measure some affective phenomenon by looking at physiological reactions like micro-expressions [10] or heart-rate[11]. On the other hand, we can emphasize Barrett's statement saying: "*for better or worse, self-report represents the most reliable and possibly only window that researchers have on conscious, subjective emotional experience*." [12]

Taking a look at Mugur-Schächter's works on theorizing the qualification and measuring process [13], [14], we would preferably agree with Barrett. Actually, to qualify implies to measure and to measure implies a measuring tool. Whatever we are trying to qualify, we first need to build a link to it starting from the conscious process. Something from the outside of our body will then be measured throughout our biosensors, our attention and our cognition. Something from the inside of our body will be measured throughout our attention to it and our cognition.

But it seems common grounded that affective phenomena like mood are both internal and external phenomena from a conscious point of view. Consequently, looking for somebody else's mood by looking at physical or physiological variation is like erasing a big part of the information. The only person able to collect information from the situation and from its inner world is then the person living the experience.

Nevertheless, that statement does not totally disqualify physiological measures like heart-rate. Actually, heart-rate variability can be a useful additive signal that can constitute a highly profitable source of information.

Asking for the mood

If the person is in the best position to determine his mood, we now have to ask him for it. According to Shannon's theory of information, a transmitter and a receiver have to speak (at least) the same language to code and decode the message. That is why spoken or written language appeared to be a good idea at their origin.

But, as we were able to notice during previous experimentation, verbalization is not the easiest way to study affective response when listening to music [15]. Actually, in an experiment we asked for people to qualify the mood of a musical excerpt using adjectives. We wanted to test an open-ended condition in order to estimate personal judgment about the music expressed in their own language.

The supra-context of the experiment was to try to establish a minimal set list of mood adjectives in order to use them as the standing ground for indexing and searching for the musical excerpt in an online library (the reason why we needed to have the mood expressed in the user's own language). In this context, the qualifier has to reach a certain threshold of consensus in order to be meaningful. Due to the impressive quantity of different adjectives we gathered, we then decided to build up some semantic proxemy analysis so as to meet an inter-subjective agreement.

In this attempt, we discovered that there was agreement to a certain degree. Actually, most of the terms seemed to coincide with a main idea as they were, for the most part, synonyms. But we could also notice that there was some confusion. Indeed, there appeared to be a discrepancy between the personal mood of the person and an extra-personal mood designating the excerpt. That sounds logical to say that in the musical signal there is no mood, only a living person can endure such an affective state, but when submitted to a *stimulus* able to induce or to produce emotions, a person is able to assign an emotional content to that *stimulus*. At that moment, there can be confusion between a "personally-felt" mood and a "perceived" one [16], [17].

The results of this experiment brought us to develop a new protocol to deal with the issues of verbalization and confusions between moods.

SYM: SPOT YOUR MOOD

Getting over the verbalization problem is a big issue. In the context of our first experiment, deleting the massive set of qualifiers would represent a more efficient tagging process resulting in a better indexation. Thus, we decided to rely on some graphical expression and representation systems to assess for the mood. Like in the Self-Assessment Manikin [18] or in the Geneva Emotion Wheel [19] everything starts with the use of dimensional theory of emotion and the valence-arousal space notably developed by Russell [20] and Thayer [21].The valence dimension consists in a hedonic value of pleasantness / unpleasantness and the arousal dimension is relative to a psycho-physiological wakefulness. According to Russell, every possible affect is representable by a set of two values in those dimensions.

The SAM allows the user to give information about his emotional state through valence, arousal and control/potency (the last is not always represented) discrete 10-step-Likert-scales. The GEW consists in a flower of circles representing the emotional universe cut into discrete families of emotions. Both give a discrete representation of the mood space. In our approach, we decided to stay more faithful to the dimensional approach described by Russell.

Thus, we directly present a valence-arousal space to the user and ask him to express his mood throughout one or many points. In our representation, the valence dimension is the abscissa axis and the arousal is the ordinate axis. *Extrema* are then characterized with smileys representing high valence, low valence, high arousal and low arousal in order to avoid verbal influence.

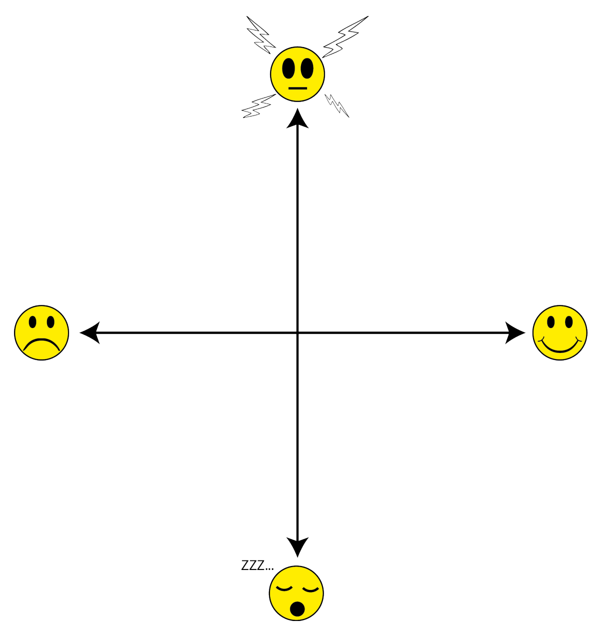


Figure 1 : SYM's Valence-Arousal diagram

Until then, the user is free at any time to designate a point on the figure to give us information on mood. We let the user have the impression of a continuous scale, however the indications are digitally stored in the discretized interval[-

100;100][-100;100]. The software architecture is a traditional client-server routine, but the idea is to allow the researcher to use plug-ins such as indoor localization or synchronization with audio or video, depending on the experimental needs.

Also, the major utility of SYM is to provide a dynamic dictionary of emotional states linked to valence-arousal locus. This dictionary consists in a database, which links nouns, verbs, adverbs and expressions related to the same emotional concept. With semantic proxemy analysis between the terms, it is possible to also link some emotional states in order to create a kind of net.

The user spots a mood which is represented as a couple of (x,y) coordinates. These coordinates are transmitted to a server, which interrogates a data-base and finds the 3 closest mood terms and sends them to the client. If the user is satisfied, the server stores the tuple (x, y, word). If not, the process is the same, excepting the last words refused by the client. Doing that, SYM can be seen as a verbalization helper. As we were able to notice during previous experimentations, it is quite hard to express clearly how we feel when asked to do so. Moreover, direct verbalization also presents polysemic issues and the use of adjectives that a computer cannot disambiguate.

Here, as every situation can lead to the usage of the same words with different meanings, a section of the master dictionary can be set up to fit the expression of the users. Also, some words will not be used on some occasions. For example, you will rarely find "anger" expressed as a musical mood. Thus, with the help of some specialists, for example in museography, we can determine the terms that would fit or not to the situation, then generating custom dictionaries. These custom dictionaries bring more simplicity for the user, avoiding "noise words" being proposed while spotting on the diagram.

Periodically, the server makes a separate update of the position of words on the valence-arousal based on the user's feedback, to include the folksonomy generated throughout the different experimentations. As a result, the dictionary used for each setup is a part of the master dictionary with custom placements of the mood words regarding the conditions.

This function allows us to make a difference for the usage of words, which can differ for certain activities (shopping, cultural exhibitions,…) or within different populations (street language, rich language, etc.).

The architecture permits us to use the client terminal or to visualize in a browser the results of the experiments, placed on the diagram for each user or in a cloud of points. The raw data can also be imported as *.csv* files on any statistics software for treatment and synchronization with physiological signals or experiment props (music, video…)

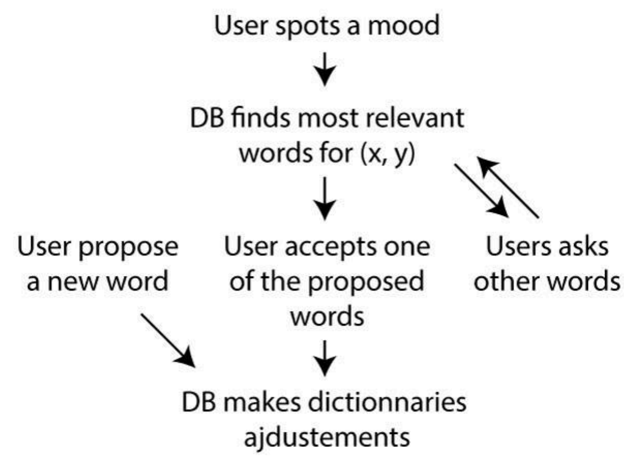


Figure 2 : SYM simplified workflow

MOOD SPOTTING

A validation of mood spotting on Valence-Arousal space

The usage of a valence-arousal space with a similar approach can be found on the experiments led by Fanny Bougenies during her PhD thesis [23]. The protocol was using it to assess the emotional state of about 130 children with or without deficiencies (deaf, mental impairment, autism…) before and after a visit to a museum. Each of them was given a tablet with an educational software program designed to enhance their visit. At that time, valence-arousal graphs were paper-made and consisted in a part of the questionnaire.

The conclusions showed that not a single child endured difficulties understanding, they managed to use the diagram without indications. Moreover, it allowed autistic children, generally not willing to talk much, to express their emotions smoothly.

The problem was that all these indications were taken on a calibrated sheet of paper. It meant precise measurement, normalization and reporting on a computer (with a computer-aid design software) to get a usable corpus of data. This underlined the necessity of finding an appropriate computerized workflow to gather and process the moods.

Mood spotting for determining visitor experience in museum exhibition

The testing experimentation was focused on the ability of the interface to be used by the public without altering their experience with a significant impact. It took place in the "Musée des Beaux-Arts" of Lille in France. The visitors wanting to visit the museum were offered a free pass just before the act of payment for their tickets, but they were required to be equipped with eye-tracking glasses and a tablet with an android version of SYM. At any time, they could spot their mood on the device. For 50% of them, a word was also provided to qualify their mood, based on the point they spotted, which they had to validate or not. After their visit, they were shown their field of view during their visit (from the eye-tracking glasses) and were asked to describe their experience.

The result was that all of the interviewees were able to spot their mood easily, and without any long explanations of the system. Also, the impressions and comments given after the visit correlated with the tuples (x, y, word) spotted on the application. If not as precise as a full interview after the experiment, the SYM approach was able to give the shape of the experience of the visitor just as the affective state shapes the user experience. This could be the first step of an evaluation protocol to distinguish the points in the exhibition that have issues.

CONCLUSION

SYM was originally designed as a "verbalization" helping tool when it comes to expressing mood. Even if we are still prototyping the tool in itself, we managed to build a complete integrative workflow that we are still amending. The experiment we presented showed us that SYM was very efficient as a non-pervasive experimental protocol when inquiring about affective states.

The participants were only told to spot their mood at the beginning and at the end of the visit and then were free, for the rest of time, to inform us or not about their state of mind. This results in a very small probability of being exposed to the social desirability bias [22] which could be feared as we penetrate into the very private area of affective states.

Before the experiment took place, we had already noticed the efficiency of pre-SYM attempts. In a still unpublished study about qualifying music throughout mood spotting, we noticed that the listeners were able to understand the diagram at a glance and give a point for their mood.

We also noticed that users were able not to mistake between their own mood and the musical one. Actually, in order not to fall into this trap, we decided to inquire into their mood before and after the listening and about the "musical" mood during the listening. It is interesting to see that there is no possible inter-personal common-ground about the own-mood (nor about its evolution) although we can recognize a certain degree of convergence for the inter-personal musical moods. That can be seen as a successful distinction between the personal and the musical mood. Furthermore, that gave us two interesting pieces of information about the music; on one hand, the difference between the beginning and the end of the experiment related to the emotional effect of the music and, on the other, a mood assigning for the "emotional content" of the musical excerpt in itself.

Admittedly, SYM is still a prototype and it needs to be improved and amended to fit the different experimental setups, nevertheless, according to the first results we had in the field, SYM can open new perspectives in user mood determination.

ACKNOWLEDGMENTS

SYM is a part of a co-directed program between the university of Mons in Belgium and the university of Valenciennes in France. We would like to thank Caroline Guillaume-Blaydes for her corrections to our text and all the participants in the different experiments we are still working on.

REFERENCES

1. H. A. Simon, ‘Motivational and Emotional Controls of Cognition’, *Psychol. Rev.*, vol. 74, no. 1, pp. 29–39, 1967.
2. P. R. J. Kleinginna and A. M. Kleinginna, ‘A categorized list of emotion definitions, with suggestions for a consensual definition’, *Motiv.Emot.*, vol. 5, no. 4, pp. 345–379, 1981.
3. J. A. Sloboda, ‘Music in everyday life : the role of emotions’, in *Handbook of music and emotion : theory, research, applications*, Oxford, England: Oxford University Press., 2011, pp. 493–514.
4. H. P. Weld, ‘An experimental study of musical enjoyment’, *Am. J. Psychol.*, vol. 23, no. 2, pp. 245–308, 1912.
5. P. N. Juslin and J. A. Sloboda, *Handbook of music and emotion : theory, research, applications*. Oxford: Oxford University Press, 2011.
6. J. P. Forgas, ‘Mood and Judgment: The Affect Infusion Model (AIM)’, *Psychol. Bull.*, vol. 117, no. 1, pp. 39–66, 1995.
7. J. D. Green, C. Sedikides, J. A. Saltzberg, J. V. Wood, and L.-A. B. Forzano, ‘Happy mood decreases self-focused attention’, *Br. J. Soc. Psychol.*, vol. 43, pp. 147–157, 2003.
8. A. M. Lane and P. C. Terry, ‘The nature of mood : development of a theoretical model’, *J. Appl. Sport Psychol.*, vol. 12, no. 1, pp. 16–33, 2000.
9. M. Heidegger, *Being and Time: A Translation of Sein und Zeit*. SUNY Press, 1996.
10. P. Ekman, ‘Facial expression and emotion’, *Am. Psychol.*, vol. 48, no. 4, pp. 376–379, 1993.
11. F. Bousefsaf, C. Maaoui, and A. Pruski, ‘Remote assessment of the heart rate variability to detect mental stress’, in *Proceedings of the 7th International Conference on Pervasive Computing Technologies for Healthcare*, 2013, pp. 348–351.
12. L. F. Barrett, ‘Hedonic tone, perceived arousal, and item desirability : Three components of self-reported mood’, *Cogn. Emot.*, vol. 10, no. 1, pp. 47–68, 1996.
13. M. Mugur-Schächter, *Sur le tissage des connaissances*. Paris, France: Hermès, 2006.
14. M. Mugur-Schächter, ‘Les leçons de la mécanique quantique’, *Le débat*, vol. 2, no. 94, pp. 169–192, 1997.
15. W. Yvart, T. Dutoit, and S. Dupont, ‘Une approche info-communicationnelle des librairies musicales en ligne’, in *Proc. SFSIC’14*, Toulon, France, 2014.

16. L.-O. Lundqvist, F. Carlsson, P. Hilmersson, and P. N. Juslin, ‘Emotional Responses to Music : Experience, Expression, and Physiology’, *Psychol. Music*, vol. 37, no. 1, pp. 61–90, 2009.
17. P. N. Johnson-Laird and K. Oatley, ‘Emotions, Music, and Litterature’, in *Handbook of emotions*, 3rd Edition., New York, NY, USA: Guilford Press, 2008, pp. 102–113.
18. M. M. Bradley and P. J. Lang, ‘Measuring emotion: the self-assessment manikin and the semantic differential’, *J. Behav. Ther. Exp. Psychiatry*, vol. 25, no. 1, pp. 49–59, 1994.
19. K. R. Scherer, ‘What are emotions? And how can they be measured?’,*Soc. Sci. Inf.*, vol. 44, no. 4, pp. 695–729, 2005.
20. J. A. Russell, ‘Affective Space is bipolar’, *J. Pers. Soc. Psychol.*, vol. 37, no. 3, pp. 345–356, 1979.
21. R. E. Thayer, *The Biopsychology of Mood and Arousal*. New York, NY, USA: Oxford University Press, 1989.
22. D. P. Crowne and D. Marlowe, ‘A new scale of social desirability independent of psychopathology.’,*J. Consult. Psychol.*, vol. 24, no. 4, p. 349, 1960.
23. F. Bougenies, ‘Expérience de visite muséale for all : visite augmentée et construction de sens. Le cas d'enfants avec et sans handicaps au Palais des Beaux-Arts de Lille.’, *Université de Valenciennes et du Hainaut-Cambrésis*, 2015
24. G. H. Bower, S. G. Gilligan, and K. P. Monteiro, ‘Selectivity of learning caused by affective states’, *J. Exp. Psychol. Gen.*, vol. 110, no. 4, pp. 451–473, 1981.
25. J. R. Anderson and G. H. Bower, *Human associative memory*. Washington, D.C., USA: Winston & Sons, 1973.
26. S. Jhean-Larose, N. Leveau, and G. Denhière, ‘Influence of emotional valence and arousal on the spread of activation in memory’, *Cogn. Process.*, pp. 1–8, 2014.
27. B. Ziegarnik, ‘Das Behaltenerledigter und unerledigter Handlungen’, *Psychol. Forsch.*, vol. 9, pp. 1–85, 1297.
28. M. M. Bradley, M. K. Greenwald, M. C. Petry, and P. J. Lang, ‘Remembering pictures: pleasure and arousal in memory.’,*J. Exp. Psychol. Learn. Mem. Cogn.*, vol. 18, no. 2, p. 379, 1992.

Position Statements

The Sound Challenge to Visualization Design Research

Niklas Rönnberg
Division for Media and Information Technology
Linköping University
SE-581 83 LINKÖPING
SWEDEN
niklas.ronnberg@liu.se

Jonas Löwgren
Division for Media and Information Technology
Linköping University
SE-581 83 LINKÖPING
SWEDEN
jonas.lowgren@liu.se

ABSTRACT
This paper is an introduction to the emotional qualities of sound and music, and we suggest that the visual and the aural modalities should be combined in the design of visualizations involving emotional expressions. We therefore propose that visualization design should incorporate sonic interaction design drawing on musicology, cognitive neuroscience of music, and psychology of music, and identify what we see as key research challenges for such an approach.

Author Keywords
Emotion; music; sound; multimodal visualization design.

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous.

INTRODUCTION
User interaction is of importance in many types of visualizations. As emphasized by recent developments in interaction design and UX, the potential significance of users' emotional states for the outcomes of the interactions cannot be ignored. Consequently, a user's possible emotional state should be considered already during the design phase. Design is not only about enabling users to solve problems; it operates also on emotional levels, stimulating affective and visceral reactions in users. Visualization research has traditionally limited itself to a rather utilitarian stance, effectively missing out on the emotional potential inherent in conscious and purposeful design. Extending the design palette to multimodal representations, and specifically to sound and music, is a particularly powerful approach when it comes to emotional impact. This position paper is an introduction to what is known about the emotional qualities of sound and music, and a few corresponding research challenges for visualization design.

THE EMOTIONAL QUALITIES OF SOUND AND MUSIC
Research on sound, music and emotion is interdisciplinary and covers music psychology, music cognition, and cognitive neuroscience of music [1], as well as musicology and philosophy of music. According to the *expression theory*, the expressiveness and meaning of music arises from emotions experienced by the composer. In contrast, the *arousal* or *evocation theory* states that emotions expressed in music are aroused by expectations of the listener [2]. Even though the origin of the emotions evoked by the music differs between these two standpoints, what they have in common is that emotions are evoked by music and that these emotions are part of the music. This suggests that music, and to some degree sound in general, evokes emotions in the listener. Furthermore, it has been suggested that experiencing emotions when listening to music is the main reason why people listen to music [3]. Research has also shown that everyone has the possibility to have meaningful musical experiences [4], which in turn implies that the emotional state of every listener can be affected by music.

It has been shown that music conveys emotions, and that music as well can elicit the same emotions in the listener [5]. This was shown by comparing two musical excerpts, one with fast tempo and in major tone and one with slow tempo in minor tone as these musical structures are known to convey happiness and sadness, respectively. Subjective ratings of the listener's emotions showed elevated levels of happiness after listening to music expressing happiness as well as elevated levels of sadness after music expressing sadness. As the intended emotion was aroused in the listener, the music conveyed as well as elicited the emotion. Music is one of the most common ways of having peak experiences [6]. A peak experience is an intricate, transcendent, and intense experience that is crucial for the development of the person. Intense musical experiences can result in lasting changes in a person's values, social relationships, and development [7].

Musically induced emotions elicit similar frontal brain regions compared to emotions activated by other stimuli [8]. For example, joyful and happy sections of a musical piece are associated with increased activity in the left frontal lobe, while fearful and sad musical sections are associated with increased activity in the right frontal lobe activity [9]. This corresponds to research showing that the right hemisphere of the brain is of importance for

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linköping University Electronic Press.

processing primary emotions such as fear or sadness, while the left hemisphere is important for preprocessing social emotions [10]. Social emotions, as compared to basic emotions, are emotions that require the interpretation of mental states of other people [11]. This might indicate that the joyful sections in music increased the activity in the left frontal because of feeling happy is associated with social sharing of a happy event or a happy memory, while feeling sad rather is an emotion connected to loneliness. Furthermore, the intensity of the perceived emotions as measured by the overall frontal region activity, increased as the musical stimuli became more intense [9]. This suggests that the degree of emotional response can extensively be manipulated by the musical stimuli.

Music and sound can be seen as high priority communication, as the psychoacoustic elements affect the listener immediately and unconsciously. Psychoacoustics is the scientific study of sound perception; that is the psychological and physiological responses associated with sound. Psychoacoustic elements of sound are: pitch, loudness, time, and timbre to mention a few [3, 12, 13]. Pitch is the perceived frequency of the sound. Tones of different pitch are combined together to form melodies, and an ascending pitch is generally perceived as more positive (such as happiness) while a descending pitch is perceived as more negative (like sadness or anger) [3]. Loudness is the perception of amplitude, the sound level. Loud sound levels might, for example, lead to increased blood pressure or increased annoyance [14], and thus might be more activating and engaging compared to less loud sounds. The feeling of fear is usually associated with a low sound level, while happiness and sadness are expressed at medium sound level, and anger at a high sound level [3]. Time is about the duration of the sound, but it also involves rhythm, meter, and tempo. A faster tempo is more arousing and engaging compared to a slower and more relaxed tempo [14] and a slow tempo evokes more calm emotions (such as sadness or tenderness), while a faster tempo evokes more active emotions (for example happiness or anger) [3]. Timbre means the harmonic content of the sound, which creates the character of the sound. A more complex timbre is more captivating with a greater emotional response as a result, compared to a simpler timbre [14]. Emotions such as tenderness and sadness are aroused with a soft or dull timbre, while happiness is evoked with a bright timbre, and anger with a sharp timbre [3].

Listening to music starts and affects physiological activities such as heart rate, perspiration, and respiration in a similar way as other emotional responses [15]. Furthermore, when listening to music, neurochemicals such as dopamine [16] and serotonin [17] are released, leading to intensely pleasurable feelings. Studies have shown that varying musical performance variables in the music, such as tempo, dynamics (differences in loudness), and articulation (how the music is expressed via the musical instrument), communicates different emotions [18]. Some studies have used experienced musicians in the experimental settings, which might suggest that these individuals have a level of expertise and musical understanding that is not necessarily representative for the general population, as musical training has been considered one of the most influential aspects for emotional responses to music [19]. This suggests that it is of great importance to know the target audience to be able to choose the right level of musical complexity when designing. Furthermore, a listener is fast, as fast as less than an eighth of a second, in judging whether the music is good or bad [20], and in determining the musical genre [21]. Additionally, the listener's music preferences influence the experienced emotions elicited by music [19]. This further underlines the importance of choosing the correct music in the design process to achieve the right emotional impact on the target audience.

Music and sound have been used successfully to create mood and emotional settings in film for quite a long time [22]. The dramatic underscore, the so-called unheard melodies used in film and computer or video games [23], can be seen as an invisible, unheard, carrier of emotions that creates narrative and connotative clues to help the viewer interpret and understand the visual information and the course of events, while simultaneously contributing to continuity and cohesion [22]. Even though the sonic stimuli in this text so far have mainly been phrased as music, the same theoretical considerations also concern sound meaning and interpretation [24].

VISUALIZATION DESIGN AND EMOTIONS

Measuring emotion among users of visualizations could lead to more effective and efficient communication. Our position extends this idea in two ways.

First, we would like to draw attention to the fact that visualization traditionally addresses mainly visual expressions, whereas it is known that expressions in other modalities offer even greater potentials to evoke affective reactions. Specifically, our interest lies with aural expressions – musical sounds and other types of sounds – in combination with visual expressions (compare how Tweedie et al. [25] argued already in the late 1990s for the relabeling of visualization to externalization to accommodate the by-then new possibilities for multimodal expression).

Secondly, we want to emphasize the vast body of existing knowledge and best practices when it comes to designing visual and aural expressions with specific affective goals in mind. In graphic design as well as in sound design and music composition, the ability to anticipate the affective outcomes of certain expressions, and the ability to design an expression in order to achieve a certain affective outcome, are highly developed and accurate in many design situations. However, from an academic point of view, these abilities have traditionally been somewhat overlooked since they are based on a body of experiential knowledge that is passed on through apprenticeships and practice rather than through scientific communication and academic literatures.

In recent years, however, an independent field called design research is starting to form [26]. That field is characterized by exploring forms of scientific communication in which designerly knowledge becomes a first-class topic. It is further characterized by its recognition of design practice as part of knowledge production, and its orientation towards producing designerly yet well-grounded knowledge.

At this point, we have summarized some of the existing academic knowledge on the emotional qualities of music and sound, demonstrating the great potential of aural expression in evoking emotion. We have argued that the field of visualization should incorporate aural expressions in order to achieve greater emotional expression potentials, and finally we have argued that the focus of research in visualization for emotion should lie with designerly knowledge every bit as much as evaluative knowledge. To elaborate our position, we conclude with a number of research challenges that would follow from such a stance.

RESEARCH CHALLENGES

The first challenge we identify is simply to combine the visual and the aural modality in the design of effective and efficient visualizations involving emotional expressions. We therefore suggest that visualization design should incorporate sonic interaction design [27] but also encompass musicology, cognitive neuroscience of music [28], and psychology of music [12].

The two modalities have different inherent characteristics, to some extent. Visual information can be persistent, whereas aural information is necessarily transient. Visual information can convey much greater quantities and ranges of symbolic and propositional content than aural information. Visual information is spatially located, whereas aural information can pervade a room, which in turn relates to focal versus peripheral communication. From a genre-conventional point of view, also affecting the users' interpretations, visual information in the context of visualization is expected to be mostly factual or instrumental – whereas aural information is expected to be mostly about mood and affective qualities (cf. music, as discussed above, and specifically the role of soundtracks in movies or video games).

The visualization literature is scarce when it comes to multimodal expressions; early examples like Gaver's 1991 Arkola simulation for process control [29] or his Sonic Finder did not have much impact on the academic field of visualization, despite the promising nature of the results he showed through what he called "sonification". One of the most interesting findings in the Arkola factory simulation was the emergent, holistic nature of the soundscape generated by a number of independent sources of non-musical sounds from different parts of the manufacturing process. Essentially, Gaver was able to replicate to some degree through sonification the robust results from previous process control research showing how experienced operators could pinpoint problems in a large physical factory by listening to the holistic soundscape on the factory floor. Furthermore, it was already known that the holistic assessment skill was jeopardized by implementing centralized control rooms remote from the actual processes, and using only visual information to convey process state.

We find early work like this quite promising, and we would like to argue that we have now perhaps reached the technical maturity required to make visual plus aural a more common choice in visualization design. The modality characteristics outlined above point to a number of approaches worth exploring, for example:

1. Using musical sounds to induce the required arousal level in monitoring tasks. For instance, if an intelligent air traffic control system predicts a conflict between two entities within a time frame that will shortly necessitate action on behalf of the operator, the ambient musical soundscape of the air traffic control room could change its mood from relaxed towards more tense to attract more attention and preparation for action.

2. Using non-musical sounds from individual entities of a large information set to form an emergent soundscape with combined rhythm and timbre that arouses different emotional qualities from which the current overall state can be experienced holistically and peripherally. This would enable full concentration on focal tasks with simultaneous peripheral awareness of significant changes in the overall state, for example in control and monitoring tasks as well as visual analytics.

3. Using musical sounds for sonification when visualizing complex data to support interpretation and comprehension of the visualization. These sounds could create emotional responses by means of harmony or disharmony to make relations between data variables subconsciously and pre-attentively perceptible. Also, these sounds might be used to illustrate the density of, as well as the blend between, the variables.

4. Using musical sounds when giving feedback during and after a user interaction could further improve the interaction by influencing the emotional state of the user. Musical elements, such as rhythm, amplitude, and harmonic content could be relaxing or soothing and thus reduce an excited and stressful tendency in the interaction. On the other hand, if swiftness is of importance, musical elements could urge the user to interact by means of faster rhythm and increased amplitude.

The second overall challenge to visualization research is a meta-challenge, and it involves adopting a design research perspective in order to capitalize on the vast body of existing designerly knowledge concerning the emotional qualities of visual and aural expressions. Similar calls to the visualization research community have been voiced before

[30-32], but apparently to little effect. Thanks to recent developments in design research, the time might be ripe for appropriating design research concepts and methods into visualization research, or to put it differently, to address multimodal visualization as a design research discipline.

REFERENCES

1. Eerola, T. & J.K. Vuoskoski, A Review of Music and Emotion Studies: Approaches, Emotion Models, and Stimuli. Music Percept, 2013. 30(3): p. 307-340.
2. Davies, S., Musical Meaning and Expression. Vol. 46. 1994: Cornell University Press. 533.
3. Juslin, P.N. & P. Laukka, Expression, Perception, and Induction of Musical Emotions: A Review and a Questionnaire Study of Everyday Listening. J New Music Res, 2004. 33(3): p. 217-238.
4. Peretz, I. & R.J. Zatorre, Brain Organization for Music Processing. Annu Rev Psychol, 2005. 56(1): p. 89-114.
5. Hunter, P.G., et al., Feelings and Perceptions of Happiness and Sadness Induced by Music: Similarities, Differences, and Mixed Emotions. Psychol Aesthet Creat Arts, 2010. 4(1): p. 47-56.
6. Maslow, A.H., Music Education and Peak Experience. Music Educ J, 1968. 54(6): p. 72-171.
7. Schäfer, T., et al., How music changes our lives: A qualitative study of the long-term effects of intense musical experiences. Psychol Music, 2014. 42(4): p. 525-544.
8. Croom, A.M., Music, neuroscience, and the psychology of well-being: a précis. Front Psychol, 2012. 2: p. 393.
9. Schmidt, L.A. & L.J. Trainor, Frontal brain electrical activity (EEG) distinguishes valence and intensity of musical emotions. Cognition Emotion, 2001. 15(4): p. 487-500.
10. Lane, R.D., et al., Cognitive neuroscience of emotion. Series in affective science. New York: Oxford University Press, 2000.
11. Burnett, S., et al., Development during Adolescence of the Neural Processing of Social Emotion. J Cognitive Neurosci, 2009. 21(9): p. 1736-1750.
12. Seashore, C.E., Psychol Music. NY, US: Dover, 1967.
13. Levitin, D.J., This is your brain on music: The science of a human obsession. NY, US: Dutton/Penguin Books, 2006.
14. Iakovides, S.A., et al., Psychophysiology and psychoacoustics of music: Perception of complex sound in normal subjects and psychiatric patients. Ann Gen Hosp Psychiatry, 2004. 3.
15. Coutinho, E. & A. Cangelosi, Musical emotions: predicting second-by-second subjective feelings of emotion from low-level psychoacoustic features and physiological measurements. Emotion, 2011. 11(4): p. 921-37.
16. Menon, V. & D.J. Levitin, The rewards of music listening: response and physiological connectivity of the mesolimbic system. Neuroimage, 2005. 28(1): p. 175-84.
17. Evers, S. & B. Suhr, Changes of the neurotransmitter serotonin but not of hormones during short time music perception. Eur Arch Psychiatry Clin Neurosci, 2000. 250(3): p. 144.
18. Bresin, R. & A. Friberg, Emotion rendering in music: range and characteristic values of seven musical variables. Cortex, 2011. 47(9): p. 1068-81.
19. Kreutz, G., et al., Using music to induce emotions: Influences of musical preference and absorption. Psychol Music, 2008. 36(1): p. 101-126.
20. Ashley, R. Affective and perceptual responses to very brief musical stimuli. In Proc. of International conference on music perception and cognition. 2008.
21. Mace, S.T., et al., Genre identification of very brief musical excerpts. Psychol Music, 2011. 40(1): p. 112-128.
22. Gorbman, C., Unheard melodies: narrative film music. London: Bloomington: BFI Pub.; Indiana University Press, 1987.
23. Phillips, W., A composer's guide to game music. Cambridge, Mass.: MIT Press, 2014.
24. Pauletto, S., Film and theatre-based approaches for sonic interaction design. Digit Creat, 2014. 25(1): p. 15-26.
25. Tweedie, L., et al. Externalising abstract mathematical models. In Proc. of CHI 96. ACM (1996).
26. Koskinen, I.K., Design research through practice: from the lab, field, and showroom. Waltham, Mass.: Morgan Kaufmann, 2012.
27. Franinovic, K. & S. Serafin, Sonic interaction design. Cambridge, Mass.: MIT Press, 2013.
28. Peretz, I. & R.J. Zatorre, The cognitive neuroscience of music. Oxford: Oxford University Press, 2003.
29. Gaver, W.W., et al., Effective sounds in complex systems: the ARKola simulation. In Proc. of CHI 91. ACM (1991).
30. Burkhard, R.A., et al., Visualization Summit 2007: Ten research goals for 2010. Inf Vis, 2007. 6(3): p. 169-188.
31. Kerren, A., et al., Workshop Report: Information Visualization: Human-Centered Issues in Visual Representation, Interaction, and Evaluation. Inf Vis, 2007: p. 189.
32. Moere, A.V., et al., On the role of design in information visualization. Inf Vis, 2011. 10(4): p. 356-371.

Beyond Detection: Investing in Practical and Theoretical Applications of Emotion + Visualization

R. Jordan Crouser
Smith College
Northampton, MA
jcrouser@smith.edu

Lane Harrison
Worcester Polytechnic Institute
Worcester, MA
lane@cs.wpi.edu

Daniel Afergan
Google Inc.
Mountain View, CA
afergan@google.com

Evan M. Peck
Bucknell University
Lewisburg, PA
evan.peck@bucknell.edu

ABSTRACT
Emotion is a dynamic variable that modulates how we perceive, reason about, and interact with our environment. Recent studies have established that emotion's influence carries to data analysis and visualization, impacting performance in ways both positive and negative. While we are still in the infancy of understanding the role emotion plays in analytical contexts, advances in physiological sensing and emotion research have raised the possibility of creating emotion-aware systems. In this position paper, we argue that it is critical to consider the potential advances that can be made even in the face of imperfect sensing, while we continue to address the practical challenges of monitoring emotion in the wild. To underscore the importance of this line of inquiry, we highlight several key challenges related to detection, adaptation, and impact of emotional states for users of data visualization systems, and motivate promising avenues for future research in these areas.

Author Keywords
Emotion, affect, visualization, adaptation, theory.

ACM Classification Keywords
H.5.m. Information Interfaces and Presentation (e.g. HCI): Miscellaneous

INTRODUCTION
Emotion is a central facet of human experience, coloring our perception and interactions with our environment. Often considered an impediment in analytical domains, emotion may steer us towards biased positions or flawed reasoning, even when the source of emotion is irrelevant to the task at hand [4]. Yet emotion is also known to play a key role in rational reasoning, allowing us to quickly recognize and act on preferences, for example, or to choose favorably among uncertain outcomes in complex planning situations. As neuroscientist Antonio Damasio summarized: "[Emotion] allows the possibility of making living beings act smartly without having to think smartly" [4].

EmoVis 2016 Workshop, March 10, 2016, Sonoma, CA, USA.
Workshop proceedings published by Linköping University Electronic Press.

Given the pervasive role emotion plays in human behavior, it is plausible that emotion would also impact how we analyze data and interact with data visualization systems. Several studies have already begun to substantiate this connection. Recent experiments from Harrison *et al.*, for example, combined affective priming with a classic visualization perception experiment and found that emotions with positive valence (*i.e.* happiness) led to better performance when analyzing charts [7]. Other studies in data visualization and human-computer interaction have pointed to emotional states as a possible explanation for their results [1, 2]. Researchers have specifically suggested connections between emotion and visualization performance in the context of task engagement, memorability, as well as higher level facets such as creativity and decision-making.

These findings raise the question of whether it is possible to design systems that reliably detect emotion and respond to it in an intelligent way. Detection remains challenging, however, even after decades of research investigating diverse information sources and analysis techniques. Common approaches include analyzing facial features, posture, and various physiological signals. Yet none of these approaches work perfectly, nor are they available in all contexts in which data visualization systems are used.

We argue that even moderately accurate sensing techniques, such as today's consumer-grade physiological sensors, can cater to the user in beneficial ways through proper adaptation techniques. Over the past decade, studies using brain sensing have yielded many compelling examples of beneficial adaptation with imperfect sensors (for a review, see [11]). This success is largely due to the development of better models of adaptation alongside better consumer-level hardware and detection algorithms. Taken together, these elements allow system designers to readily connect user tasks and goals with adaptation strategies in creating emotion-aware systems.

Breakthroughs in sensors that detect the body's natural signals allow the user to wear lightweight sensors while having normal interactions with a computer and make physiological sensing possible at consumer-level scale. *Physiological computing* "has the potential to extend the communication bandwidth of

HCI and enable a dynamic, individualised dialogue between user and system," [6] without any effort on the part of the user. By monitoring user biological signals, we can extract information about the user's cognitive state and use this as a system input to control the visual elements onscreen.

In order to create systems that successfully interpret and respond to human emotion, the research community must develop robust models to describe the dynamic influence of emotional state on visualization performance. At the same time, it is critical to consider the potential advances that can be made even in the face of imperfect models and sensing modalities. In this position paper, we highlight several key challenges related to detecting and adaptating to human emotion. We also discuss the impact of emotional states for users of data visualization systems, and motivate several promising avenues for future research in these areas.

DESIGNING WITH IMPERFECT INPUT

One of the challenges of defining the design space of emotion and visualization is that we often turn to interaction mechanisms that necessitate high levels of accuracy. For example, biofeedback applications are valuable only to the extent that they are accurate. Providing users with visual feedback that actively works against their intuition may cause users to lose a sense of control with the interface, reducing their trust levels. This places innovation in a difficult place - should we wait for detection mechanisms to reach a certain threshold before focusing on certain applications?

Instead, we draw on our experience designing passive brain-computer interfaces (BCIs) with physiological monitoring. Looking more broadly at user state, we built several adaptive applications that relied on unreliable classifications of user workload [9]. The accuracy of our models varied widely between individuals, and there was no way to get ground truth and figure out when the model was misclassifying the user. These systems focus on *implicit input*, user contexts that the system knows is input but that the user does not actively choose to share with the system [10]. However, despite these challenges, our systems quantifiably improved user interaction. This enables us to explore the real-world challenges of passive BCI despite being years away from robust sensing of cognitive state.

We believe that it is critical for researchers to explore the interaction space of emotion + vis *now* instead of waiting until models improve beyond an arbitrary threshold. Because laboratory or consumer-grade physiological monitors are prone to noise and artifacts, and emotion is a complex construct with multiple dimensions that are difficult to capture, we must consider and anticipate the inevitability of these misclassifications, and continue to advance the field using strategies to minimize their impact.

One promising direction for dealing with imperfect input is sensor fusion. Sensor fusion systems integrate multiple physiological sensors into one classification model, with the hopes of improving accuracy or specific inferences that could not be determined with only one sensor [8]. By using techniques such as moving averages, confidence value of predictions (label smoothing), or hybrid sensor fusion, we can increase the reliability of classifications and limit the harm of a small number of misclassifications. As long as we are fairly confident of user state before triggering an adaptation, we can limit the rate of mistakes and provide an overall gain to the user.

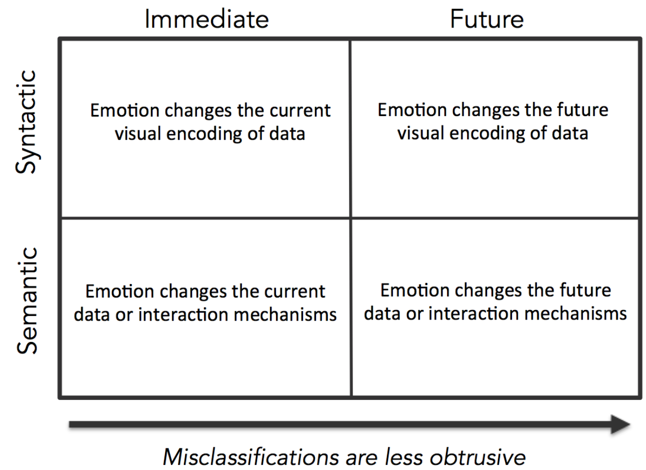


Figure 1. Framework of adaptive systems.

WHEN CAN WE RESPOND TO EMOTIONS?

The challenge of identifying emotion in the wild is not trivial. Because of this, there is a danger that the focus of the community overemphasizes the improved detection of affect, with little to show for itself on the other end. To counteract this, we also need to contextualize those improvements within the goals of the visualization or interface: how will better sensing lead to *meaningful improvement* in people's analytical tasks and workflows?

As we move forward applying emotions to visualizations, we can look to the physiological computing community for lessons. Stephen Fairclough has this to say about the state of the art: "Constructing a system that can detect a range of psychological states is pointless if [the] adaptive repertoire of the machine is unable to respond to those psychological states in an intelligible fashion" [5].

We want to make sure that the user is not perturbed by the adaptations a system makes, and that the adaptations be resilient to misclassifications. Because the prediction might not always be correct, a visualization designer should view the physiological input as an augmentation to traditional input devices (keyboard, mouse, touchscreen), and not as the main source of input [11]. The designer should avoid irreversible or mission-critical adaptations, and instead make subtle, helpful changes, that the user might not even recognize or attribute to the system adaptations. Zander proposes that these passive systems can be evaluated along three key dimensions: *complementarity*, or lack of interference with other input mechanisms; *composability*, or potential to stack with other monitors; and *controlled cost*, or the effort of calibration and price of mispredictions [12].

Categorizing Implicit Interactions with Emotion

Building upon the framework for implicit interaction in passive brain-computer interfaces by Solovey et al. [11], we can begin to consider how emotion can be used to potentially drive mixed-initiative systems for visualization. We categorize adaptive mechanisms into four categories - immediate syntactic, immediate semantic, future syntactic, and future semantic - as shown in Figure 1 and discuss applications in visualization.

- **Immediate syntactic:** Emotion is used to change or refine the visual form of the data. This can be accomplished by zooming, filtering or other visualization tasks. For example, a system might detect that a user is frustrated trying to understand data of a particular form, and dynamically alter its representation to aid understanding. Or a visualization could "tag" data with emotional responses as a form of meta-data to be used in later exploration or analysis.

- **Immediate semantic:** Emotion is used to trigger changes in the interaction mechanisms in a visualization. If the user seems frustrated, we can zoom in on a particular subset of data to allow her to focus on a small subset of high-priority information.

- **Future semantic:** A person's emotion is used to inform the system as to what data should be shown next. For example, based on a person's emotional response while engaging with a subset of datapoints, a system might suggest visualizing other data that is related or contains similar properties.

- **Future syntactic:** A person's emotion is used to inform the visual encodings of data seen in the future. For example, an intelligent system might learn that particular affective states modify interaction with a visualization. During a later interaction, the system could recommend a more optimized representation of the same data.

By expressing this design space more explicitly, we may be able to identify opportunities for innovation even as we continue to develop our models for detecting emotion. Looking at the classification above, we identify opportunities to provide for the value for the user, but only if we do it in unobtrusive ways.

For example, constant updates across the entire interface or changes in the display format may be jarring and unsettling for users and disrupt their ability to form cohesive mental models of the system. While immediate adaptations have potential, they requires a high level of accuracy since the adaptations occur directly within a person's focus. This is particularly dangerous in the context of visualizations since our perceptual system is sensitive to changes in certain visual features such as movement or color changes.

Future changes can be effective because we can completely change the system without alerting or disrupting the user. However, it may be more difficult to predict what state the user will be in when the change occurs, making it challenging to evaluate the efficacy of these changes.

MODELING EMOTION'S ROLE AND IMPACT

In addition to the challenges raised in determining when emotion plays a role in data visualization, we presently lack appropriate mechanisms for evaluating the systems we build. In-house experimentation and in situ studies can help us determine whether or not our systems are useful, but they fall short of explaining *why* we see the results we do. This often results in the recycling of known techniques; generalizing our results to new domains is challenging, and we are left to speculate about the role of emotion in producing observed effects on user behavior. To overcome this, it is important that we develop theoretical models and corresponding language of the role of emotion in visualization systems to improve our ability to reason about their performance and design.

In visualization, and particularly in the arena of incorporating and designing for human emotion, interest in the development of real-world implementations [3] has far outpaced the development of theoretical measures. The current trajectory for research in emotion+visualization is largely focused figuring out how to leverage emotion in visualization in ways that measurably impact performance. However, in the absence of a rigorous theoretical framework in which to ground the development of new algorithms, researchers must rely on intuition and some deeply-rooted assumptions about the role of human emotion in order to design new systems. Using tacit knowledge regarding emotional responses to which we believe humans are predisposed, we build systems that capitalize on these responses. These systems will then be used as evidence that the chosen method for adapting to (or exploiting) human emotion works; a sort of "proof-by-example".

We argue that developing a theoretical language for describing the role of emotion is of critical importance to the study and design of visualization systems. Mechanisms for drawing parallels at the algorithmic level and identifying areas where existing approaches are redundant or inefficient will enable us to design more effective systems in the future. In addition, reporting theoretical arguments along with the observed performance of the system would greatly improve study reproducibility, as well as help isolate the effects of interface design and other implementation details.

The importance of understanding human emotion as part of a larger computational framework is not limited to improving visualization design. Augmenting our existing models to incorporate human emotion can expand our understanding of what can be computed, as did the development of probabilistic and parallel computation. The development of theoretical measures for human emotion may play a significant role in the broader acceptance of emotion as an important component of visualization systems.

CALL TO ACTION

Over-emphasizing the detection of emotion is tempting. Models of emotion can be clearly evaluated, and progress can be quantifiably defined by improvements in classification accuracy. However, we should avoid the trap of improving emotion models just for the sake of improving the model. We propose that by bolstering our efforts on the theoretical foundations and practical applications, we can identify high-impact applications of emotion, and begin to answer the question of "how good is good enough?" with respect to our models. We may discover that the most compelling intersections of emotion and visualization do not in fact require high accuracy to achieve the desired results.

From an applied perspective, we should be careful not to view emotion exclusively as a tool for evaluating visualization. While this application is compelling, it requires high levels of trust and accuracy in our interpretation of signals - something that may not be reliable for years. Instead, we believe that focusing on the design space of *emotion-driven* applications has the potential for significant impact even using models with modest accuracy. In the near term, we can consider the applications of implicit interaction on future modifications of a visualization, both at the semantic and syntactic levels. By using emotion to inform the design or presentation of future data presented to the user, we can build personally attentive portraits of data that do not necessarily rely on robust models. Finally, investing in practical applications has the added potential of capturing the imagination of other researchers or developers, propelling this young field forward.

From a theoretical perspective, we need to internalize that emotion exists as part of a larger human-machine interactive system, in the context of a specific task and environment. We cannot consider human performance and cognitive workload as a static measure that is unaffected by the human's current state. The role of emotion cannot be ignored or be considered in a vacuum, nor is it only one-dimensional. Instead, we must strive to consider all aspects of the closed-loop system, and strive to make modifications in real-time as a user's emotional state changes and as the task requirements change.

In assessing the merits of various tools for developing emotion-aware systems, we should not allow the perfect to become the enemy of the good. At the same time, we challenge this interdisciplinary community use these systems as a lens to better understand *how humans fit* into the grand scheme of computational tools, and to develop models that incorporate not only their power but their inherent messiness as well.

REFERENCES

1. Scott Bateman, Regan L Mandryk, Carl Gutwin, Aaron Genest, David McDine, and Christopher Brooks. 2010. Useful junk?: the effects of visual embellishment on comprehension and memorability of charts. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2573–2582.
2. Nick Cawthon and Andrew Vande Moere. 2007. The effect of aesthetic on the usability of data visualization. In *Information Visualization, 2007. IV'07. 11th International Conference*. IEEE, 637–648.
3. Daniel Cernea, Andreas Kerren, and Achim Ebert. 2011. Detecting insight and emotion in visualization applications with a commercial EEG headset. In *SIGRAD 2011 Conference on Evaluations of Graphics and Visualization-Efficiency, Usefulness, Accessibility, Usability*. 53–60.
4. Antonio Damasio. 2008. *Descartes' error: Emotion, reason and the human brain*. Random House.
5. Stephen Fairclough. 2015. We Need To Talk About Clippy. (2015). http://physiologicalcomputing.org/2015/03/we-need-to-talk-about-clippy/
6. Stephen H Fairclough. 2009. Fundamentals of physiological computing. *Interacting with computers* 21, 1 (2009), 133–145.
7. Lane Harrison, Drew Skau, Steven Franconeri, Aidong Lu, and Remco Chang. 2013. Influencing visual judgment through affective priming. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM, 2949–2958.
8. Danilo P Mandic, Dragan Obradovic, Anthony Kuh, Tülay Adali, Udo Trutschell, Martin Golz, Philippe De Wilde, Javier Barria, Anthony Constantinides, and Jonathon Chambers. 2005. Data fusion for modern engineering applications: An overview. In *Artificial Neural Networks: Formal Models and Their Applications–ICANN 2005*. Springer, 715–721.
9. Evan M Peck, Daniel Afergan, and Robert JK Jacob. 2013. Investigation of fNIRS brain sensing as input to information filtering systems. In *Proceedings of the 4th Augmented Human International Conference*. ACM, 142–149.
10. Albrecht Schmidt. 2000. Implicit human computer interaction through context. *Personal Technologies* 4, 2-3 (2000), 191–199.
11. Erin Treacy Solovey, Daniel Afergan, Evan M Peck, Samuel W Hincks, and Robert JK Jacob. 2015. Designing implicit interfaces for physiological computing: Guidelines and lessons learned using fNIRS. *ACM Transactions on Computer-Human Interaction (TOCHI)* 21, 6 (2015), 35.
12. Thorsten O Zander, Christian Kothe, Sabine Jatzev, and Matti Gaertner. Brain-Computer Interfaces. In *Brain-Computer Interfaces (Human-Computer Interaction Series)*. Springer London, 181–199.