An Approach to Analysis of Physiological Responses to Stimulus
From Electrodermal Activity to Combined Physiological Responses

Reza Tasooji, Nicole Buckingham, Denis Gračanin$^{(\boxtimes)}$, and R. Benjamin Knapp

Department of Computer Science, Virginia Tech,
2202 Kraft Drive, Blacksburg, VA 24060, USA
`gracanin@vt.edu`

Abstract. Recent advances in wearable devices capable of measuring physiological signals such as Electrodermal Activity (EDA) support affective computing and related applications. We present an algorithm that uses EDA signals to detect highlights of a stimulus. To test the accuracy of our method, two different mediums, a scene from a movie, as our ground truth, and a scene from a video game, for testing the algorithm, were selected. We conducted a user study with 20 participants, analyzed the differences between mediums and validated the accuracy of our method for detecting the highlights of the stimulus using only EDA signals. Our approach uses commonalities among users based on their phasic responses for detecting highlights. The result of the study shows a F1 score of 0.93 and 0.89 for movie and video game respectively. We are in a process of conducting a user study with several sensory devices to explore combined physiological responses using IAPS data set as a stimulus.

Keywords: Emotion recognition · Affective computing ·
Human Computer Interaction · Electrodermal Activity · Fear

1 Introduction

Our emotions play an important role in our decision making and the outcome of our decisions can change the emotions we experience. This interplay between emotions and decisions plays a crucial role when defining normal and abnormal human behaviors.

Emotion is a sensory or behavioral reaction that is caused by any object or event. In psychology, the term "affect" describes the experience of emotion. The object or event that causes such changes is called "stimulus". The two fundamental dimensions of emotion are valence and arousal. Various studies show that these dimensions are related to physiological aspects and brain activities [14, 18, 26].

Extensive research in this field has revealed that it is possible to detect emotion through physiological signals. One study, done by Feldman [11], showed that the degree of arousal is related to the amount of physiological activities during affective experience. Additionally, Chênes et al. [6] stated that physiological signals are difficult to hide or fake and with appropriate biometric sensors, it can be possible to detect the presence of emotion.

For many years, medical and health services have used biometric sensory devices to monitor the behavior of many human organs, such as the heart and brain. These sensory devices are now becoming available to the general public in different forms, including wristbands. Wristband-like sensory devices aid in data collection and can be easily integrated into the process of analyzing physiological signals.

In the field of human computer interaction (HCI), the use of such devices can help enhance user interaction by removing the cognitive load, through solely analyzing affective responses. The current databases used for the majority of emotion recognition systems, such as IAPS [20], are heavily dependent on participants' self-reports. Our goal is to explore methods for detecting highlights of the video stimulus that might cause affective responses, without fully relying on users' self-reports.

2   Related Work

The process of automatically detecting the highlights of a video is a far-reaching subject. With the huge variation in genres and different mediums such as video game, movies, virtual reality (VR) and augmented reality (AR), defining what is considered as a highlight can be diverging. Several studies [6,16] show that it is possible to detect highlights of the stimulus with high accuracy using physiological changes, but designing such a system comes with limitations and it will not be able to handle all possible outcomes.

Studies by Levenson [21,22] show that it is possible to distinguish emotion based on physiological responses. Specifically, fear stimulus will cause significant metabolic demands on the heart, which increases the maximum heart rate, and fear and disgust will cause higher skin conductance compare to happiness.

Joho et al. [16], analyzed facial activities to detect personal highlights of the content. This team developed a real time facial expression recognition system that outputs a vector of motion features of certain regions of the face. In a user study, they used eight video clips from different genres with varying duration and showed the content to ten participants. To extract the highlight, the team searched for a high level of consensus on personal highlights for the different videos. Results revealed one video had a common highlight among all users except one.

Unlike speech recognition and facial expression [7,9,10,12], using biosensors to measure and analyze physiological signals for identifying emotion is not well researched. Chênes et al. [6] present a technique to obtain user-independent summary of a given video. This approach did not require emotion recognition.

Instead, the system concept was based on the physiological linkage between different participants' emotional responses. Four physiological signals were used in the study: Electromyogram (EMG) which measures the activities of the muscles, Blood Volume Pulse (BVP), measures the change of blood pressure, Electrodermal activity (EDA) and skin temperature. The results showed that skin temperature with a window response of eight seconds returned the best correct classification rate (77%) and the EDA signal can be considered the most promising signal for detecting highlights in video clips.

EDA signals are a great candidate to detect highlights; EDA devices are easy for users to wear and it that provides sensory data with high accuracy. Two main components of an EDA signal are tonic and phasic responses. Tonic response, also known as Electrodermal Level (EDL), is the slow background changes. It can be described as slow changes that happen in absence of stimulus or during response-free recording intervals. Phasic response, also known as Electrodermal Response (EDR), is a distinct response to a stimulus [8].

We are using the method proposed by Benedek and Kaernbach [2] to deconvolve EDA signal into two contentious signals of tonic and phasic responses. This method is based on the precondition that there exists a stable Impulse Response Function (IRF). The assumption is that sudomotor nerve activity will show peaks, known as sudomotor bursts, with short time constants which leads to a larger time constant exhibition in EDRs. IRF represents the basic EDR shape that would result from a unit impulse. Further, we focus on the phasic component of EDA signal to find the highlight of the video stimulus.

To validate our model, we used video stimulus that might cause fear. Fear can be described as an evolutionary necessity that notifies a person to proceed on their current direction or take another course of action in order to increase the likelihood of survival. Excessive fright can cause a cognitive dissonance state, which can cause an attitude change, or inconsistent thoughts, that might turn anxieties into phobias. Additionally, fear is an essential emotional response to societal conflicts. For example, the fearful expression is thought to serve as a social cue [24].

The result from the study by Trubanova et al. [31] shows the positive correlation between recognizing fear and perspective-taking abilities, while perspective-taking was not significantly associated with recognition of the other basic emotions. The result shows there is something unique about fear that might separate it from other basic emotions.

A study by Lynch [23] examined fright experiences caused by video games. This study was based on a self-report by 269 undergraduate students; no physiological signal measurements were included in this study. The results showed that elements such as darkness, zombies, being surprised, etc. in video games are causes of fear. There was no significant gender-based differences in experiencing fear. These studies demonstrate that fear could help to detect the cause of affective response with more confidence.

3   Problem Definition

The exponential growth of video media increases the necessity of algorithms to summarize the content of the video. Personalized recommendations from content providers, such as Netflix's "Top picks for you", helps consumers watch the content that they care about and helps advertisement companies by providing them with better insight on what gets watched, and by whom. Additionally, VR and AR are finding their way into the entertainment industry. This industry is growing more than ever and with many variations of this medium, the consumers want the content to be their way.

One approach to put the consumer first in selecting the content is to build a system capable of automatically detecting what a consumer wants to see based on their physiological changes. Detecting the highlight of the media and analyzing the content can help us to take the first step for building a system that is capable of selecting the content based on the users' current affective response.

One of the problems with current approaches for detecting highlight is that the process for collecting signals is often laborious and exhausting for average users and, consequently, not feasible for daily tasks. To solve this problem, we are using a wristband capable of recording physiological signals, thus making the process of collecting physiological data much easier.

Most of the studies in the field of affective computing use multiple physiological signals. Although this approach increases the accuracy of the results, it requires more processing and it is not always possible to analyze the data in real time. Our assumption in this study is that EDA signals have the essential parameters such as tonic and phasic responses that are sufficient for detecting highlights of the stimulus. Our focus is on stimulus that might arouse fear in order to reduce the number of independent variables, gain a better insight and provide consistency among all subjects.

To test our approach, we conducted a user study with two stimulus that might arouse fear. We selected a scene from the movie "Halloween" as our ground truth for our method. The result of the study by Philippot [27] shows that the selected section of this movie can arouse fear. Additionally, the selected scene has elements of surprise, anticipation, loud sound and music that can be considered as highlight of the video.

To test the accuracy of our model, we selected another scene from a video game "Evil Within" [1] that might arouse fear but with all elements that has been reported by Lynch et al. [23] that might cause fear. The model needed to be able to detect these elements as the highlights of the video content that might cause affective response by analyzing EDA signal. Our goal for selecting a video game was to select a different medium, unrelated to movie, while sharing the same highlights.

In this study, highlights are any event in the stimulus that might cause physiological changes. This includes any sudden change in sound volume, change of scene or any content that might surprise the audience.

4   Methods and Techniques

In our previews work [30] we explored a real-time system for processing physiological signals. We divided the system into two parts: collector and analyzer. The collector provides the infrastructure for gathering the data when conducting the user study. The purpose of the collector is to store streamed data and synchronize the input data with the video stimulus. Before sending the data to the processing unit, the physiological data and the information about stimulus needs to be transferred to a local machine to be synchronized using the same clock time of the local machine. The process of synchronization involves checking the quality of the recorded signal and syncing the input sensory data with the video stimulus. The synchronized data then will be send to the analyzer. The purpose of the "Analyzer" is to run the algorithm and report the highlight of the stimulus.

Figure 1 illustrates the system. The top section of the figure shows the process for collecting the data and the bottom section demonstrates how we process the signal for detecting highlights. Different sub units, such as a clustering unit, can be added in the processing output layer for later studies.

The biometric sensor that we are using for this study is E4 wristband by Empatica [13]. One of the main advantages of using the E4 compared to other devices is that Empatica E4 is easy to use. This makes it possible for people to use the device with ease during daily activities. Also, the study by Ragot et al. [29] compares Biopac MP150 to Empatica E4 in terms of emotion recognition accuracy using Support Vector Machine as their learning algorithm. The result from the study shows there is not significant differences in accuracy between two devices in recognising emotion. One of the capabilities of E4 device is to measure electrical conductance of the skin in the $[0.01, 100]\mu S$ (micro-Siemens) range with digital resolution of 1 digit per 900 pS (pico-Siemens) [13].

5   Experiment Design

The first step is to provide the stimulus that might arouse emotion. In the review article about how to measure emotion by Mauss [25], the author concludes "there is no gold standard measure of emotional responding. emotions are constituted by multiple, situationally and individually variable processes".

Each stimulus includes three parts (Fig. 2). The first part (30 s) named as "Baseline" is a black screen without any sound, to collect the baseline of the user. The second section (90 s) is called "Relaxed", and the purpose is to ease the user. The content for this section was selected from the video game named "Flower" (2009). Finally, the "Fear" section is the "Halloween" movie scene for movie stimulus and "Evil Within" for video game stimulus (210 s).

To be able to detect the highlights in both movie and video game, we removed any element of interaction between user and the content by asking participants to watch a prerecorded section of the video game, passively, instead of asking them to play the content of the video game.
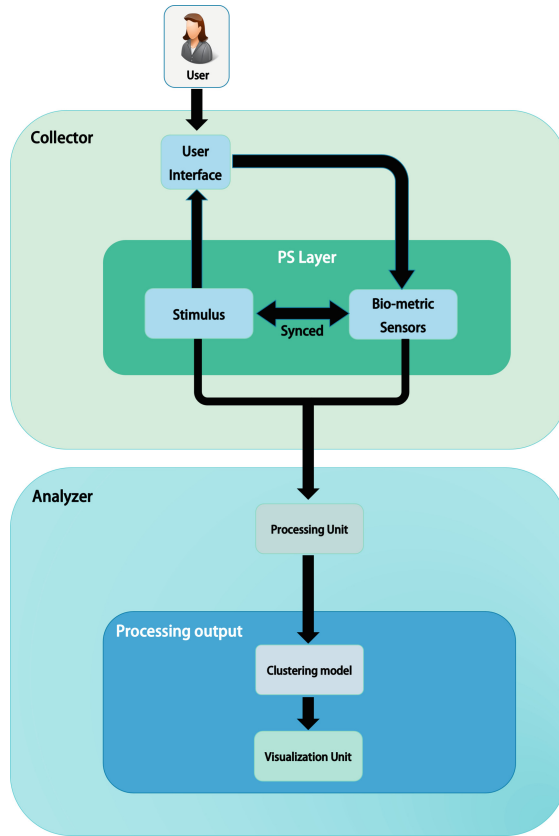
Fig. 1. The system is divided into two subsystems. The collector subsystem is capable of storing and streaming different biometric sensory data. After the data is synced with the video stimulus, the data will be send to analyzer subsystem. Different processing units can be added to analyze input signals.
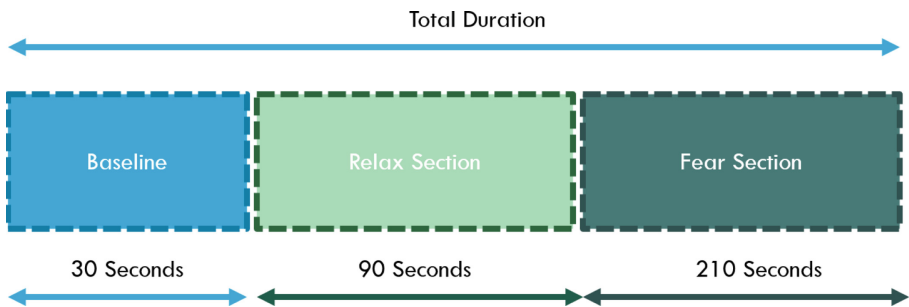
Fig. 2. Three parts of the video stimulus.

5.1   Running the Experiment

A total of 20 users (female = 8, male = 12), with ages between 18 to 40, ($\mu = 27.15$, $\sigma = 6.62$) participated in the study. All the participants were college students; 7 were undergraduate and 13 were graduate students. All the participants but one reported that they had not seen the content of the stimulus before the study.

Each section includes pre-and-post questionnaires. The pre-questionnaire includes questions about age, gender and nationality of the participants. The post-questionnaire includes six questions asking for a rank from 1 to 5. The goal of the post-questionnaires is to be able to compare two mediums and find out the differences between them and if there is any differences, does physiological signals capable of showing this differences. Additionally, the mean of the answers can be considered as a score that shows how strong the content is to arouse emotion.

Knowing that both stimulus might cause fear, our hypothesis is there is no significant differences between two mediums. From this hypothesis we can then explore the performance of our proposed model of detecting the highlights.

The content of the stimulus can cause fear based on what is considered as scary in entertainment industry. For ethical reasons, we encouraged participants to withdraw from the study at any stage if they find the content too strong. There was no withdrawal from the study.

6   Data Processing

During the study we measured EDA signal using Empatica E4 wristband and after the study we asked the participants to fill the post-questionnaires. The recorded data has the frequency of 4 Hz with total duration of 330 s for each section. The required steps for processing the raw data follow.

6.1   Standardizing the Data

One problem with quantifying EDA signal is the existence of large variability because of individual differences. For example, the amplitude of 0.5 μS EDR might be high for one person while it may be the baseline for another person. To correct these differences, the process of standardization is required.

There is no universal approach on methods to standardize EDA signals. For our system, we use *transforming raw data into Z-scores* [4] to standardize a raw data (Eq. 1).

$$X_n = \sum_{n=1}^{N} \frac{x_n - \bar{\mu}}{\bar{\sigma}} \tag{1}$$

where x_n is the value of raw signal at time n, $\bar{\mu}$ is the mean of the raw signal, and $\bar{\sigma}$ is the standard deviation of the raw signal.

After standardizing raw EDA data, we decompose it into phasic and tonic responses using the continuous decomposition analysis (CDA) method [2].
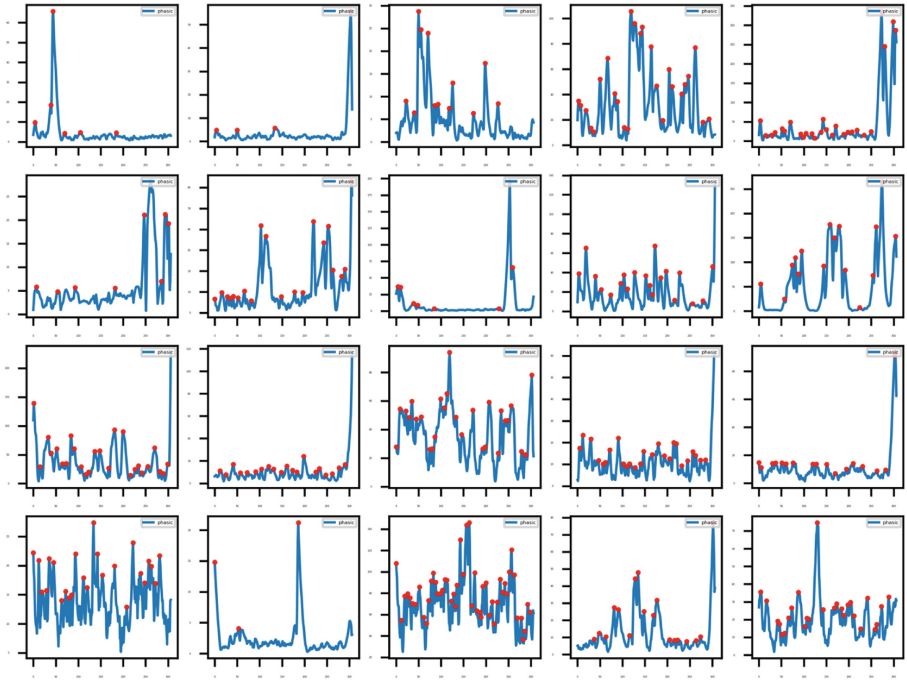
Combined data from 20 subjects

Fig. 3. Detected peaks from Z-score of the phasic response for 20 subjects. The x axis is the duration of the stimulus while y axis is the Z-score of the phasic response.

6.2 Detecting Highlights

We considered *Area Under the Curve* and *Sum of Squares*, with the window size of three seconds [4] and moving window of one second, as possible features for detecting highlights. After exploring and analyzing user responses of each, we decided to use *Area Under the Curve* for phasic responses as our main feature for detecting highlights.

After calculating this feature for all users, we detect and locate the peaks of the feature for all users. Figure 3 shows the detected peaks for 20 users. The x axis is the total duration of stimulus, while the y axis is the Z-score calculated from phasic response.

The histogram of the calculated peaks can be used to locate the time of the stimulus (Fig. 4). Specifically, bins with higher values show a larger group of people having higher phasic responses during bin's duration. By adjusting the threshold for the number of users in the bins, we have a selection of bins that can be analyzed further for detecting highlights.
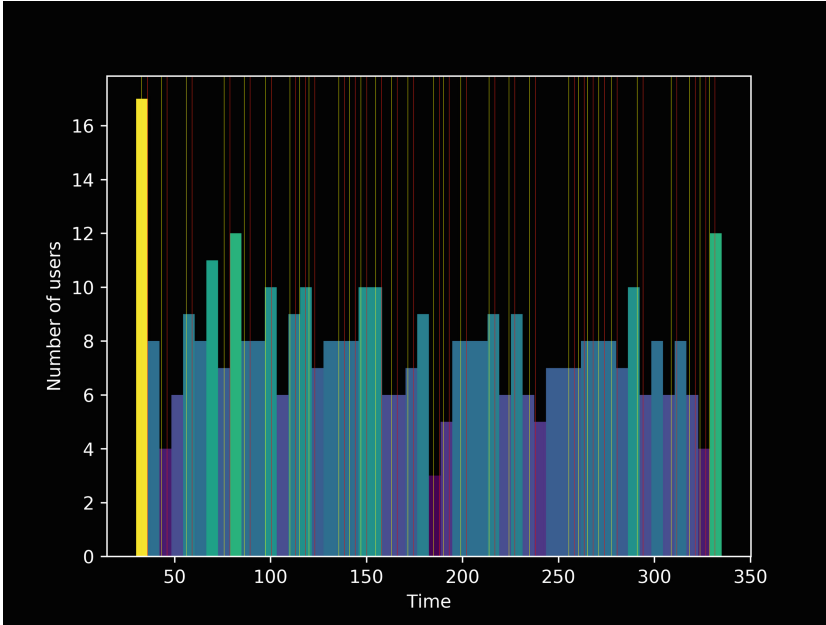
Fig. 4. Histogram from the calculated peaks.

To analyze the bins, we calculate the *Jaccard similarity* between bins within the same level of users. **Jaccard similarity coefficient** [15] is a method for measuring similarity of a sample set. It is defined as the size of the intersection over the size of the union of the set (Eq. 2). The result from Jaccard dissimilarity is between zero and one, and the value closer to one represents higher similarity. Jaccard distance was used because it is a good indicator to measure commonality between users.

$$J(A, B) = \frac{|\, A \bigcap B \,|}{|\, A \bigcup B \,|} \tag{2}$$

By defining the threshold for Jaccard dissimilarity, we can select bins on the same level of the histogram that have similarities to each other. We define a graph to connect similar members on the same level together. For every similar bin, we find the intersections between two bins. Using stack data structure, we push the intersections onto the stack, then we pop one member and connect that member to the remaining items in the intersection and add the connected edge to the graph. We repeat the same process for all the bins in different levels.

The graph with the highest number of edges has highest probability to locate the time that might be considered as highlight. By defining a threshold for the number of edges in the graph, we can use bins that might include some highlights from the stimulus. Algorithm 1 shows the steps taken to report the highlight.

Figure 5 shows the result as a karate graph [32], from the histogram of size 30 bins. In this example, bins with a size bigger than four were considered as bins

---

Algorithm 1. Reporting highlights.

---

input : Bins at the same level of the histogram, Jaccard Dissimilarity
        Threshold, Edge density Threshold
output: Possible time stamp of the highlights

for $bin \in InputBins$ **do**
  for $bin \in InputBins$ **do**
    Calculate Jaccard Dissimilarity between bins
  end
end
graph ← new Graph()
while *not at the end of bin* $\in InputBins$ **do**
  CurrentBin ← bin
  for $bin \in InputBins$ **do**
    if $jaccardDissimilarity(CurrentBin, bin) \geq JaccardThreshold$ **then**
    intersection ← $Intersection(CurrentBin, bin)$
    while $length(intersection) \geq 1$ **do**
      CurrentPerson ← intersection.pop()
      for $members \in intersection$ **do**
        graph.addEdge(CurrentPerson, members)
      end
    end
    end

  end
end
if $graph.EdgeNumber \geq EdgeDensityThreshold$ **then** Report timestamp of
the bins in the level.

---

with potential highlights in them. The density of the edges in each graph, shows the similarity between bins in that level. Bins with higher number of edges have higher probability to be considered as highlight of the stimulus.

The three threshold that we use in this approach are as follow:

– Number of bins in the histogram
– Jaccard dissimilarity threshold
– Edge density threshold for karate graph

Different threshold levels can provides different level of certainty. For example, with low number of bins, low Jaccard value and low edge density threshold, the timestamp of highlights may point to the whole duration of the stimulus. Figure 5 shows the output with total number of 30 bins, with 0.45 Jaccard dissimilarity threshold and edge density of 40.
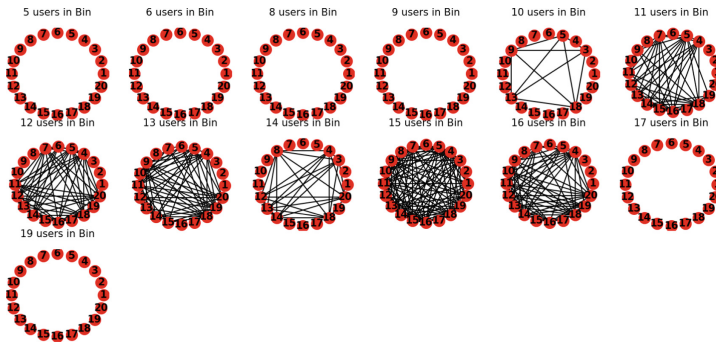
Fig. 5. Karate graph shows the timestamp of the highlights based on the number of the edge in each bin.

7   Result and Discussion

To implement our method, we assumed EDA signal by itself is sufficient for detecting highlights of the stimulus without using other physiological signals. Further, the phasic component of EDA signal played an important role in detecting the highlights.

We selected two different medium, "Halloween" movie as our ground truth and "Evil Within" as our testing stimulus. To explore the differences between the two mediums, due to the small sample size and, therefore, uncertain underlying distribution of the data, we considered both parametric and non-parametric approaches for analyzing the data.

We used paired t-test for parametric and Wilcoxon test for non-parametric approach. The result for both approaches shows that participants are more engaged with the video game content compare to the movie (p-value < 0.05). Other than this, as we expected, there are no significant differences between two contents.

Table 1 shows specific means and standard deviations for each questions posed to participants after the study. The questions were ranked between one to five, where a value of 1 represented negative feedback to the question and a value of five represented positive feedback.

Table 2 shows the overall mean and standard deviation for all questions. The overall mean for the movie and video game can also be considered as the rank for how strong the content was. A value closer to 5 indicates the content might have higher emotion arousal on the users.

We can conclude, overall, the two stimulus are very similar to each other. The unique characteristics in the video game, for example using camera in first person perspective, caused the users to become more engaged to the content. Regardless, the mean and standard deviation for "Being in Control" shows that the level of interaction for both stimulus are identical.

Table 1. Mean and standard deviation between movie and video game.

	Movie		Video Game	
	Mean	Std	Mean	Std
Being engaged and involved	3.40	0.99	4.00	0.79
Negative or positive feeling	2.80	0.69	2.75	1.07
Being passive or active	2.75	1.02	2.95	1.14
Being relaxed or tensed	3.50	0.94	3.10	0.23
Dislike or like	3.15	0.18	3.15	0.26
Being in control	3.2	0.28	3.2	0.26

**Table 2.** Overall mean and standard deviation between movie and video game.

Movie		Video game	
Mean	Std	Mean	Std
3.13	0.30	3.19	0.42

Table 3. Highlight detection including all signals for movie.

Number of bins	Movie					
	Jaccard threshold = 0.0			Jaccard threshold = 0.4		
	Precision score	Recall score	F1 score	Precision score	Recall score	F1 score
100	0.93	0.84	0.88	0.93	0.84	0.88
90	0.89	0.80	0.84	0.90	0.93	0.92
80	0.93	0.87	0.90	0.96	0.83	0.89
70	1.0	0.74	0.85	0.95	0.74	0.83
60	0.93	0.90	0.92	0.96	0.93	0.95
50	1.0	0.87	0.93	1.0	0.87	0.93
40	1.0	0.84	0.91	1.0	0.83	0.91
30	1.0	0.64	0.78	1.0	0.74	0.85

The result for detecting the highlights based on the commonalities among bins is displayed in Fig. 6 using a karate graph. For each bin, we selected the top three graphs with the highest number of edges. From the selected graphs, we extracted the time stamp of the bins in that graph. Finally, we compared the extracted time stamps to our pre-defined time stamps for highlights.

The result show, for both movie and video game stimulus, the reported highlights point to the locations that can be considered as highlights. These highlights include sudden change of scene, starting of loud sound and element of surprise, that was common in both stimulus.

Number of Bins	Movie	Video Game
30		
40		
50		
60		
70		
80		
90		
100		
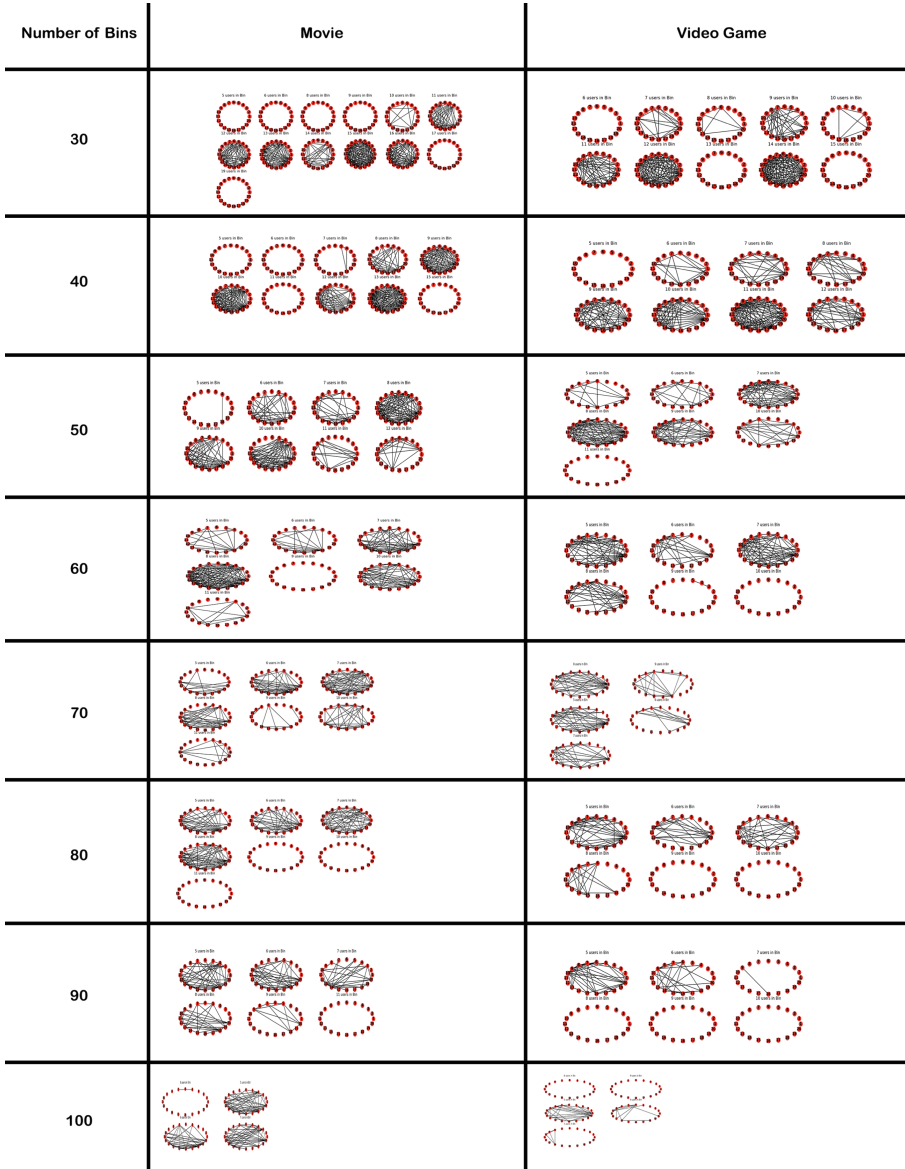
Fig. 6. Karate graph for measuring commonalities among size of bins for different number of bins.

Tables 3 and 4 show the precision, recall and F1 score for each bins, with and without including Jaccard threshold. Looking at the result, the two bins, number 60 for movie and number 50 for video game, show the Jaccard threshold caused improvement in F1 score. The selected threshold for Jaccard dissimilarity shows some improvement but overall, it is not significant.

Table 4. Highlight detection including all signals for video game.

Number of bins	Video game					
	Jaccard threshold = 0.0			Jaccard threshold = 0.4		
	Precision score	Recall score	F1 score	Precision score	Recall score	F1 score
100	0.96	0.67	0.79	0.92	0.7	0.8
90	0.93	0.75	0.83	0.9	0.78	0.84
80	0.96	0.78	0.86	0.96	0.81	0.88
70	0.92	0.83	0.87	0.96	0.78	0.86
60	0.91	0.86	0.89	0.89	0.95	0.92
50	0.96	0.75	0.85	0.97	1.0	0.98
40	0.96	0.67	0.79	1.0	0.62	0.76
30	1.0	0.73	0.84	1.0	0.7	0.84

Table 5. Highlight detection per individual.

Movie			Video game		
Average precision	Average recall	Average F1 score	Average precision	Average recall	Average F1 score
0.43	0.27	0.31	0.47	0.21	0.28

To further test our method, we ran the method for per-individual signals instead of including all the signals. Table 5 shows the average precision and F1 score among all participants.

The result from our study show our method has high accuracy when including EDA signals among all users and low accuracy when using per individual. The precision of this methods depends on the size of the bins and number of edges in karate graph.

By analyzing the time interval between each pre-defined highlights, the mean of 9.85 with a standard deviation of 4.51 for movie stimulus and mean of 8.65 with standard deviation of 5.70 for video game stimulus provides the best time interval in detecting highlights.

7.1 Combined Physiological Responses

WE are in a process of conducting a user study with several sensory devices, using IAPS [20] data set as our stimulus. The goal is to categorize users based on their affective responses to stimulus.

There are available data set that can be used for this purpose [19]. However, to the best of author knowledge, in these studies users are 'wired' to connect different sensory devices. Such a setup can reduce users' emotional response.
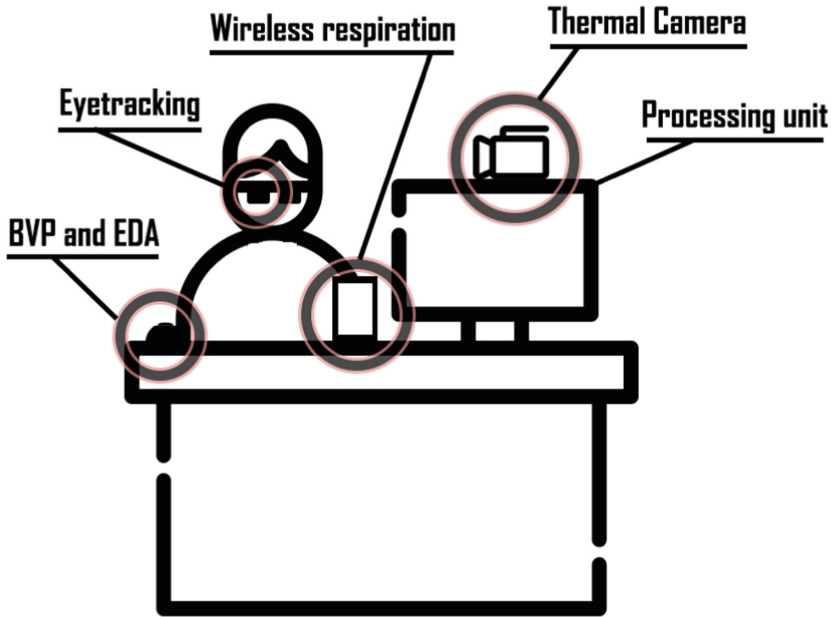
Fig. 7. The processing unit synchronizes the data and stores combined physiological signals. The system can provide real-time data analysis.

Also, the process of collecting these data sets does not consider the effect of precedence that might cause by strong contexts in the stimulus. For example, a physiological response caused by an unpleasant image followed by a pleasant image might not fully descend. This will cause to record the effect of viewing an unpleasant image when viewing a pleasant image.

We limit the use of wearable devices to a wristband and an eye-tracking device. The rest of the sensory devices are stationary.

The devices used in include:

– Empatica E4 [13] wristband: records a user's blood volume pressure and electrodermal activities.
– Jenoptik stationary thermal camera: records the users' facial skin temperature.
– Pupil [17] eye tracking headset device: records users' eye movement.
– Walabot [28], an RF sensory device: records user respiration.

Fig. 7 shows the setup and different devices that being used in the study.

Our stimulus is a collection of 120 images (60 neutral, 30 unpleasant, 30 pleasant) selected from IAPS. To reduce the effect of precedence caused by stimulus, we are using the proposed method of detecting highlights for selecting a sequence of images.

With this approach, the likelihood of affective response and the duration of physiological response for each image is known before hand. This gives us an

approximation of a time interval that is needed to reduce the effect caused by unpleasant or pleasant images. During this time a sequence of neutral images will be shown. With this approach each image with pleasant or unpleasant context will have its own physiological responses.

Finally to reduce the cognitive load from affective response during our study, we divide the study into two sections. The purpose of the first section is to focus on physiological response cause by the stimulus. During this section, users' physiological responses are recorded without asking any self-report questionnaires. During the second section, we ask the users to only rate each image using an affective slider [3] rather than using the self-assessment Manakin (SAM) [5]. Using affective slider is easier for users to understand, compared to SAM.

We hope that this approach for collecting data can gives us a representative and more reliable data that can be used to explore causation and correlation between stimulus and users' emotion. The data set can be used to create a model capable of clustering users based on their response to the images and predict users' behavior.

8   Conclusions

Affective computing is growing and the idea that one day an application can make decisions for us without our cognitive front is not out of reach. Today, biometric sensors are becoming smaller with better accuracy that can easily be used in daily tasks.

These advancements require the need to explore new methods and approaches in affective computing that can help us to explore new ways of interaction between human and computers. The results from our study show it is possible to detect the highlight of the stimulus using only the EDA signal of the group of people. This opens a new door for creating a system capable of recommending the stimulus based on the current physiological changes of the users.

Although the result of the study shows that by only using EDA signal it is possible to detect the highlights of the stimulus, a symmetric histogram may occur by increasing the number of participants. This might lead to the case that the algorithm include the whole duration of the stimulus as highlight, which is meaningless. Studies with larger sample sizes are needed to explore the result in more detail.

We mainly focused on video stimulus that can cause fear. However, it would be useful to run the same study with other basic emotions to find the physiological differences between emotions.

Overall the result from this study, demonstrate the potential of EDA signals, independently, in affective computing. We were able to gain a better understanding of physiological signals and their correlation and causation to stimulus. Although it is always possible to provide self-report in study, but the results might not fully represent our affective response.

In studies with several sensory devices, having method such as this can be beneficial to validate the effect of stimulus before hand. This can be helpful for

models to know the time intervals of the stimulus with higher likelihood of users' affective response. This information can be used to categeorize users based on their response or predict future responses based on the current ones. Further study needs to be done to explore if adding more sensory devices can narrows down the time interval for detecting highlights.

References

1. The evil within (2014). https://theevilwithin.bethesda.net
2. Benedek, M., Kaernbach, C.: A continuous measure of phasic electrodermal activity. J. Neurosci. Methods **190**(1), 80–91 (2010)
3. Betella, A., Verschure, P.F.M.J.: The affective slider: a digital self-assessment scale for the measurement of human emotions. PLOS One **11**(2), 1–11 (2016)
4. Boucsein, W.: Electrodermal Activity. The Springer Series in Behavioral Psychophysiology and Medicine. Springer, New York (2012). https://doi.org/10.1007/978-1-4614-1126-0
5. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. J. Behav. Ther. Exp. Psychiatry **25**(1), 49–59 (1994)
6. Chênes, C., Chanel, G., Soleymani, M., Pun, T.: Highlight detection in movie scenes through inter-users, physiological linkage. In: Ramzan, N., van Zwol, R., Lee, J.S., Clüver, K., Hua, X.S. (eds.) Social Media Retrieval, pp. 217–237. Springer, London (2013). https://doi.org/10.1007/978-1-4471-4555-4_10
7. Cowie, R., Cornelius, R.R.: Describing the emotional states that are expressed in speech. Speech Commun. **40**(1), 5–32 (2003)
8. Dawson, M., Schell, A., Filion, D.: The electrodermal system. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (eds.) Handbook of Psychophysiology, 2nd edn, pp. 200–223. Cambridge University Press, New York (2000)
9. Dhall, A., Asthana, A., Goecke, R., Gedeon, T.: Emotion recognition using PHOG and LPQ features. In: Face and Gesture 2011, pp. 878–883. IEEE, Piscataway (2011)
10. Ekman, P.: Facial expression and emotion. Am. Psychol. **48**(4), 384–392 (1993)
11. Feldman, L.A.: Variations in the circumplex structure of mood. Pers. Soc. Psychol. Bull. **21**(8), 806–817 (1995)
12. Gaebel, W., Wölwer, W.: Facial expression and emotional face recognition in schizophrenia and depression. Eur. Arch. Psychiatry Clin. Neurosci. **242**(1), 46–52 (1992)
13. Garbarino, M., Lai, M., Bender, D., Picard, R.W., Tognetti, S.: Empatica E3: a wearable wireless multi-sensor device for real-time computerized biofeedback and data acquisition. In: 2014 4th International Conference on Wireless Mobile Communication and Healthcare - Transforming Healthcare Through Innovations in Mobile and Wireless Technologies (MOBIHEALTH), pp. 39–42. IEEE, Piscataway (2014)
14. Haag, A., Goronzy, S., Schaich, P., Williams, J.: Emotion recognition using biosensors: first steps towards an automatic system. In: André, E., Dybkjær, L., Minker, W., Heisterkamp, P. (eds.) ADS 2004. LNCS (LNAI), vol. 3068, pp. 36–48. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-540-24842-2_4
15. Jaccard, P.: The distribution of the flora in the alpine zone.1. New Phytol. **11**(2), 37–50 (1912)
16. Joho, H., Staiano, J., Sebe, N., Jose, J.M.: Looking at the viewer: analysing facial activity to detect personal highlights of multimedia contents. Multimed. Tools Appl. **51**(2), 505–523 (2011)

17. Kassner, M., Patera, W., Bulling, A.: Pupil: an open source platform for pervasive eye tracking and mobile gaze-based interaction. In: ACM International Joint Conference on Pervasive and Ubiquitous Computing: Adjunct Publication (2014)
18. Kim, K.H., Bang, S.W., Kim, S.R.: Emotion recognition system using short-term monitoring of physiological signals. Med. Biol. Eng. Comput. **42**(3), 419–427 (2004)
19. Koelstra, S., et al.: DEAP: a database for emotion analysis; using physiological signals. IEEE Trans. Affect. Comput. **3**(1), 18–31 (2012)
20. Lang, P.J., Bradley, M.M., Cuthbert, B.N.: International affective picture system (IAPS): affective ratings of pictures and instruction manual. University of Florida, Gainesville, FL, Technical report (2008)
21. Levenson, R.W.: Autonomic nervous system differences among emotions. Psychol. Sci. **3**(1), 23–27 (1992)
22. Levenson, R.W., Ekman, P., Friesen, W.V.: Voluntary facial action generates emotion-specific autonomic nervous system activity. Psychophysiology **27**(4), 363–384 (1990)
23. Lynch, T., Martins, N.: Nothing to fear? An analysis of college students' fear experiences with video games. J. Broadcast. Electron. Media **59**(2), 298–317 (2015)
24. Marsh, A.A., Blair, R.J.R.: Deficits in facial affect recognition among antisocial populations: a meta-analysis. Neurosci. Biobehav. Rev. **32**(3), 454–465 (2008)
25. Mauss, I.B., Robinson, M.D.: Measures of emotion: a review. Cogn. Emot. **23**(2), 209–237 (2009)
26. Murugappan, M., Ramachandran, N., Sazali, Y.: Classification of human emotion from EEG using discrete wavelet transform. J. Biomed. Sci. Eng. **03**(04), 390–396 (2010)
27. Philippot, P.: Inducing and assessing differentiated emotion-feeling states in the laboratory. Cogn. Emot. **7**(2), 171–193 (1993). pMID: 27102736
28. Pierce, A.: Walabot diy can see into walls. Tech Direct. **76**(5), 8–9 (2017). Copyright - Copyright Prakken Publications, Inc., Jan 2017; Document feature - Photographs; Last updated - 2017–11–23
29. Ragot, M., Martin, N., Em, S., Pallamin, N., Diverrez, J.-M.: Emotion recognition using physiological signals: laboratory vs. wearable sensors. In: Ahram, T., Falcão, C. (eds.) AHFE 2017. AISC, vol. 608, pp. 15–22. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-60639-2_2
30. Tasooji, R.: Determining correlation between video stimuli and electrodermal activity. Master's thesis, Virginia Polytechnic Institute and State University, Blacksburg, VA (2018)
31. Trubanova, A., et al.: The role of perspective-taking on ability to recognize fear. Curr. Res. Psychol. **6**(2), 22–30 (2016)
32. Zachary, W.W.: An information flow model for conflict and fission in small groups. J. Anthropol. Res. **33**(4), 452–473 (1977)