A Protocol for Acquisition of Emotion and Affect in Music – Dataset Collection During COVID-19 (EAM-DCDC)

Aleksandra Wołowiec and Paweł Jemioło
AGH University of Science and Technology,
al. A. Mickiewicza 30, 30-059 Krakow, Poland
{wolowiec, pawljmlo}@agh.edu.pl

Abstract. Emotion and affect recognition constitute one of the most frequently explored areas in Human-Computer Interaction. Since the credibility of data used for generating Artificial Intelligence models in the field is questioned, acquiring new high-quality datasets is essential. In this study, we aim to present a protocol for the experiment, which focuses on gathering cardiovascular data during music stimulation. Following this protocol, our objective is to provide a credible, ecologically valid and publicly available dataset, which will be further used for emotion and affect recognition with Artificial Intelligence models. Preliminary insights from our pilot research serve as the basis for protocol.

Keywords: Music · Cardiovascular · Heart Rate · Emotion and Affect Recognition · Database · Dataset · COVID-19

1 Introduction

Emotion and affect recognition from physiological signals are one of the main objectives of Human-Computer Interaction. For this purpose, the internal states of participants are evoked employing different stimuli, e.g. an audio file with sounds of nature, music, pictures, or a short movie. At the same time, people that take part in experiments are connected to devices, which measure the activity of their brain, heart, or vascular system, to name only a few. Further, data are pre-processed, and they serve as an input for Artificial Intelligence models for emotion and affect recognition.

As our recent systematic review [5] revealed, there are 18 publicly available datasets for emotion and affect recognition, which contain cardiovascular signals. However, only 2 of them may be considered high-quality ones. In VRAD [6], authors used virtual reality for emotion elicitation, while AMIGOS [3] employed audiovisual clips. Therefore, there is still a need for new credible datasets, including cardiovascular signals, to be acquired with other stimulation methods, e.g., music or pictures.

The replication crisis in psychology and reproducibility problems in computer and information sciences show that the data used to generate Artificial Intelligence models need to be collected via thoroughly planned experiments [5]. Otherwise, we may face the *garbage in, garbage out* problem, which refers to a situation when doubtful input results in producing questionable output. That is why many initiatives in this field emerged, e.g. preparation of a protocol for the study and its registration, *registered reports*, preregistration of hypotheses, or replication of previously published studies.

Additionally, Human-Computer Interaction faces another issue, which arose from COVID-19. Namely, conducting research with human subjects in laboratories during a pandemic is arduous or even impossible due to restrictions. Fortunately, alternatives exist, and it occurs that most of the experiments may be successfully carried out via the internet thanks to initiatives like Amazon Mechanical Turk or simply through passing equipment to participants, particularly when the research does not involve highly-specialised apparatus. Not only the safety of the subjects is ensured, and they may participate in the experiment at a time that suits them, but also the ecological validity of the study increases.

The purpose of this paper is to present a protocol for the experiment, which focuses on the acquisition of physiological data during music stimulation. Following this protocol, we aim at providing a credible, ecologically valid and publicly available dataset, which will be further used for emotion and affect recognition with Artificial Intelligence models.

2 Methods

In this study, participants will experience music stimuli (songs excerpts). However, the choice of emotion-evoking material needs to be justified in a separate pilot study with other yet similar participants. One needs to make sure that selected stimuli indeed evoke intended emotions.

In our first attempt at stimuli validation, we selected 60 songs from the dataset with music excerpts grouped by the Artificial Intelligence model[1] into 4 emotional quadrants (High Arousal High Valence – HAHV, High Arousal Low Valence – HALV, Low Arousal Low Valence – LALV, Low Arousal High Valence – LAHV) of Russell space. Russell's space consists of two emotional dimensions, namely arousal and valence [7]. The first one refers to the intensity of emotion, while the second one is connected with the aspect of pleasure. Unfortunately, our initial choice occurred to be improper (see section 3); thus, we incorporated another source of stimuli for validation. For this phase we selected 64 songs from DEAM [1], 16 excerpts per Russell's quadrant. From 64 songs, we will choose 32 stimuli with the most dominant valence and arousal values, 8 per category.

In our experiments, the affective rating system called Self-Assessment Manikin (SAM) [2] will be used for stimuli assessment for all participants. SAM allows measuring three emotional dimensions of valence, arousal, and dominance with different scales. Each of the scales consists of 5 drawings showing the experienced emotional states. A verbal description will precede the scales for precise interpretation of the images. The user's task will be to select an adequate value from 1-9 using the slider.

We will include participants aged 20-30 with higher education (or ongoing) who have not graduated from music schools or work in the music industry but listen to music every day. Exclude criteria involve: cardiovascular or mental disorders and inability to recognise emotions (alexithymia). We plan to examine 45 people in the main study, which is more than average in existing datasets. See our recent systematic review for more details regarding the number of subjects [5].

[1] github.com/cristobalvch/Spotify-Machine-Learning (Accessed on 28.02.2022).

Additionally, we will control potential confounders. Firstly, we will ask participants about their mood, which may influence experienced emotions. Next, as the experiment will take place at the subjects' homes, they will be prompted to certify that they were listening to the songs themselves, no one disturbed them, and they were not taking any additional activities. Moreover, we will ask participants to rate their concentration level during the whole procedure.

The pilot and main studies will be conducted using Google Forms, and PsychoPy deployed on Pavlovia servers. Participants will receive instructions in Google Docs with links to both parts of the experiment and information on how to handle the delivered device (in case of the main study).

The first part is a personal questionnaire (Google Forms) with basic information: age, gender, ethnicity, which genre of music the respondents prefer, whether they usually listen to music in line with their mood or the opposite and how they feel at a given moment (SAM).

The second part consists of the following sequences: 10-second break, a 45-second song, SAM scale rating, and yes/no question for determining whether participants liked the song. The order in which the songs are played for each subject is determined randomly.

In the main experiment, the participant's heart rate will be measured with a wearable smart-band Xiaomi Mi Band 5 during the second part (PsychoPy). Before starting the experiment, subjects will be asked to prepare a phone with a voice recorder (recording experiment for synchronisation purposes) and a computer with an active loudspeaker. After starting the experiment, participants will be asked to wait 2 minutes for the band to measure their baseline. After the subject completes the experiment, the data will be extracted from the band using the Gadgetbridge application[2] for Android smartphones, and then the band will be passed on to the next participant.

After acquiring the data, we will validate the data (to confirm they can be used for emotion and affect recognition) using basic Artificial Intelligence algorithms. We will use Support Vector Machines, Naive Bayes, Random Forest and Deep Learning techniques (e.g. following Harper and Southern [4] method) as they are most frequently incorporated [5].

After validation, the data will be shared anonymously in the form of CSV files via Open Science Framework (OSF). Every participant will be asked whether they agree to share the data.

3 Preliminary results

In Figure 1, we present results regarding our first attempt to pilot study (valence and arousal ratings). Songs selected precisely for particular quadrants do not overlap with participants' answers. Notably, we have not obtained space coverage for LVHA songs that comply with the dataset ground-truth. We thus need another pilot study with songs chosen more adequately (the DEAM dataset [1]).

[2] codeberg.org/Freeyourgadget/Gadgetbridge (Accessed on 28.02.2022).

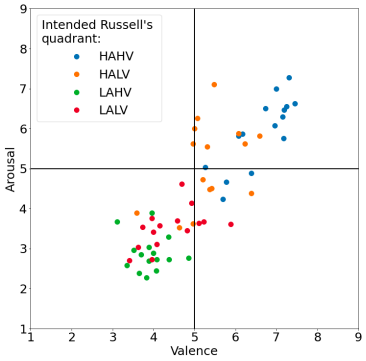

Fig. 1. Participants responses regarding heard songs (each song is a dot).

4 Conclusions and Future Works

In this paper, we comprehensively presented a protocol for the experiment, which focuses on the acquisition of cardiovascular data during music stimulation for emotion and affect recognition with Artificial Intelligence models.

Now, we want to follow this protocol, gather the data, validate it with selected models and share it. We believe that only thorough methodology allows for acquiring high-quality data for development of Artificial Intelligence solutions.

Funding Sources

Research project supported by program *Excellence initiative -- research university* for the University of Science and Technology.

References

1. Aljanaki, A., Yang, Y.H., Soleymani, M.: Developing a benchmark for emotional analysis of music. PloS one **12**(3), e0173392 (2017)
2. Bradley, M.M., Lang, P.J.: Measuring emotion: the self-assessment manikin and the semantic differential. Journal of behavior therapy and experimental psychiatry **25**(1), 49–59 (1994)
3. Correa, J.A.M., Abadi, M.K., Sebe, N., Patras, I.: Amigos: A dataset for affect, personality and mood research on individuals and groups. IEEE Transactions on Affective Computing (2018)
4. Harper, R., Southern, J.: A bayesian deep learning framework for end-to-end prediction of emotion from heartbeat. IEEE Transactions on Affective Computing (2020)
5. Jemioło, P., Storman, D., Mamica, M., Szymkowski, M., Żabicka, W., Wojtaszek-Główka, M., Ligęza, A.: Datasets for automated affect and emotion recognition from cardiovascular signals using artificial intelligence — SR. Sensors (in press) (2022)
6. Marín-Morales, J., Higuera-Trujillo, J.L., Greco, A., Guixeres, J., Llinares, C., Scilingo, E.P., Alcañiz, M., Valenza, G.: Affective computing in virtual reality: emotion recognition from brain and heartbeat dynamics using wearable sensors. Scientific reports **8**(1), 1–15 (2018)
7. Russell, J.A.: A circumplex model of affect. Journal of personality and social psychology **39**(6), 1161 (1980)