Study Design Document

Affective computing for subjective video quality assessment

Project: Affective computing for subjective video quality assessment

Author: Nida Cilasun, Mengzhu Deng, Franz Hahn, Yazan Kittaneh
Fabian Klopfer, Jan Laun-Cieslik, Anlin Liang, Zhou Long,
Daniil Matzkuhn, Andrei Sacuiu

Last Change: April 17, 2018 , Version 1.0

Contents

1 Introduction

In our daily life, many of us consume video content frequently. Be it on mobile devices, video platforms or even at public places. While consuming video content we experience a bunch of emotions that surely depends on our current mood and the video content. But is there an affective dependency between varieties in video quality and the experienced emotions? Since a lower quality typically comes up with a smaller memory size developer and users are interested to reduce buffering time and memory space. Reducing the video size is a wide research area especially in the context of streaming, videos on mobiles or wearable devices. More and more video files are not stored on devices anymore but streamed online and personal data get stored in clouds. Therefore we are interested in how the video quality does affect the experienced emotions during video consumption. Knowing the affective impact of a lower video quality could help to handle the trade-off between video quality and size in multimedia services.

This document shall provide foundations, the experimental set-up and preliminaries of the study, preparing for execution, data evaluation and ultimately a seminar paper with a more dedicated focus.

Possible Hypothesis

> Emotions positively correlate with the video quality. Emotions are more sensual in case of high video quality while emotions are less intensive when the video quality is low.
>
> The captured mood does not depend on the video quality. Mood instances stick the same with underlying high or low video quality.

In this assessment, we quantify the observer's moody reaction to video content and relate them to perceptual quality. By using facial recognition, skin conductance sensors for tracking electrodermal activity, electromyography and electrocardiography this study tracks facial expressions, face muscles, sweating, and heart rate. Beside of sensors quality assessment is used by questionnaires for evaluating the video quality and quality of service. The underlying videos are composed of divergent stimuli in a sense of high and low arousal as well as high and low valence.[29] The quality assessment is realized as a collaboration project between the University of Konstanz and the Technical University Berlin.

2 Subjective Quality Assessment

The subjective assessment of multimedia video quality may be used for video-only or audiovisual tests. Subjective evaluation of video quality is the fundamental ground truth to compare the performance of objective methods of quality assessment.[7] It is concerned with how a video is perceived by a observer and designates their opinion on a particular video sequence and therefore related to the field of quality of experience

The measurement of subjective video quality is necessary since objective quality assessment algorithms have been shown to correlate badly with ratings.[7] Subjective ratings may also be used as ground truth to develop new algorithms.[13] In order to obtain reliable results out of raw subjective scores, screening of the observers might be employed to discard observers that are considered as outliers.

In subjective evaluation, a group of subjects is asked to watch a set of video clips and to rate their quality. The scores assigned by the observers are averaged in order to obtain the Mean Opinion Scores (MOS). In order to produce meaningful MOS values, each stimuli needs to be carefully selected and randomized.

2.1 Quality

Video quality metrics can be classified according to the availability of the original video signal.[10] This original video signal is considered to be distortion-free or perfect quality.[10] For measuring video quality relatively, we take the original video signal as a reference to compare a distorted video signal.[11] One of the main goals of subjective video quality metrics is to automatically estimate the average observer's opinion on the quality of a video processed by a system.[11] Procedures for subjective video quality measurements are described in ITU-T recommendation P.910 [14].[11] Video sequences are shown to a group of observers.[11] The observers' opinion is recorded and averaged into the Mean Opinion Score to evaluate the quality of each video sequence.

2.1.1 Scale

The study shall use a continuous category-based rating scale with hidden references to derive both the mean opinion score and differential viewer scores. The calculus of those is equivalent as described for ACR-HR in the P.910 paper [14].[30] The continuous scale is chosen based on the **Pick-A-Mood system** and with respect to the continuous character of valence and arousal.

"Valence (pleasure - displeasure) and arousal (high energy – low energy).[21] Together these dimensions represent four basic mood categories, see Figure 1" [12]

Secondly we use a **five-level scale** to evaluate the recognized video quality. Thereby the scale offers five radio buttons, horizontal orders from left to right with the following combination of text and numbers: 1:Bad, 2:Poor, 3:Fair, 4:Good and 5:Excellent.

2.2 Emotional Model:[8] The Pick-A-Mood System

The Pick-A-Mood system is a cartoon-based pictorial instrument for reporting and expressing moods [11].[8] Pick-A-Mood consists of three characters, each expressing eight different mood states (and a neutral expression), representing four basic mood categories:

- Excited-Lively and Cheerful-Happy (for energized-pleasant)
- Tense-Nervous and Irritated-Annoyed (for energized-unpleasant)
- Calm-Serene and Relaxed-Carefree (for calm-pleasant)
- Bored-Weary and Gloomy-Sad (for calm-unpleasant)

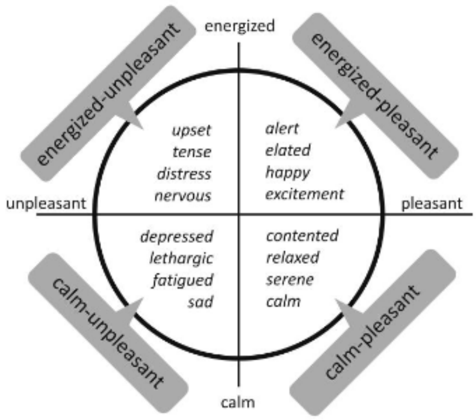


Figure 1. Four basic mood categories; based on the PANAS model by Watson and Tellegen (1985), with examples of moods (in the circle) from Russell (1980) and Barrett & Russell (1999).

Graph: Pick a mood categories [12]

3 Stimuli

3.1 Origin & License

The LIRIS-ACCEDE homepage states:
„[6]The discrete and continuous sections of LIRIS-ACCEDE have been created by a french team of researchers:"

- Yoann Baveye, Technicolor & Ecole Centrale de Lyon, LIRIS, France
- Emmanuel Dellandréa, Ecole Centrale de Lyon, LIRIS, France
- Christel Chamaret, Technicolor, France
- Liming Chen, Ecole Centrale de Lyon, LIRIS, France
- Jean-Noël Bettinelli, Ecole Centrale de Lyon, LIRIS, France
- Ting Li, Technicolor, LIRIS

„[6]All excerpts are shared under Creative Commons licenses and can thus be freely distributed without copyright issues.[6]The dataset (the video clips, annotations, features and protocols) are publicly available". [6]

3.2 Selection

It is desirable to have both - discrete and continuous data - available. We select those excepts of the discrete database, that are of movies which are also in the continuous data set. In numbers those are approximately 1600 of 9800.

In order to not influence the participants it is desirable to show uniformly many stimuli of each quadrant on a (valence, arousal)-coordinate system. To keep the time frame for a participant decent, we've decided to take 4 videos per quadrant.

Additionally all selected videos should have the same resolution and shall be colorful videos (meaning no black-white and no animated videos)

3.3 Structure

Original data-set

Discrete The discrete database contains in total 9,800 excerpts from 8 to 12 seconds long. It is labeled with raw data and rankings for arousal and valence dimensions. The data on all stimuli are in one huge table; the excerpts are all in one folder named by an ID. [6]

Continuous The continuous database contains 30 full movies with raw data (valence, arousal, time).[6]Additionally EDA measurements are also available. [6]

Intermediate data-set

Discrete The approximately 1600 stimuli are sorted into folders by movies, the data of the excerpts for one movie are in a file in the same folder.

Continuous No changes

After this the black-white and animated videos are excluded. Then we sample 4 videos from each quadrant of the valance-arousal scale.

3.4 Pre-Processing

FFmpeg provides ready-to-use functions to alter the frame rate (quality factor (QF) of an image. [5]

Stimuli Derivatives For each of the 4 4 excerpts, 3 versions are derived: quality factor 0, 22 and 35.

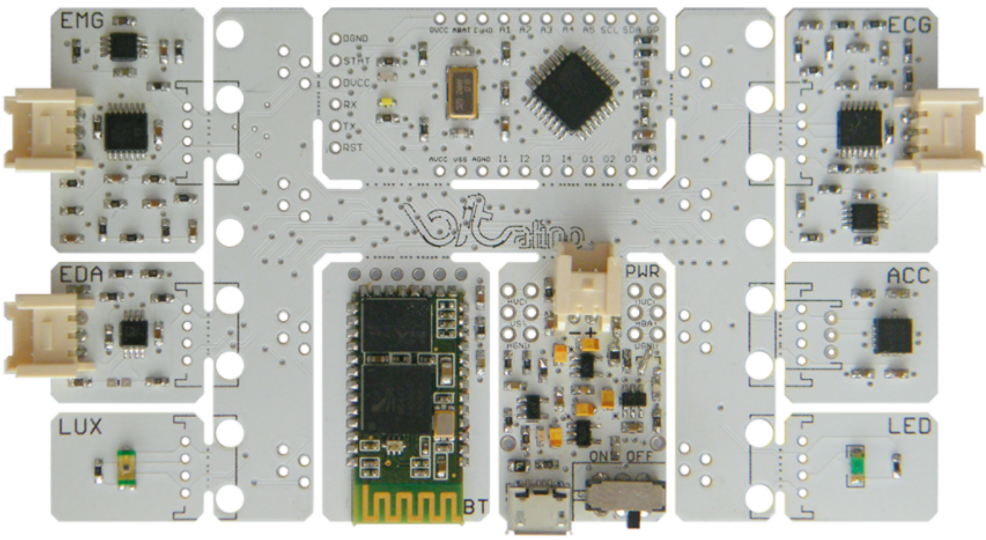
Pseudo-Randomization of shown Stimuli Sequence A python script was written that generates the videos order in which they will be shown. As a structure it generates blocks of 16 videos, each stimuli being shown only once, at a random quality. The last video of the current block is not allowed to be the last of the next block.

4 Sensors

The following is a fair description of the sensors to be used. We will be using 4 sensors in total, 2 EMG sensors, 1 EDA sensor and 1 ECG sensor, together with the BITalino for each test. Each sensor is the combination of a pair of leads, wires and the connecting chips which are used to connect the sensors and BITalino. Biosignal will be gathered using BITalino, and we will extract data from it afterwards. Data will be processed and plotted using Matlab.

4.1 BITalino Kit

BITalino is an easy-to-use, versatile and scalable hardware platform for biosignals acquisition and wireless transmission in real-time. It comes with following main modules:
Micro-controller unit(MCU): converts the analogic signals from the sensors to a digital format, and samples all of the channels. The MCU provides access to the BITalino analog and digital channels, as well as to the peripherals.
Power management block: provides energy to all the other BITalino blocks. This module also has a built-in charger that controls the battery charging.
EDA, EMG, ECG blocks: performs respective measurements of biosignals.



Graph: Overview of the Board

4.2 EDA: Electrodermal activity

EDA stands for electrodermal activity, and it can be defined as a transient change in certain electrical properties of the skin, associated with the sweat gland activity and elicited by any stimulus that evokes an arousal or orienting response. The EDA sensor

is capable of measuring the skin activity with high sensitivity measurement power in a miniaturized form factor. With low noise signal conditioning and amplification circuitry, we are able to provide accurate sensing capability and detect even the most feeble electrodermal skin response events. [10]
Equipment: One EDA sensor, Bitalino. For technical details on the sensor see [8]
Setup: Sensor will be placed at the palm of the participant.

4.3 EMG: Electromyography

EMG stands for electromyography activity. Muscular activation involves the action of muscles and nerves, which is triggered by very small electrical currents. Measuring the electrical activity in muscles and nerves can be useful for Human-Computer Interaction, control, biofeedback and many other applications. The sensor is capable of performing electromyography measurements using bipolar surface electrodes (plus a ground lead), and monitors the muscle activation.
Equipment: Two EMG sensor, Bitalino. For technical details on the sensor see [9]
Setup: We are going to apply the sensors to the cheeks and above the eyebrow.

4.4 ECG: Electrocardiography

ECG stands for electrocardiogram activity. Conduction of action potentials through the heart generates electrical currents that can be picked up by electrodes placed on the skin. A recording of the electrical changes that accompany the heartbeat is called an electrocardiogram. Variations in the size and duration of the waves of an ECG are useful in diagnosing abnormal cardiac rhythms and conduction patterns. The ECG works mostly by detecting and amplifying the tiny electrical changes on the skin that are caused during the heart muscle cycle during each heart beat.
Equipment: One ECG sensor, Bitalino. For technical details on the sensor see [7]
Setup: We are going to apply the sensors on parts of the neck, pulse and elbow.

4.5 FER: Camera and Facial Emotion Recognition

Equipment: GoPro Hero 3
Implementation: The code base used was originally coded by Mitiku Yohannes [4]. The code in the data_collect folder is used to bring the data sets into a certain folder structure. The preprocessor loads the images and provides a data generator (with optional data augmentation) that may be fed into the network as input dictionary. The mainly tested network is the multi-input neural network architecture using the image, labels and 68 point face landmarks generated by the dlib library [16]. It further implements CapsNet [18] as an experiment, either image or 68 point face landmarks input architectures, the VGGFace [17] architecture, the Inception-ResNet [19] architecture and some LSTM [13] architectures. The Inception-ResNet architecture and the LSTM architectures weren't trained, due to a lack of fitting data in the given time frame of 3 weeks. The util folder contains a logger class, a post-processor which is used for printing the result on the image and showing it and a wrapper class fot the neural net which

defines the exact classification objective. The used data were those of the 2013 Kaggle challange on facial emotion recognition [2] and the Cohn-Kanade+ database [15]

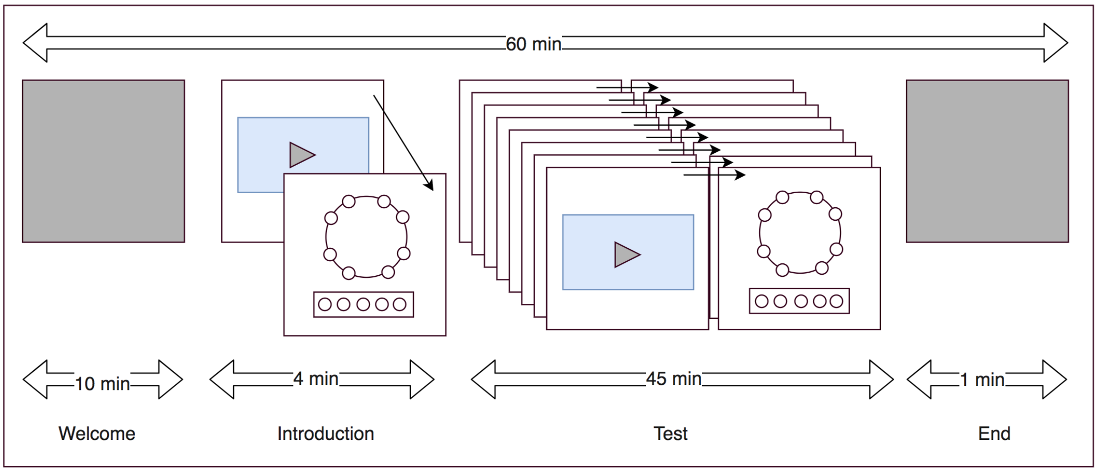
The base implementation is presented at GitHub [3].

5 Study Design

5.1 Participants

Try to reach a inhomogeneous group of participants as good as possible. Therefor we capture the sex and age of the participant. Ensure that glasses get replaced by contact lenses and participants have to be notified about make-up removal procedure. Twelve participants are minimum, a group of 15 participants would be good. The more the better.

5.2 Overview



Graph: Overview of the Study

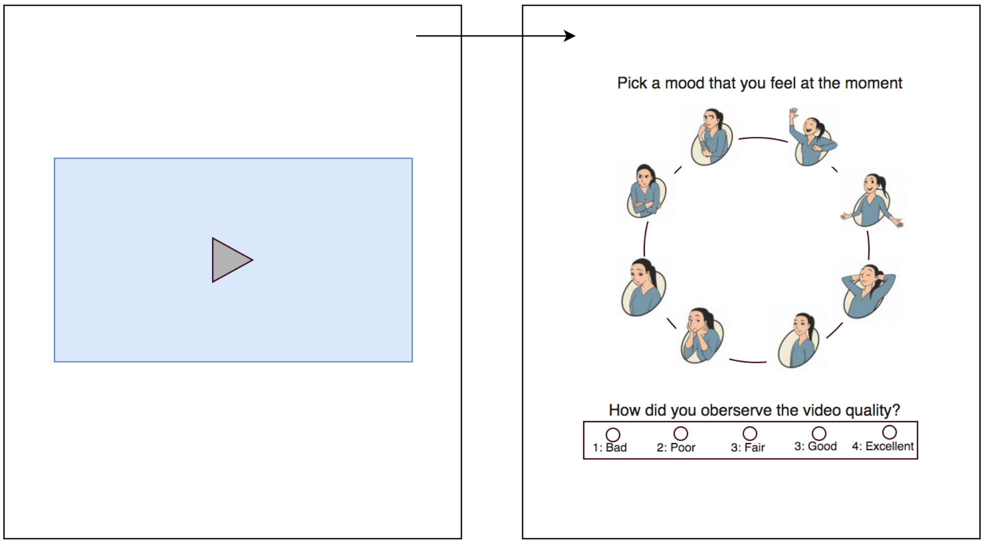
The laboratory experiment is partitioned into four phases: Welcome-, Introduction-, Test- and End phase. The overall duration will be about one hour.

5.3 Setup

The video stimuli are shown with a computer monitor. By using Psychtoolbox-3 for Matlab a iterative user interface gets offered. In total there will be 20 minutes of pure watching time by a amount of 48 random videos (4 4 3 stimuli). Each video will take about 10 to 15 seconds and five more seconds each for questionnaires. After half of the videos the participants get asked to take a 5 minutes mandatory break.

5.3.[8]1 Questionnaire

A questionnaire provides a reliable understanding of a respondent's mood state in a qualitative perspective. After a single video sequence, the participant will be asked to name their current mood state that they have at the moment by the pick-a-mood selection. Moreover the respondent will be asked to rate the perceived video quality with a five level scale.



Graph: Questionnaire for qualitative video assessment

Welcome: In the first phase, that will take about 10 minutes, it will be shown a neutral grey image. In this step the participant enters the room, get asked to take a seat and the sensors get prepared.

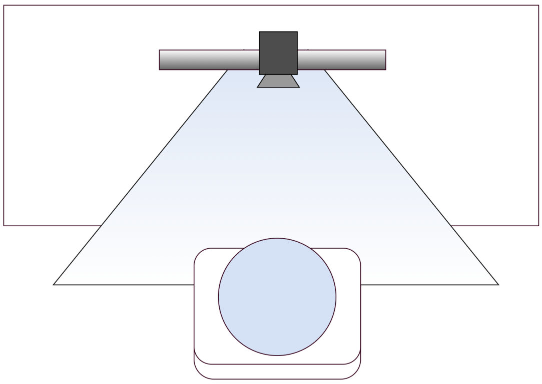
Introduction: The lab-leader guides the participant threw the first video and questionnaire as an example when the sensors are ready. Therefor the lab-leader explains according to the example how the pick-a-mood buttons and the five-level scale for rating of the overall quality works. Finally the participant gets asked to complete the questionnaire of the provided example by selecting a mood and rating the quality scale. When there are no more further questions the participant gets asked to press the "okay" button on the right bottom corner that leads to the next video.

Test: The video gets automatically played by reaching this phase. When the video sequence (10-15 seconds) is completely shown the participant comes (automatically) to the questionnaire view. In this view the participant is asked to give a spontaneous feedback on how the participant feels at the moment by selecting a mood in the pick-a-mood diagram. Moreover the participant is asked to evaluate the overall quality **during the video** with a five-level scale. This scale displays from left to right: 1:Bad, 2:Poor, 3:Fair, 4:Good and 5:Excellent. When both two feedback modules are answered the "okay" button on the right bottom corner leads to the next video which gets played immediately. After half of the videos the participant is asked to take a 5 minutes mandatory break. The circuit stops when all corresponding video sequences are shown.

End: When all video sequences are shown the screen turns automatically to a neutral grey screen.

5.3.2 Monitor and Camera setup

Place the monitor on a table desk with the same space in the left and right direction. For capturing the emotions of the participants it is necessary to record the face from a frontal perspective. Therefor position the camera directly above the monitor and adjust the camera direction to the center of the participant's face. Be aware that the participant is sitting in a straight and comfortable position. The video is recorded on a GoPro Hero camera with a resolution of 1920*1080 pixels and a frame rate of 50 frames per second. If the camera does not have a display, the viewfinder for the positioning can be accessed through the GoPro smartphone app. It can be downloaded from the Google Play Store or Apple App Store.



Graph: monitor and camera position

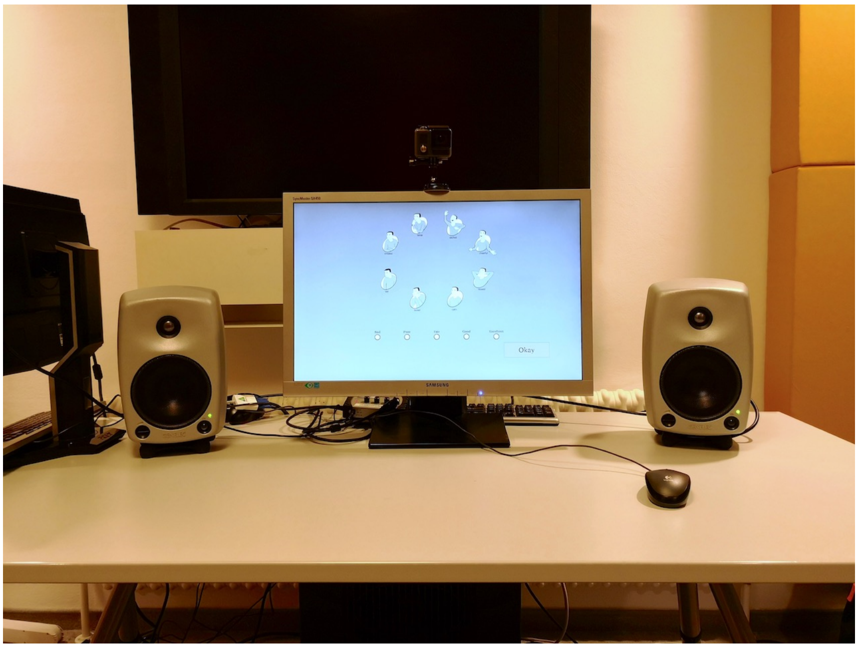


Photo: monitor, camera and speaker position

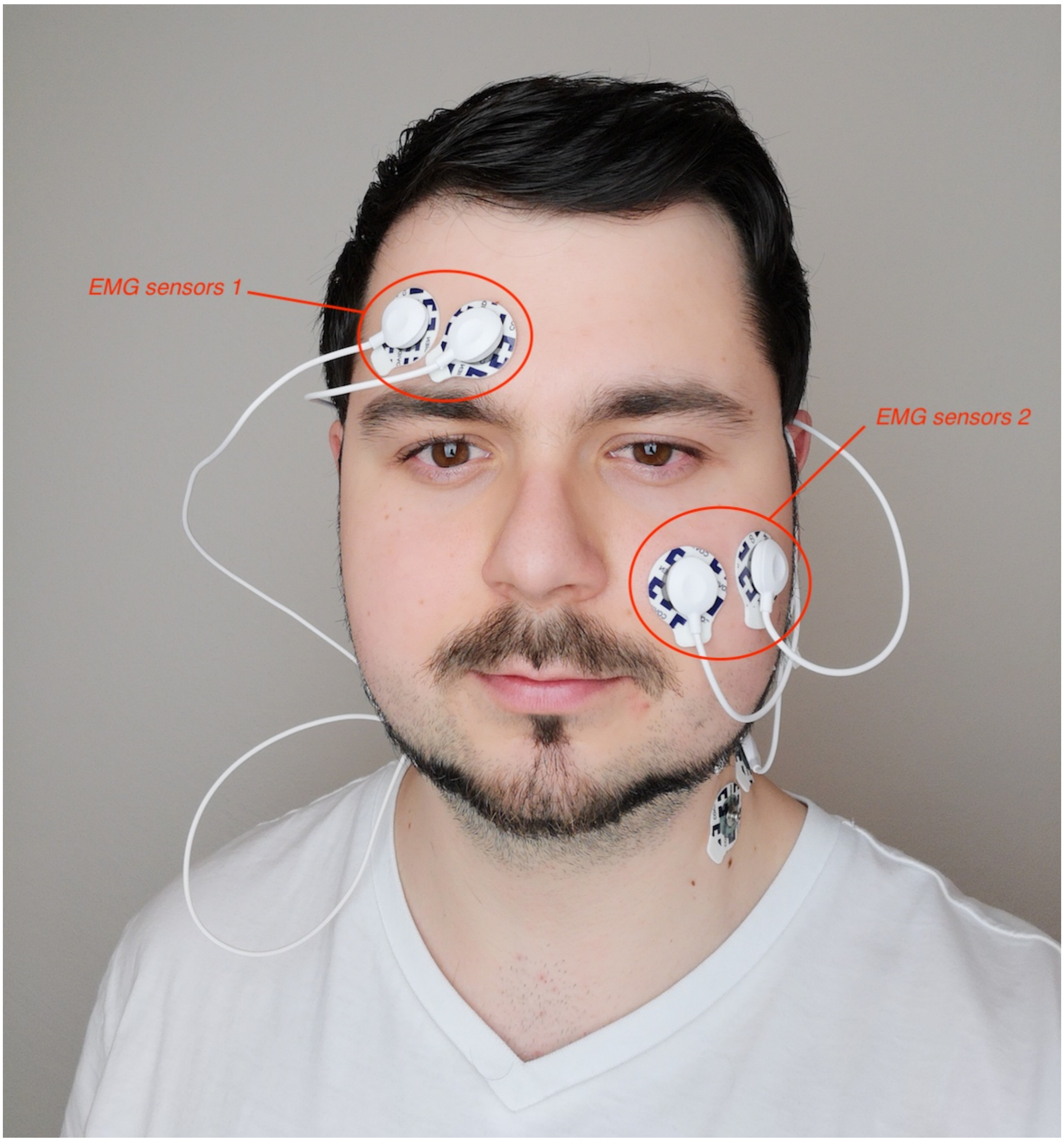
Photo: desk setup from the side with the participant in front of the screen

5.3.3 Bio-Sensor placement

Please ensure that the participants are not wearing glasses because of the facial recognition (contact-lenses are totally fine). For skin conductance prepare make-up/skin fat removal tool.

For the experiment a total of four sensors will be used.[28] Two EMG sensors (with a total of 6 electrodes) will be used in the face of the participant, one EDA sensor (with a total of two electrodes) will be placed on the left palm of the participant and one ECG sensor (with a total of three electrodes) will be placed on the left wrist and right neck part. In total, 11 electrodes will be used. Make sure to attach them to the electrode cables before sticking them to the participant, as otherwise, the procedure might end up painful.

EMG sensor 1: EMG sensor 1 consists of three electrode cables. The outer two electrodes are positioned above the left eyebrow as seen in the image below. The ground electrode (middle cable) can be conveniently positioned in the neck area, as it only needs skin contact. The electrode cable (3-pin) is then connected with the EMG sensor board. This is then connected to a 4pin - RJ10 cable with port A1 on the bitalino board. Please make sure that the plugs are solid and fully inserted.

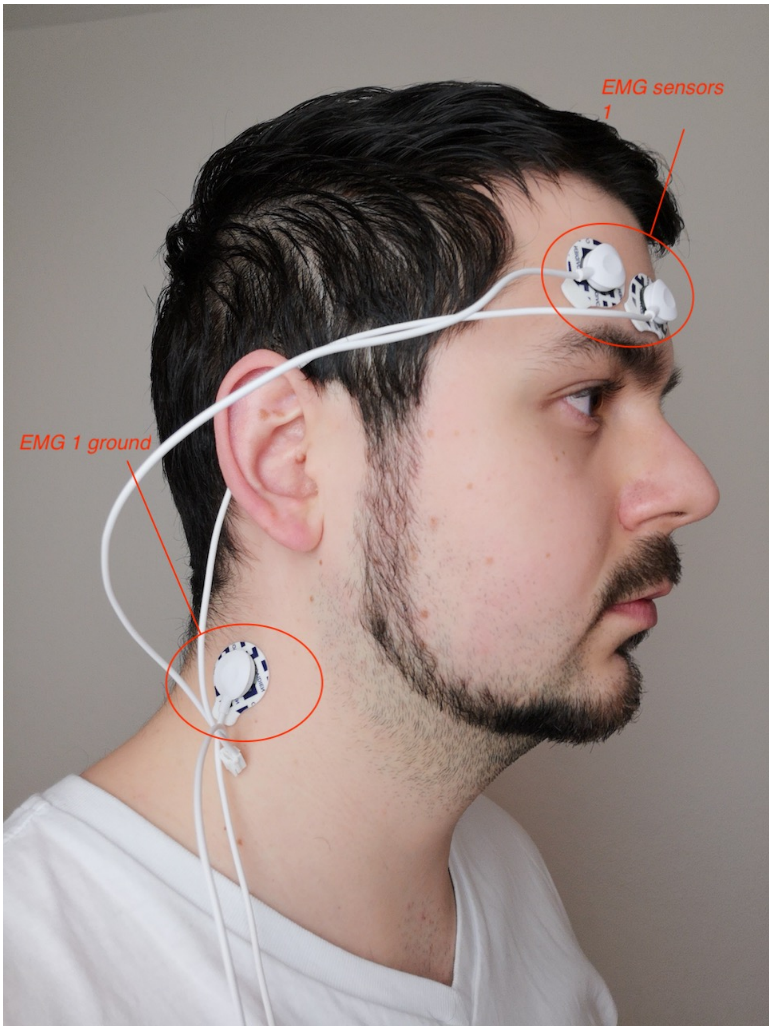


Graph: EMG sensor setup

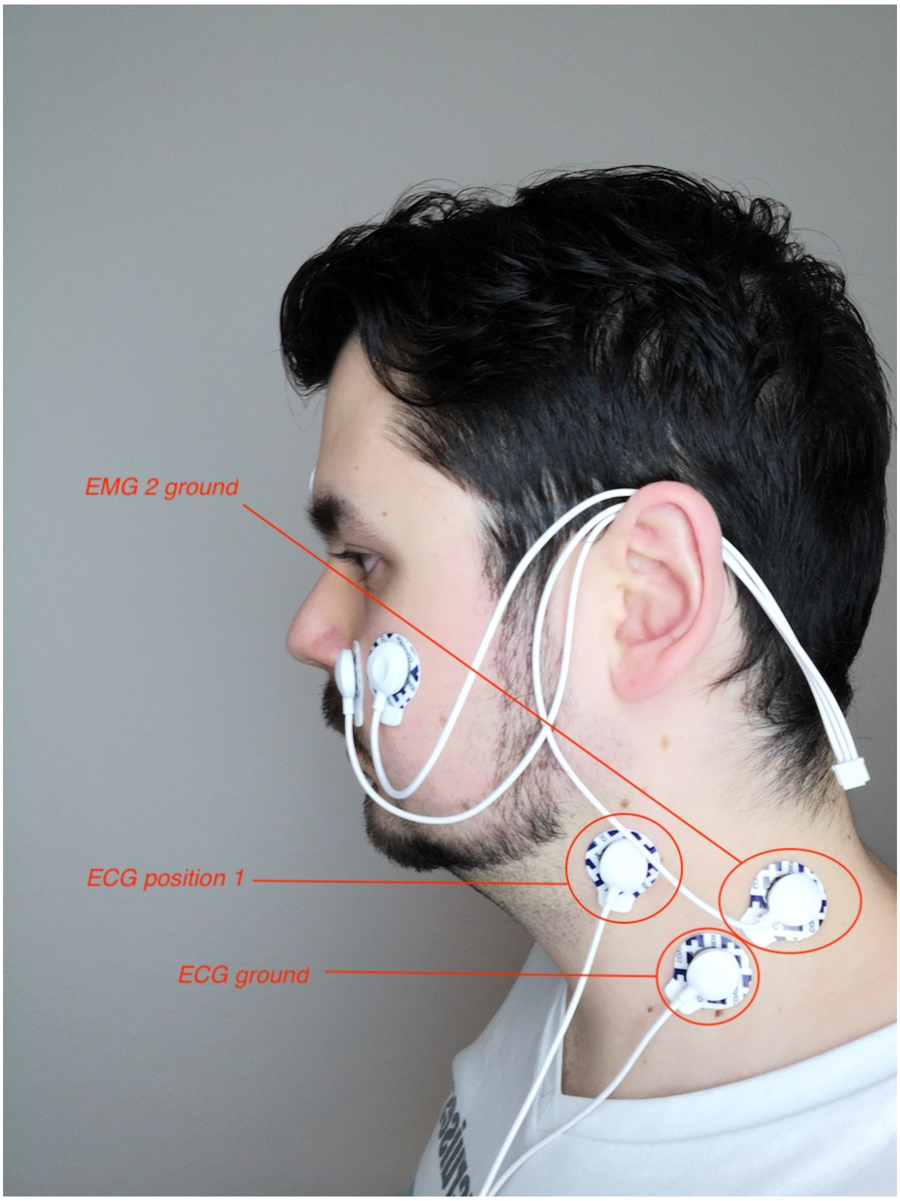
EMG sensor 2: EMG sensor 2 consists of three electrode cables. The outer two electrodes are positioned on the right cheek as seen in the image above. The ground electrode (middle cable) can be conveniently positioned in the neck area, as it only needs skin contact. The electrode cable (3-pin) is then connected with the EMG sensor board. This is then connected to a 4pin - RJ10 cable with port A4 on the bitalino board. Please make sure that the plugs are solid and fully inserted.

EDA sensor: EDA sensor consists of two electrode cables. The outer two electrodes are positioned on the palm of the left hand, as the right hand would be needed to operate mouse. The electrode cable (3-pin) is then connected with the EDA sensor board. This is then connected to a 4pin - RJ10 cable with port 2 on the bitalino board. Please make sure that the plugs are solid and fully inserted.

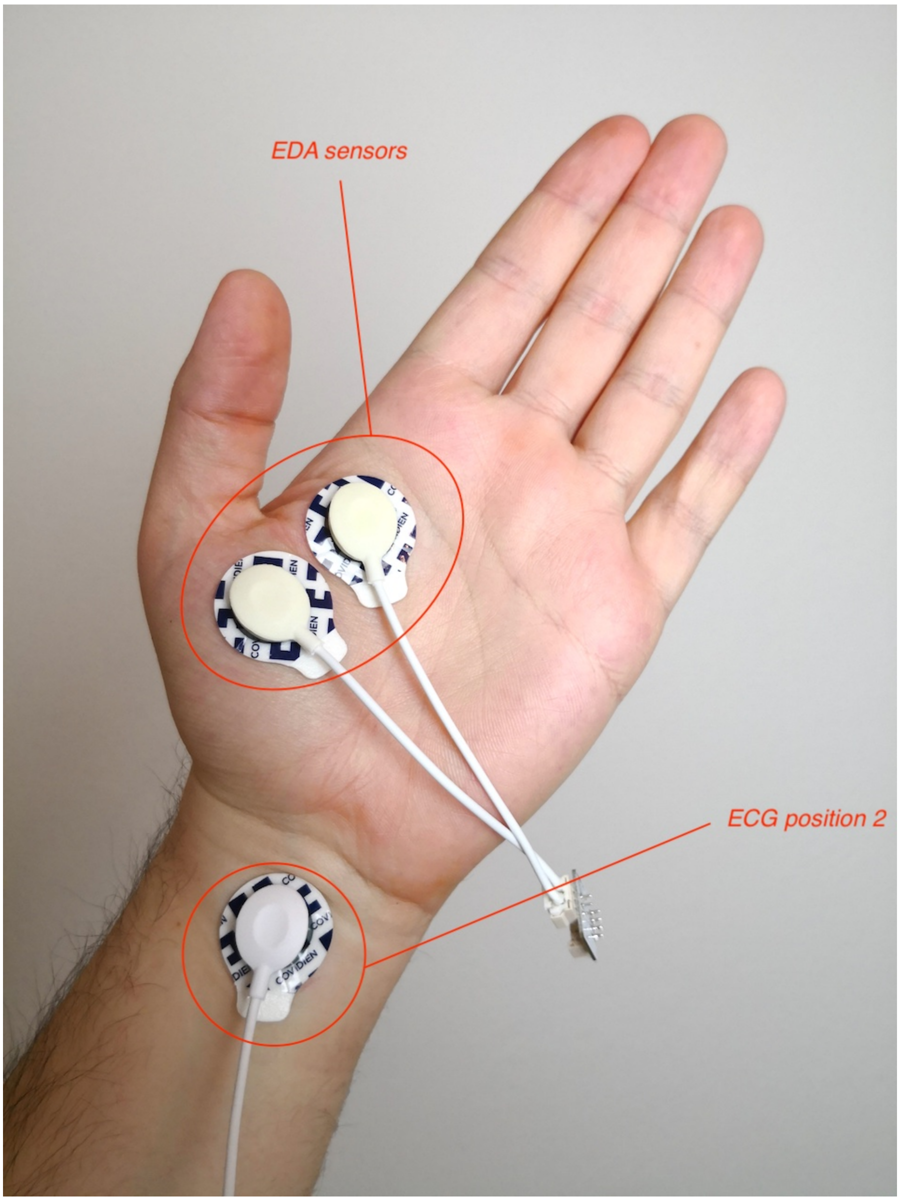
ECG sensor: ECG sensor consists of three electrode cables.[16] The outer two electrodes are positioned on skin surface where pulse can be palpated with one on the right side of the neck (carotid artery) and one on the wrist (radial artery). The ground electrode (middle cable) can be conveniently positioned in the neck area, as it only needs skin contact. The electrode cable (3-pin) is then connected with the ECG sensor board. This is then connected to a 4pin - RJ10 cable with port 3 on the bitalino board. Please make sure that the plugs are solid and fully inserted.



Graph: EMG sensor setup from the side



Graph: EMG and ECG sensor setup from the side



Graph: EDA and ECG sensor setup

5.3.4 Bitalino setup

After plugging in the RJ10 connectors for the sensors, a hardware recording switch is plugged in to I1 port of the Bitalino. The recording switch should be turned off (showing in the white dot direction). The Bitalino is now being switched on. This is indicated by a glowing orange LED. Open the Opensignals app on a smartphone and press the magnifying glass icon to search for the Bitalino (the app may ask to turn on Bluetooth, follow these directions). Click on the first entry in the list and choose the following settings:

Sampling rate: 1000Hz

Visualization rate: 100Hz

Channels: A1, A2, A3, A4

Visualization Channels: A1, A2, A3, A4

Digital: I1

Pressing the red record button will start the recording. The recorded signals will be shown in landscape mode. Before starting the experiment, it may be useful to make a test recording to see, if the sensors are connected properly. Ask the participant to tense his facial muscles, to see EMG activity. These will appear in the signal as shown in the screenshot below.

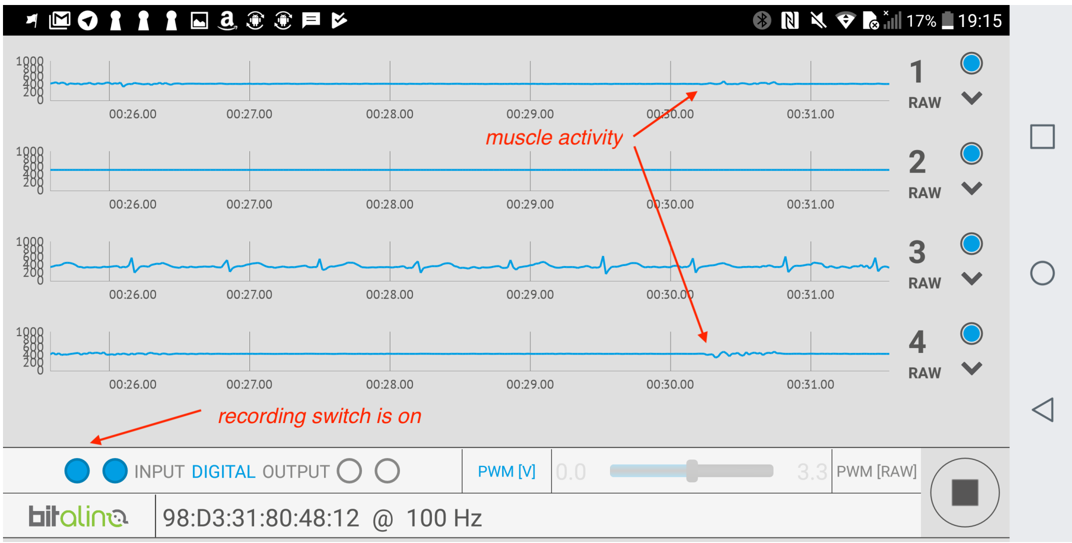


Image: Screenshot of the Opensignals app during recording, showing facial muscle activity with the hardware recording switch on

5.4 Instructions

1. Welcome the participant. At the same time, simply explain to them about the experiment purpose. After that, kindly ask the participant to sign the consent form to make sure their voluntary particiption and agreement with our usage of experiment data. disclaimer document

2. When the participant is ready, kindly ask him/her to take a seat in front of the monitor, and ensure that the participant is sitting in a straight and comfortable position. The participant will face to the monitor and camera as we described in 5.3.2.

3. When the participant is seated, set up the sensors as described in 5.3.3. Firstly, help clean the skin with wetwipes to make sure there is accurate skin conductance. Use new electrodes for every experiment. Describe shortly that the participant can expect to watch short video sequences with various content. Explain that we are going to measure the body reaction during the study by using skin sensors and a camera. Tell the participant that the whole experiment will take about one hour.

4. Calibrate the bio-sensors and start to record with the camera (ensure that participant is in the center of the record). By calibration, first of all, we start running the program from matlab and follow the instructions on screen. Then input the age and gender information of participant.

5. Switch on the Bitalino and examine if the signal recording is working well.

6. Show the user interface and introduce the participant to the system by going through the prepared example sequence and questionnaires. Ask the participant to pick their current mood in the example. Ensure that the participant has the correct seat position and answer open questions if existing.

7. Give the mouse to the participant and turn on the Bitlino recording fuction, which means recorded data from this moment will be used to be processed. Start the test in the matlab program. During the test ensure that the participant keeps focused on the test and guarantee a silent atmosphere with no disturbance.

8. During the test, one of us will wait outside in case the participant needs any help. After half of the video sequences ask the participant to take a 5 minutes mandatory break. During the break, turn off the camera and stop the bitalino processing data record. Restart when break is over.

9. When all video sequences are completely played the finishing screen appears. Please help the participant to discard the bio-sensors. Turn off the camera and switch off Bitalino.

10. Thank the participant for attendance.

11. Finish experiment

6 Data processing

Through the biological sensors and subjective rating through pick-a-mood, we can have two aspects of data, subjective data and physiological data. By analyzing all kinds of data together at the same time, we can see the relation between videos and people's affective reaction.
BITalino offers some software and tools to help analyse the data, see [1]. Matlab or Python can be used for that.

6.1 Data preprocessing

After the experiment, we can have a large number of physiological data from participants. There are many noise and useless parts in raw data. And information of sensor data, videos, and participants are separated. Before data analysis, data should be filtered, we only extract the useful part that we want.

allData
1948266x9 double

	1	2	3	4	5	6	7	8	9	10	11	12	13
1	0	1	1	1	1	430	540	297	448				
2	1	1	1	1	1	428	540	295	451				
3	2	1	1	1	1	430	540	287	449				
4	3	1	1	1	1	432	540	284	449				
5	4	1	1	1	1	433	540	275	448				
6	5	1	1	1	1	435	540	268	448				
7	6	1	1	1	1	440	540	259	448				
8	7	1	1	1	1	440	540	255	448				
9	8	1	1	1	1	438	540	255	448				
10	9	1	1	1	1	435	540	259	448				
11	10	1	1	1	1	433	540	270	448				
12	11	1	1	1	1	432	540	280	448				
13	12	1	1	1	1	432	540	295	449				
14	13	1	1	1	1	433	540	312	449				
15	14	1	1	1	1	435	540	323	451				
16	15	1	1	1	1	433	540	337	448				
17	0	1	1	1	1	432	540	350	449				
18	1	1	1	1	1	428	540	355	448				
19	2	1	1	1	1	425	540	359	448				
20	3	1	1	1	1	428	540	355	448				
21	4	1	1	1	1	432	540	350	448				

Graph: Raw data from BITalino

The picture shows the example raw data file which output from BITalino in matlab. We can see the column 2 recorded switch information, it has two state, "0" and "1". When switch is on, the switch value is "1", it means the data recorded after is what we want. When switch is off, the switch value is "0", it means the data recorded after is meaningless.

Data Separation and Regroup

6.1.0.1 For Physiological data Chanel 6, 7, 8, 9 respectively represent the different data of different sensors. Chanel 6 is EMG 1, Chanel 7 is EDA, Chanel 8 is ECG and Chanel 9 is EMG 2. In the data preprocessing part, we only use the data when switch is on. A Matlab script is run for every data file to discard the rest of the data and separate the different sensor channels. For physiological part, we abstract data for each ecg,emg,eda chanel separately while switch is "1". And saved as a out put file.

6.1.0.2 For subjective data For subjective part, there are four variables that we can use, "pick a mood", "video quality", age and gender. "Pick a mood" is the rating that participant thought what kind of mood they had after watching a particular video. "Video quality" is the rating that the participant thought was the quality of current video. Age is the age of participant, and gender is the gender of participant.

Participant				Subjective Measurements		Objective Measurements			
ID	Age	Gender	Video ID	PAM	VQ	EMG1	EMG2	ECG	EDA
1	25	m	1						
1	25	m	2						
1	25	m	3						
1	25	m	4						
1	25	m	5						
1	25	m	6						

Graph: Data sheet

The idea for the data analysis is to put all the data, subjective and objective in a data sheet as seen above, including the ID of the participant, his/her age and gender, video id, rated video quality and the pick-a-mood selection as well as the measurements of the physiological signals.

We began for each participant correlating the physiological signal data to the videos shown in the experiment. The physiological data is recorded as a continuous data stream in 1000Hz during the complete experiment. However, the relevant physiological signal is being produced only while the videos are running, not during the assessment after individual videos. For that we had to split the stream of data into subsections, which represent the affective reaction during the time when the videos are shown.
To find out the length of the videos shown and the assessments, we had to subtract the stop time from the start time of the video (taken from the Matlab assessment file). For the assessment time, we subtracted the start of the next video from the stop time of the previous video.
To correlate the video time with the physiological signal, we considered 1000 rows from the Bitalino file as 1 second from the experiment (1000Hz - 1000 data sets per second). The plan was to exclude the physiological data during the assessment time after each video. However, we encountered technical and organizational difficulties. We lacked the knowledge and time needed to implement this part of data processing.

7 Future Work

At present we have already set up a experiment system. It include the following part:

Stimuli generator (Video play)
Display external device (Monitor,speaker, mouse,keyboard)
Facial measurement (go pro camera)
Physiological measurement (with BITalino Kit)

We could collected a subjective rating and philological data from participant while they are watching the video. And we could simple plot the output data from BITalino sensors. But We still not do more in-depth data feature extraction and mood analysis. Mainly due to we lack of time, we finish it in one semester(5 month) and cooperate with Konstanz university. There's some waiting time between communication.

In the future, researcher can go on this work with this measure hardware system to collect raw data. Through a more detailed analysis, something like feature abstraction and some model analysis to find the influence between video content,quality with human mood.

8 Summary

For this research, we know that lower quality and fps will reduce the size of the video. This reduces the transmission bandwidth need and playback pressure. But if people's emotion will be influences or how people's emotion will be influenced by different quality of videos. That is the question we want to know. We researched the influence between video content, video quality, emotions and human mood. Two parts are considered, subjective part and physiological part. The experiment was set up as an experiment among 16 participants, the measurements were gathered using several methods as follows: "pick a mood" and video quality questions between every video for subjective measurements, Bitalino sensors attached to measure the EMG, EDA, and ECG signals, and finally the camera recording a video of the participant (that is to be analyzed by the Konstanz University) for objective measurements of the physiological reaction.

We associated all the data together for each participant, and worked on a code to split the matlab data to be synced with the videos played to be able to measure signal's output compared with the correct video associated with it. Unfortunately, we could not finish it in time and could not start the data processing.

References

[1] Bitalino data analysis software. .

[2] Challenges in representation learning: Facial expression recognition challenge.

[3] Emopy.

[4] Emopy codebase.

[5] Ffmpeg. .

[6] Y. Baveye, E. Dellandrea, C. Chamaret, and Liming Chen. Liris-accede: A video database for affective content analysis.[6] , 6(1):43–55, Jan.-March 2015.

[7] BioSignalsPlux. Ecg sensor datasheet. .

[8] BioSignalsPlux. Eda sensor datasheet. .

[9] BioSignalsPlux. Emg sensor datasheet. .

[10] Wolfram Boucsein. . Plenum Press, 1992.

[11] Pieter Desmet, MH Vastenburg, and Natalia Romero Herrera. 'mood measurement with a pick-a-mood:[8] Review of current methods and design of a pictorial self-report scale.[8] , 14(3):241–279, 2016.

[12] P.M.A Desmet, M.H. Vastenburg, and N. Romero. Pick-a-mood manual. , Oct 2016.

[13] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. , 9(8):1735–1780, 1997.

[14] ITU-T. P.910 :[7] Subjective video quality assessment methods for multimedia applications. .

[15] Takeo Kanade, Yingli Tian, and Jeffrey F. Cohn.[15] Comprehensive database for facial expression analysis.[15] In , FG '00, pages 46–, Washington, DC, USA, 2000. IEEE Computer Society.

[16] Davis E. King. Dlib-ml: A machine learning toolkit. , 10:1755–1758, 2009.

[17] O. M. Parkhi, A. Vedaldi, and A. Zisserman. Deep face recognition.[26] In , 2015.

[18] Sara Sabour, Nicholas Frosst, and Geoffrey E. Hinton. Dynamic routing between capsules. , abs/1710.09829, 2017.

[19] Christian Szegedy, Sergey Ioffe, and Vincent Vanhoucke.[26] Inception-v4, inception-resnet and the impact of residual connections on learning. , abs/1602.07261, 2016.