



# Automatic detection and classification of emotional states in virtual reality and standard environments (LCD): comparing valence and arousal of induced emotions

Martin Magdin<sup>1</sup> · Zoltán Balogh<sup>1</sup> · Jaroslav Reichel<sup>1</sup> · Jan Francisti<sup>1</sup> · Štefan Koprda<sup>1</sup> · Molnár György<sup>2</sup>

Received: 16 June 2020 / Accepted: 15 February 2021 / Published online: 9 March 2021  
© The Author(s) 2021

## Abstract

The following case study was carried out on a sample of one experimental and one control group. The participants of the experimental group watched the movie section from the standardized LATEMO-E database via virtual reality (VR) on Oculus Rift S and HTC Vive Pro devices. In the control group, the movie section was displayed on the LCD monitor. The movie section was categorized according to Ekman's and Russell's classification model of evoking an emotional state. The range of valence and arousal was determined in both observed groups. Valence and arousal were measured in each group using a Self-Assessment Manikin (SAM). The control group was captured by a camera and evaluated by Affdex software from Affectiva in order to compare valence values. The control group showed a very high correlation (0.92) between SAM and Affdex results. Having considered the Affdex results as a reference value, it can be concluded that SAM participants evaluated their emotions objectively. The results from both groups show that the movie section is supposed to evoke negative emotion. Negative emotion was perceived more intensely than its counterpart, positive emotion. Using virtual reality to evoke negative emotion (anger) has confirmed that VR triggers a significantly stronger intensity of emotion than LCD.

**Keywords** Virtual reality (VR) · Emotions · Classification · Valence · Arousal

---

✉ Zoltán Balogh  
zbalogh@ukf.sk

Martin Magdin  
mmagdin@ukf.sk

Jaroslav Reichel  
jreichel@ukf.sk

Jan Francisti  
jan.francisti@ukf.sk

Štefan Koprda  
skoprda@ukf.sk

Molnár György  
molnar.gy@eik.bme.hu

<sup>1</sup> Department of Informatics, Faculty of Natural Sciences, Constantine the Philosopher University in Nitra, Tr. A. Hlinku 1, 949 74 Nitra, Slovakia

<sup>2</sup> Department of Technical Education, Budapest University of Technology and Economics, Budapest, Hungary

## 1 Introduction

Recently not only psychology but other scientific disciplines have placed the emotional state in their focus of inquiry. Emotional state can play a significant role in such areas as education, driving a motor vehicle, health care, but in case of smart homes as well. Current automatic recognition systems are based on three basic phases: user detection, extraction of areas of interest and subsequent classification of emotional state.

In general, we can apply the following methods to the classification of the emotional state by the help of the automatic recognition system (Magdin et al. 2019a; Marín-Morales et al. 2018):

1. Face capturing and areas of interest using a camera,
2. Placement of sensors on different parts of the body (e.g. Galvanic Skin Response—GSR, Heart Rate—HR, Temperature, Electroencephalography—EEG) and processing the respective data
3. Voice expression analysis,

#### 4. Abnormal characteristics activity capturing (body movements, speed of response to stimuli ...).

Camera-assisted classification of emotional state has become standard practice nowadays (Ekman et al. 2002; McDuff et al. 2016). A number of Software Development Kit (SDK) solutions developed over the last 20 years (Microsoft AI as Face API, Kairos, Face ++, BetaFace, OpenCV, Eye Face and others) include algorithms for detection, extraction and classification. Affdex from Affectiva is also a standardized SDK solution (McDuff et al. 2016). This complex approach relies on a multipoint mask for implementing the first two phases of the recognition process. Based on the classification algorithm (Support Vector Machine—SVM is used for classification) corresponding with the standardized Ekman classification we can determine the emotional state of the user (Ekman et al. 2002; Ekman and Friesen 1978).

In the article “A Case Study of Facial Emotion Classification Using Affdex” (Magdin et al. 2019a) it was demonstrated that the Affdex SDK works predominantly with instructed emotional states—it can capture and classify emotional states that are external to the user.

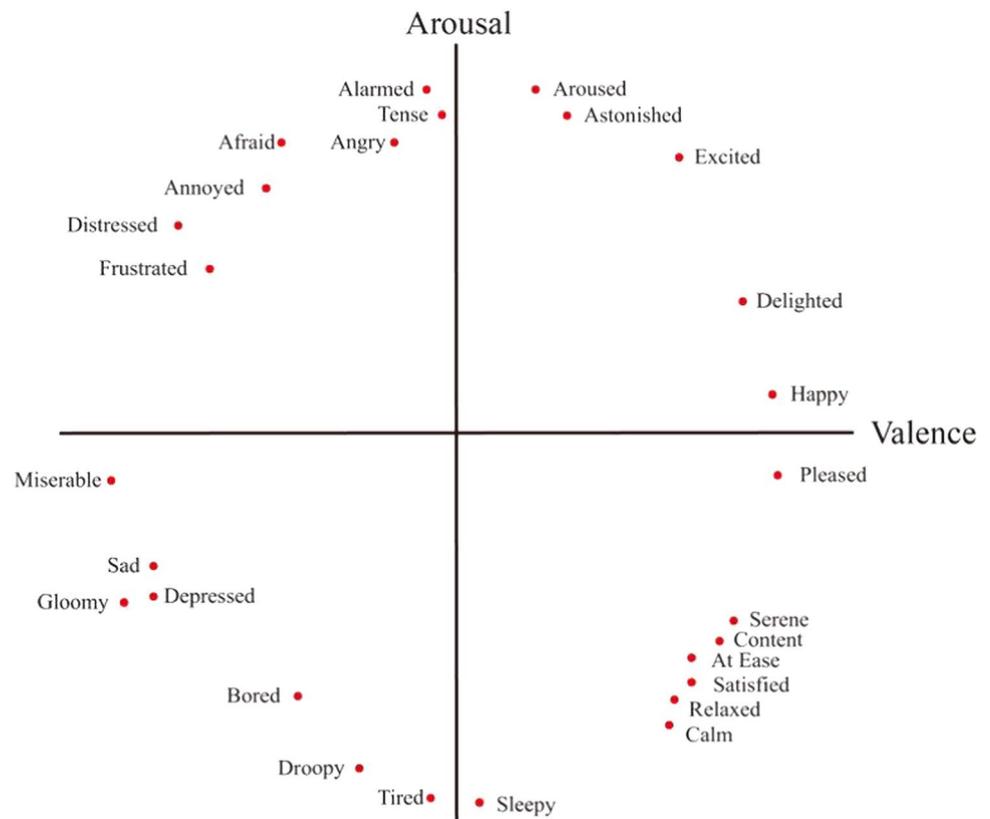
In the case study described in this article, we subjected the participants to an experiment in which we projected sections of movies from the standardized LATEMO-E database from Micheline (Micheline et al. 2019). The

individual sections of the movies are categorized in the LATEMO-E database according to the Ekman classification (Ekman et al. 2002), containing the appropriate value of valence and arousal for each movie section. The Russell model (Russel, 1979) was applied to extend the Ekman classification. The value of valence is an important feature in determining the positive or negative aspects of a given emotion (Fig. 1).

The research aimed to compare two groups of participants—in the experimental group we projected movie sections through VR (Oculus Rift S, HTC Vive Pro), the control group watched the movie sections on the LCD monitor. By evoking an emotional state, we determined the range of valence and arousal in both observed groups. We measured valence and arousal in both groups using the Self-Assessment Manikin—SAM (Bradley and Lang 1994). In addition, the second group was captured by a camera and evaluated by Affdex software (McDuff et al. 2016), which allows to measure the size of valence in the range  $[-1; 1]$ , where values approaching  $-1$  express negative emotion and values approaching  $1$  express positive emotion (Russell 1979). Based on the above, research questions and related hypotheses were formulated (listed in subchapter titled Results of Experiment):

Q1: Does virtual reality evoke a more pleasant emotion than a classic monitor?

**Fig. 1** Russell’s circumplex model of emotional states (Seo & Huh 2019)



Q2: Does virtual reality evoke a more intense emotion than a classic monitor?

Q3: Are the evoked negative emotions more intense than the evoked positive emotions?

In the Related Work section, we present an overview of the current research effort. The combination of exploring emotions and virtual reality is currently a much-discussed topic as virtual reality is gradually becoming part of our daily lives [VR is not only actively used for playing games but also in the learning process, see e.g. (Allcoat and Mühlhelen 2018; Garzotto et al. 2019; Demitriadou et al. 2020; Chang et al. 2020)]. In the Related Work section, we focus on works that examine the induction and classification of emotional states from different perspectives, among them psychology, but also data processing—artificial intelligence, computer vision. In the Material and Methods section, we present the methodology of performing the given experiment. The findings of the experiment are described in the Results section. In the Discussion and Conclusion section, we present a comparison of our achieved results with the findings of other authors.

## 1.1 Related work

Virtual Reality presents an interactive (simulated) environment. In this environment, users can interact with various objects. These objects and environment scenes are generated with a computer (Chang et al. 2020). Pan et al. (2006) mention that VR is generated by using the computer as a 3D environment. Virtual reality has developed at a breakneck pace in recent years and introduced major improvements in education, therapy, media, medical training and other areas (Ding et al. 2018; Loureiro et al. 2019).

Virtual reality can be used in a variety of ways. This technology is found in many segments such as education, medicine (Zinchenko et al. 2020; Javaid and Haleem 2019), tourism (Loureiro et al. 2020), but it also serves leisure time purposes (Mohamed Elias et al. 2019; Jang and Park 2019).

Research over the last decade suggests that the use of VR devices also affects human emotions (Vera et al. 2005; Han et al. 2009; Pizzi et al. 2019; Garzotto et al. 2019). There is a significant amount of research focusing on the change in the emotional state during the use of virtual reality (Han et al. 2009; Pizzi et al. 2019).

Authors (Felnhofer et al. 2015) examined whether five different virtual park scenarios would evoke a specific affective state (joy, sadness, disinterest, anger and anxiety) in the respondents. The results of the research showed that almost all five scenarios of the virtual park were able to induce a given affective state. Felnhofer (Felnhofer et al. 2015) also asserts that it is very important that participants are in a relaxed state when they enter a virtual park scenario. This condition is a crucial reference point for measuring

other affects on emotion. As far as hardware application is concerned, this study brings interesting results. The author underlines that a VR model Oculus Rift is very comfortable for human usage as a device facilitating sufficient immersion in VR (Felnhofer et al. 2015). Conversely, it is argued that the overall experience in VR may not be related to the specific technology but rather to the sensory properties of the user.

Diemer suggests that emotional immersion related to VR is not always unambiguous (Diemer et al. 2015). According to Visch, it depends on the type of VR devices and on the valence of the examined emotional states and the value of the evoked arousal (Visch et al. 2010).

These realized studies (Felnhofer et al. 2015; Diemer et al. 2015; Visch et al. 2010) are very important, because they underline the need to determine the so-called reference (zero) point. It is the point in time when the participant is in a state of peace (this emotional state is defined as "neutral"). Any deviation from this point representing a change is actually an affect (or emotion, depending on the duration). These studies lead to the conclusion that the presence of emotional states is unequivocally related to the human perception of a specific environment.

According to Ding (Ding et al. 2018) the video clip has a different emotional impact on humans in the VR environment as compared to a more traditional context (LCD). This difference is related to immersion and human presence in the given environment types. The degree of immersion is connected with sensors and parameters of the environment (e.g. field of view). Presence entails the human perception of a specific environment.

While emotions are basic aspects of the human experience, they are difficult to define, recognize and classify. People are different from each other and each person expresses emotions in different ways. According to the way emotions are expressed three human types were distinguished in the past—sanguine, choleric or pessimist (Makki et al. 2019). Today's classification of human types is not that simplistic because of the interpersonal differences in perception, processing and expression of emotions. Emotions should reflect a person's inner state, how he feels. This should most often be seen on his face (Ekman's classification is also based on this). However, based only on the face it is not always easy to recognize this condition. Recognition of emotions and their classification is especially problematic, for example, in some people with paralysis in the facial area. In this case, it is necessary to use other assistive technologies (sensors) that can reveal hidden emotions (Makki et al. 2019; Vanderlind et al. 2020).

The issue of recognizing and classifying human emotions (not only by using a computer) is a very interesting topic that has received increasing attention in recent years. Recognition and classification of human emotions are used

not only in education and psychology but also in the field of applied informatics—e.g. smart homes and others (Francisti and Balogh 2019a, b). In 1980, Plutchik described wheel-shaped emotions (Plutchik 1980). There were eight primary emotions paired in the circle: joy / sadness, anger / fear, trust / resistance and surprise / expectation (Wang et al. 2019; Merta 2019).

Another model for classifying emotions is the Lövheim emotional cube (Lövheim 2012). This model consists of three active substances (dopamine, norepinephrine and serotonin) that evoke emotions and the eight basic emotions are located in the corners of the cube (Leukhin et al. 2018; Lövheim 2012). Ekman and Friesen developed the Facial Coding System (FACS) for coding facial expressions in 1978, in which facial movements are described by a set of action units (AUs). Each AU is based on a muscle relationship. Assessing AU at a given time provides the ability to identify up to seven types of emotions: anger, disgust, fear, happiness, sadness, surprise and being neutral (indifference) (Waller et al. 2020; Ekman et al. 2002; Ekman and Friesen 1978).

Ekman's model can be used especially in direct face view; it can be deployed in artificial intelligence and computer vision for the design and creation of an automatic recognition system (ARS), which has been developed by researchers dealing with artificial intelligence for almost 40 years. Three phases are used in ARS: face detection on the image, extraction of areas of interest from the face and subsequent classification of the emotional state (Goshvarpour and Goshvarpour 2018; Hossain and Muhammad 2019; Imani and Montazer 2019; Kerkeni et al. 2019; Sharma et al. 2020; Tonguç and Ozaydın Ozkara 2020; Zhang et al. 2020).

The problem with Ekman's classification is that emotions often arise quickly (affect), based on the automatic cognitive evaluation, with little involvement of consciousness coupled with unintentional changes in expression and physiology. We experience them as they happen, but we cannot change them (e.g. stress). Therefore, Ekman's model does not capture all the potentially expressed emotions. If we consider stress as another factor that affects our emotions and their distribution, then it is not possible to use the Ekman classification in ARS, but it is important to understand how people react to events that can cause stress (Cruz-Albarran et al. 2017; Bahreini et al. 2016).

The problems associated with the use of the Ekman model necessitated another approach in obtaining input data for the detection phase. In addition to looking at the camera, various wearable physiological sensors were deployed in the practical application of the Ekman classification in which valence and arousal are applied from the Russell Circumplex Model (Fig. 1).

This model describes emotions as a circle in a two-dimensional bipolar space—it is divided by valence and

arousal. According to Russell, valence can be either positive or negative, being in the range  $[-1; 1]$ . This means that with the help of valence we can determine a neutral state, negatively perceived emotions (e.g. sadness, anger, fear, etc.) or positive emotions (joy, surprise, tenderness and others). Arousal is determined by an equally large interval. Anger can be identified as a combination of extreme resentment and medium–high arousal (Russell 1979).

Russell's model is used wherever emotions can be identified and classified by other methods than just looking into the camera. Russell's approach can also explain emotions that are not captured by default (directly visible) in a person's face (Alberdi et al. 2016; Quazi and Mukhopadhyay 2011; Shalini and Vanitha 2013; Kaklauskas et al. 2011; Yang et al. 2019).

From a psychological point of view, we are using various methods for evaluation of emotions. The most commonly used psychological methods include the self-evaluation techniques based (typical SAM) on various, especially psychological questionnaires (Isomursu et al. 2007). We can often rely on machine learning methods for the evaluation of the measurement parameters of human physiological signals (Gross 2015; Mahlke et al. 2006). According to a research (Gonçalves et al. 2017; Scherer 2005; Molnár et al. 2018), emotion has various properties and we can therefore analyse them via behavioural or physiological reactions but also subjective feelings (these are usually assessed by only using SAM).

Currently various software can facilitate the automatic recognition and classification of emotions. This software was developed with the help of psychologists and researchers in the field of applied informatics (cybernetics, artificial intelligence and others) (Picard 2009). Different physiological parameters of the human body (often typical electrical impulses—electroencephalography, skin resistance, blood pressure, heart rate or temperature) are measured in the recognition and classification phases. The measurement of physiological sensors is a complex process in which a large number of questionnaire items are asked (Poria et al. 2017). Francisti and Balogh present good examples of how we can gauge the structure of emotions based on various physiological signals (Francisti and Balogh 2019a, b). These authors researched the possibilities of using smart wristbands to obtain physiological states and subsequent classification of emotional states. More attempts have been made to classify emotions, sensors and universal selection algorithms. The first step consists of the selection of measurement parameters and methods, while the second step entails the selection of sensors and devices.

In the next section, we present our material and methods for the classification of emotions, the measurement of valence and the arousal of emotions.

## 2 Materials and methods

The experiment was performed under the supervision of a psychologist in March 2020. The groups were comprised of an equal number in terms of sex and age. Each group consisted of 20 participants (students). In each group, we had 5 women and 15 men and the age of each participant was in the interval of 21–25 years. Each of the participants was informed about the course of the experiment (not about the content). At the beginning of the experiment, participants completed a questionnaire concerning age, sex and possible mental or other diseases. Participants were shown various clips representing high arousal and valence value intervals. Our aim was to capture the affect and then classify the emotional state. The affect is a very short but intense state of emotion. In the case of a participant suffering from heart disease, it would not be possible to perform this experiment since the resulting emotion of fear is likely to damage one's health. The same applies in case of mental illness, which in a sense can affect a person's emotional state. Prior to each launch of the movie section from the standardized LATEMO-E database, participants in both the experimental and control groups were instructed by psychologists to breathe calmly (their steady-state ECG (electrocardiogram) signal was measured every 25 ms). We measured the ECG signal similarly to Felnhofer (Felnhofer et al. 2015) to ensure that participants are in a relaxed state (reference point from which we measured other induced emotions). Therefore, various time gaps were installed for each participant between the presentations of the clips to complete the respective questionnaire. When measuring the parameters, we did not notice any significant changes between women and men, therefore we considered the data to be mutually consistent.

In this way, we eliminated random input data (possible influence of emotional state). In the control group, the data from the Affdex measurement were synchronized with each other—after the start of the movie sections all acquired data were automatically recorded on the basis of a multipoint mask in the MySQL database at an interval of 25 ms. Our own EmoSens application was used for measuring (Magdin et al. 2019b) which synchronized all data with each other. The experimental group watched the movie sections via the VR device Oculus Rift S and HTC Vive Pro. The control group watched the movie sections through an LCD monitor.

The hardware used in case of the control group was standard PC with the Windows 10 Educational 64bit operating system and with the following parameters:

- CPU Intel(R) Core(TM)2 Duo CPU E7500 @ 2.93 GHz
- RAM 4096 MB

- GPU NVIDIA GeForce GT 710
- HDD SSD 1 TB
- LCD ASUS 27" VG27WQ Gaming—Quad HD 2560×1440, 16:9

Participants sat at a distance of 1 m from the monitor (calculated from the diagonal and resolution of the monitor). We tried to keep this distance throughout the implementation of the experiment. For the VR device, we set the "same parameters" as for the classic environment. People wearing glasses did not participate in the experiment. The viewing angle of the LCD Asus monitor is 178° on a curved display. The problem of HTC Vive Pro is its field of vision being only 110° but with a range of 360°. Therefore, we applied the same canvas degree in VR as well as in the classical environment.

Only one participant at a time was shown a movie clip in a separate room in order to prevent influencing the respective results. Movie sections from the standardized LATEMO-E database were selected on the basis of their highest values of valence and arousal, in order to evoke a given emotional state as much as possible (we selected clips with high values of valence and arousal). The sequence of the clips varied. The participants perceived audio from the clips in the native language. After watching the movie section (data were automatically recorded in the database), participants were asked to indicate the type of emotional state that the movie section evoked in them (to check back whether they can actually identify the emotion) and evaluate this state through SAM for valence and arousal on a scale from 1–9. A 5-point scale is used in SAM by default, but this scale can be extended (Tsonos et al. 2008). The advantage of the 9-point scale is that respondents can express their views in more detail, how they feel by placing an "X" either on a specific image or between images (e.g. radiobutton). All values of valence and arousal were recalculated to the interval  $[-1; 1]$ , where the value 5 represented the number 0 in the interval, the value 9 stood for the number 1 and the value 1 represented the number  $-1$ . In the case of the experimental group, data were not recorded using Affdex, as neither Oculus Rift S nor HTC Vive Pro contains any standard cameras for face capture, hence it is not possible to capture the participant's face or a specific area of interest (for example mouth). However, in the case of the control group this method was applied. Based on a 24-point mask, the participant's face was captured, areas of interest were evaluated, and the Affdex software subsequently classified a specific emotional state with valence. The control group, as well as the experimental group, evaluated their evoked emotional state by SAM after watching each of the movie sections for valence and arousal on a scale from 1–9. The following movie sections were selected from the standardized LATEMO-E database (Table 1):

**Table 1** Selection of movie sections according to the type of evoked emotion from the LATEMO-E database

Induced emotion	Name of the movie	Length of the movie [min:sec]
Anger	12 Years a slave	04:51
Disgust	The human centipede II	01:26
Fear	The conjuring II	03:23
Sadness	The impossible	07:05
Joy (Happy)	Blended	02:55
Tenderness (in Russell's model this emotional state is marked as at ease)	He's just not that into you	02:59
Neutral	Holy motors	01:01

### 3 Results of the experiment

The obtained data were analysed to identify a potential significant difference in the perception of one's emotion in the movie experience evoked by VR and traditional LCD monitor. We examined several partial facts and relationships as well. The hypotheses we tested were as follows:

H1: There is a statistically significant difference between the assumed emotion (level of Valence) evoked by virtual reality and the monitor.

H2: There is a statistically significant difference between the intensity (level of Arousal) of emotion evoked by virtual reality and the monitor.

Hypotheses H1 and H2 were tested overall for the whole data set and also separately for the individual emotions examined.

H3: There is a statistically significant difference between the intensity of perception of negative and positive emotions.

It is necessary to verify the reliability of our measurement of valence through SAM. Because we also used face and facial recognition as well as subsequent emotion classification through Affdex software for LCD video playback, we can use the Affdex detection results as a reference. We used TIBCO Statistica software for statistical data processing.

The correlation between the results of valence given by the respondents through SAM and the results shown to us by Affdex was examined by Pearson's selection correlation coefficient:

$$r_{XY} = \frac{s_{XY}}{\sqrt{s_X^2 s_Y^2}}$$

where  $s_{XY}$  is the covariance of variables  $X$  (SAM) and  $Y$  (Affdex) and  $s_X^2$ ,  $s_Y^2$  are the variances of the variables  $X$  (SAM) and  $Y$  (Affdex).

By evaluating the correlation between the results of valence given by the respondents through SAM and the

results shown to us by Affdex, we can identify similarities or overlaps between these results. We used correlation analysis to evaluate it. The variables SAM and Affdex were found to be strongly correlated,  $r(138) = 0.92$ ,  $p < 0.001$ .

Correlation analysis showed that there is a 0.92 dependence between them. This means that the results obtained from the SAM are credible and the respondents objectively assessed their own emotional state. To verify hypotheses H1 and H2, we used a t test for independent variables. The precondition for the use of the t test is equality of variance, which was verified using Lavene's test and based on which the condition was met. The second condition is the normal distribution of the compared groups of data, which is also met.

To verify hypotheses H1 and H2 the t test for two independent variables was used (Munk and Drlík, 2011) and the following formula was applied:

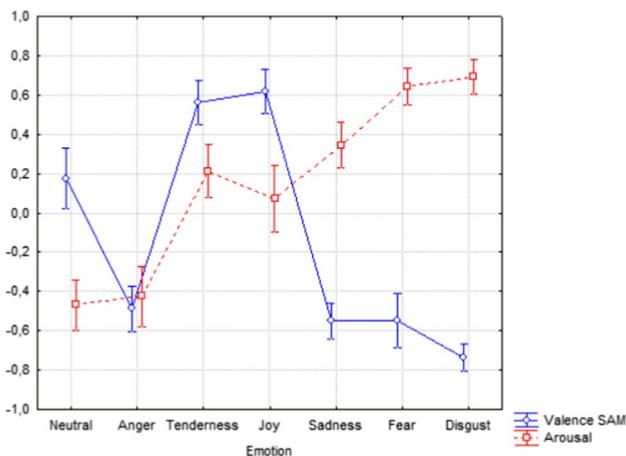
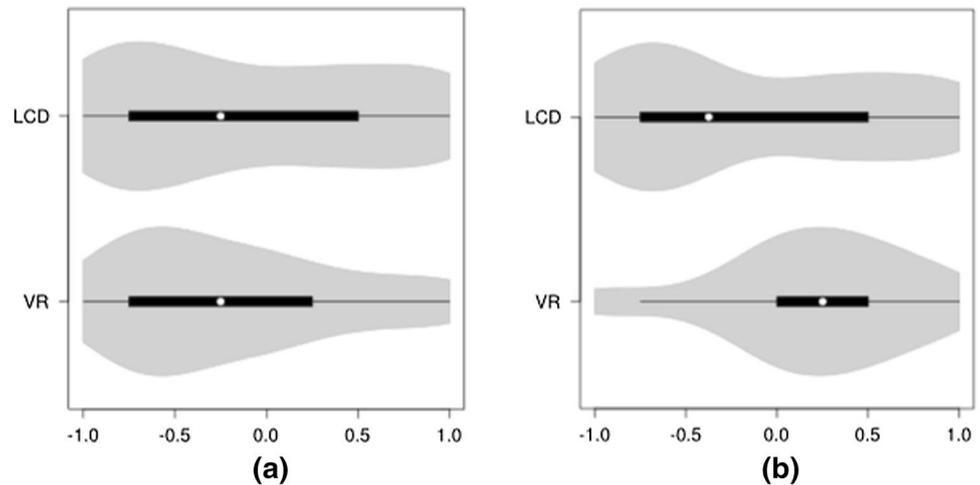
$$T = \frac{\bar{x} - \bar{y} - (\mu_1 - \mu_2)}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}},$$

where  $\mu_1$ ,  $\mu_2$  are estimates of the mean values of the groups of values,  $\bar{x}$ ,  $\bar{y}$  are the means of groups of values,  $n_1$ ,  $n_2$  are the numbers of values of groups of values and  $s_1^2$ ,  $s_2^2$  are estimates of the variances of groups of values.

Using a t test, we compared the mean values between the VR and LCD groups (Fig. 2) in valence (from SAM) and Arousal (for all affects together). The exact results are:

- there was no statistically significant difference for *Valence SAM* in the used method for VR
- ( $M = -0.19$ ,  $SD = 0.58$ ) and LCD ( $M = -0.09$ ,  $SD = 0.69$ ); conditions ( $t(278) = -1.37$ ,  $p = 0.171$ ),
- there was no statistically significant difference for *Arousal* in the used method for VR ( $M = 0.2$ ,  $SD = 0.5$ ) and LCD ( $M = 0.1$ ,  $SD = 0.66$ ); conditions ( $t(278) = 1.42$ ,  $p = 0.156$ ).

**Fig. 2** Violin plot of *Valence SAM* (a) and *Arousal* (b) values for independent VR and LCD variables



**Fig. 3** Interaction graph of *Valence SAM* and *Arousal* values for individual emotions

In both cases the analysis showed there was no statistically significant difference between the group VR and LCD. We do not reject either H1 or H2 with 95% confidence. This means it generally does not affect the level of the identified valence or arousal whether the experiment was carried out using virtual reality or using a monitor.

Since the expectation was that using a device of virtual reality would have a bigger significant effect at least on the intensity of emotions (*Arousal*), we also looked at the results of testing individual emotions (Fig. 3).

Figure 3 visualizes average values with 95% confidence intervals. Accordingly, the difference in the value of valence for positive emotions such as joy or tenderness compared to negative ones, i.e. anger, sadness, fear and disgust can be identified.

For the emotions of tenderness (VR ( $M = 0.25$ ,  $SD = 0.34$ ), LCD ( $M = 0.18$ ,  $SD = 0.5$ );

( $t(38) = 0.55$ ,  $p = 0.584$ ), sadness (VR ( $M = 0.33$ ,  $SD = 0.36$ ), LCD ( $M = 0.36$ ,  $SD = 0.37$ );

( $t(38) = -0.32$ ,  $p = 0.747$ )) and fear (VR ( $M = 0.58$ ,  $SD = 0.3$ ), LCD ( $M = 0.71$ ,  $SD = 0.26$ );

( $t(38) = -1.53$ ,  $p = 0.133$ )), it was confirmed that there was no difference between the testing methods.

Figure 4 shows a better visualization of the comparison of interaction graphs for the variables *Valence SAM* and *Arousal* when comparing the VR and LCD methods for individual emotions separately.

Interesting results were shown in case of the emotion, joy, although the difference between the testing methods revealing valence and arousal was also not confirmed. The overall results of the correlation between the *Valence SAM* and *Arousal* values showed a dependence,

$$r(278) = -0.19, p = 0.001.$$

At a significance level of 5%, this relatively low correlation of  $-0.19$  is also statistically significant. This is a negative correlation which means that the higher the valence, the lower the arousal and this is valid vice versa.

Thus, when negative emotion was evaluated (such as fear, sadness, etc.) arousal, which is a certain form of intensity, acquired higher values, and in case of positive emotions, arousal was rather lower (on the whole sample—VR and LCD together).

It follows that negative emotions were felt more intensely than positive ones. However, with the emotion of joy, it was the opposite.

Correlation between the variables *Valence SAM* and *Arousal* was positive and even statistically significant for joy,  $r(38) = 0.33$ ,  $p = 0.037$ . This means that the positive emotion of joy was perceived more intensely than other positive emotion, such as tenderness,  $r(38) = 0.11$ ,  $p = 0.49$ .

If we only look at the results of anger testing, the data show a statistically significant difference between the *Arousal* values, which is also evident in the averages of

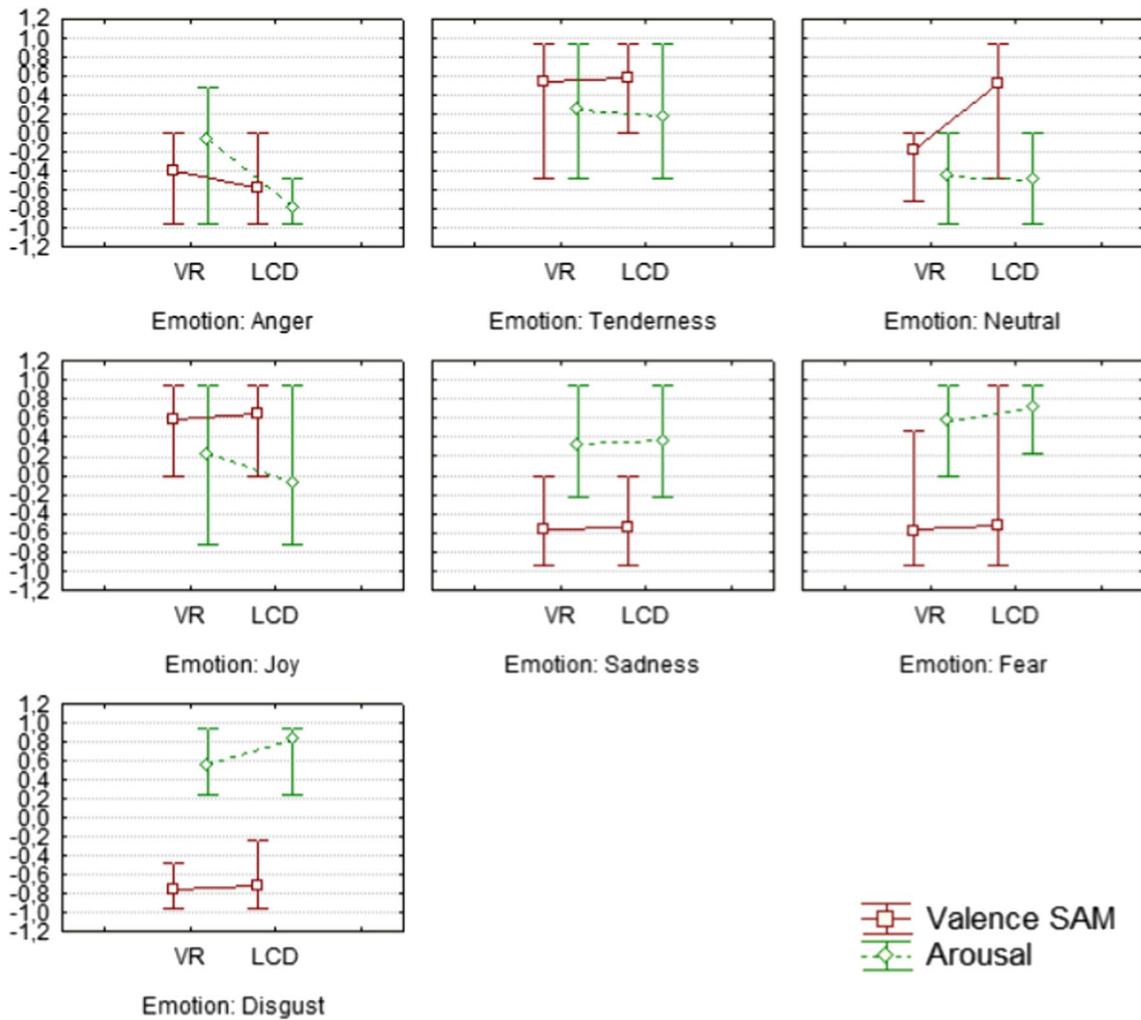


Fig. 4 Interaction graphs on comparing *Valence SAM* and *Arousal* values obtained by using VR and LCD methods separately for each emotion

the VR and LCD variables. The results of the independent samples t test, the preconditions of which were also used for this analysis, shown in Fig. 5 that the emotion of anger was perceived significantly weaker through the monitor than through VR. There was a statistically significant difference in *Arousal* for the emotion of anger for VR ( $M = -0.06$ ,  $SD = 0.4$ ) and LCD ( $M = -0.79$ ,  $SD = 0.19$ ) conditions;  $t(38) = 7.28$ ,  $p < 0.001$ .

The analysis of data for the emotion of disgust also showed relatively interesting results. In Fig. 6, we can see a statistically significant difference in *Arousal* values, which means through the monitor this emotion is perceived more intensively. There was a statistically significant difference in *Arousal* for the emotion of disgust for VR ( $M = 0.56$ ,  $SD = 0.25$ ) and LCD ( $M = 0.83$ ,  $SD = 0.24$ ) conditions;  $t(38) = -3.32$ ,  $p = 0.002$ .

The video database also contained movie sections that were associated with neutral emotion. Therefore we expect

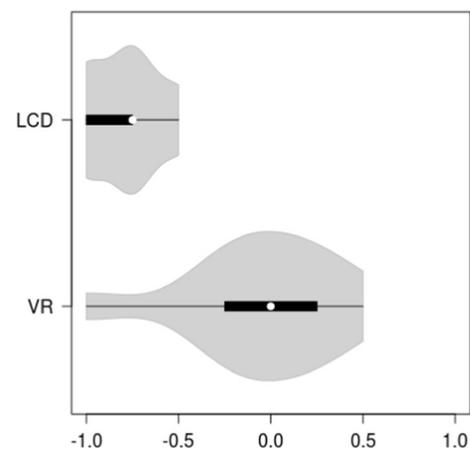
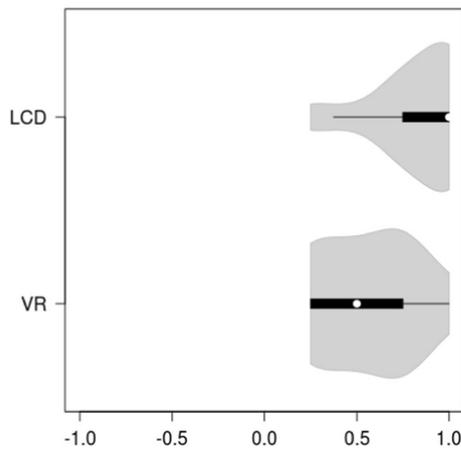


Fig. 5 Violin plot of *Arousal* values for the independent variables VR and LCD for the emotion of anger



**Fig. 6** Violin plot of *Arousal* values for the independent variables VR and LCD for the emotion of disgust

that the values of valence should be around 0, as neither positive nor negative emotions should exceed. However, testing showed that this movie section was perceived rather positively through VR and slightly negatively through the monitor. There is also a statistically significant difference between these averages, meaning a significant difference can be discerned between how positively/negatively a movie section is perceived when using VR and when using a monitor (Fig. 7). However, arousal came out relatively weak in both cases. Despite the discrepancy in whether the movie section is perceived positively or negatively, it displayed a relatively weak intensity of emotion (Table 2).

Hypothesis H3 "there is a statistically significant difference between the intensity of perception of negative and positive emotions". We verified this by dividing the records in two ways:

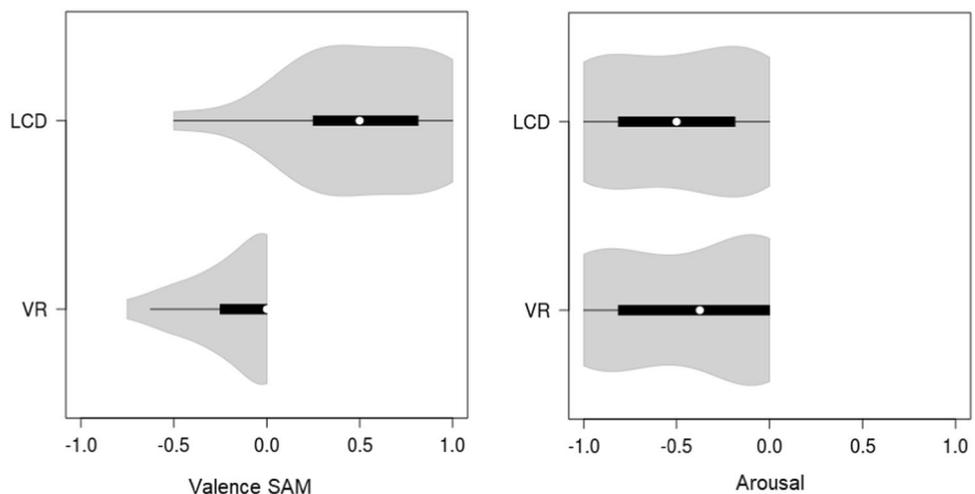
**Table 2** Arousal t test for independent variables VR and LCD for neutral emotion

	Mean VR	Mean LCD	<i>p</i> value
Valence SAM	0.525000	−0.175000	0.000000
Arousal	−0.487500	−0.450000	0.771735

1. The first way was the subjective division of records into those relating to positive (Positive1) emotion (tenderness, joy) and those relating to negative (Negative1) emotion (fear, disgust, anger, sadness). Based on this division into positive and negative we compared the arousal values.
2. The second way was to divide by the value of *Valence SAM*, which was given by the respondent himself. If the participant stated a value from −1 to −0.25, the value of arousal from his record was included among the negative records (Negative2). If the participant stated a value from 0.25 to 1, we included the value of arousal among the positive records (Positive2).

For both sample distribution methods, those with a value of 0 were excluded from the analysis because they could not be assigned to any of the mentioned groups. For both methods of distribution [subjectively to researchers (1.) and according to valence (2)], we used a t test for independent variables, where we examined the mean values of the variable *Arousal*. The conditions for using the t test for independent variables are also met in this case. The results show that there is a statistically significant difference in both cases between the averages for group Negative and group Positive. There was a significant difference in the *Arousal* for group Negative1 ( $M=0.31, SD=0.57$ ) and group Positive1 ( $M=0.14, SD=0.48$ ) conditions; ( $t(238)=2.28, p=0.023$ ). As in the previous case, there was a significant difference in

**Fig. 7** Violin plot of *Valence SAM* and *Arousal* values for the independent variables VR and LCD for the emotion of neutral



the *Arousal* for group Negative2 ( $M=0.29$ ,  $SD=0.59$ ) and group Positive2 ( $M=0.07$ ,  $SD=0.53$ ) conditions; ( $t(250)=-2.98$ ,  $p=0.003$ ). Thus, we do not reject hypothesis H3 and we can also say that respondents perceived negative emotions significantly more strongly than positive ones.

## 4 Discussion

We focussed on the comparison of the subjective user experience and the evoked emotional states during the implementation of the experiment, when watching movie sections from the standardized LATEMO-E database. The experiment was carried out in a VR environment and a 2D environment (watching a movie section on an LCD monitor). Based on the same input data (movie sections), whose task was to evoke a specific emotional state in the participants, we designed a comparative experiment. An equally large sample of participants was used for both the experimental (VR environment) and control group (LCD monitor). We recorded each experiment and used subjective questionnaires for valence and arousal values for statistical analysis, which the participants always filled in after the end of the movie section (after displaying a movie section with the appropriate affect). Statistical analysis showed several significant differences that arose between the experimental and control groups in terms of valence and arousal. Russell's circumplex model presents a description of evoked emotional states using the range of valence and arousal. The value of valence represents the positivity/negativity of the perceived emotion, and arousal refers to the extent it is maintained.

If we compare the results, where the correlation between the variables valence and arousal was positive and even statistically significant, the values measured by us and the results from the statistical evaluation correlate with Russell's circumplex model. Even in this model, the positive emotion of joy was perceived more intensely than other positive emotions (e.g. at ease), while the valence of this value is very high according to Russell's model and also acquires a positive value of arousal.

In the evoked emotion of anger, the measured data give rise to a statistically significant difference between the arousal values, which is also evident in the averages of the VR and LCD variables. The results visualized in Fig. 5 show that the emotion of anger was perceived significantly weaker through the monitor than through VR. Russell's model shows that it is an emotion with a negative value of valence, but a high value of arousal. By evoking this emotion while watching *12 Years a Slave* through VR (senseless whipping of a slave girl), it is understandable that participants felt the emotion affect more intensely.

The results measured by our experiment show that Russell's model is mainly justified in determining a particular

emotion in terms of valence, i.e. its positivity/negativity, which was confirmed by the averages, where a statistically significant difference could be discerned. This means there is a significant difference between how positive/negative movie sections are perceived when using VR and when using a monitor.

The results we obtained correlate with the results of a similar experiment (Niu et al. 2019), during which the authors of the experiment also compared the valence and arousal of evoked emotions in VR and 2D environments. The evoked emotions were evaluated on the basis of a subjective questionnaire (similar to ours) and by changing the values measured by physiological signals ECG, SKT (skin temperature) and EDA (electrodermal activity). They used signal measurements as a source of reference data to determine whether physiological signals actually changed with the evoked emotion. We used the standardized Affdex software from Affectiva as a source of reference data for determining valence and arousal. From the other authors' experiment (Niu et al. 2019) it follows that the average score of positive emotions (mainly for joy) is significantly higher in the VR environment than in the 2D environment (when viewed via LCD). Niu (Niu et al. 2019) also states that although there are no significant differences in valence and user arousal values between the two environments, ultimately all average valence values in the VR environment are slightly higher than in the 2D environment. We can agree with this statement, as we achieved the same results in the implementation of our experiment. Niu (Niu et al. 2019) at the same time states that during the implementation of the experiment through the VR there are symptoms of the so-called simulator disease, where participants may reach a state of fatigue or dizziness, which may affect the final results. This situation occurred to us only in one case, when we had to interrupt the experiment and repeat it after the participant calmed down [therefore, we repeated the experiment with this participant with a time interval (1 week) in order not to distort the measured data].

The reason why we could not apply Affdex even with VR is based on the limited current capabilities of VR devices that do not include standard face cameras (technically impossible). At present, technological solutions are already emerging that would make it possible in the future to measure the value of valence and arousal based on the movement of pupils. According to Nam (Nam et al. 2019), it is possible to add sensors to the HTC Vive headset, which would enable this feature. However, since it is still an experimental modification of the device, the results of the experiment could not be considered relevant. Other approaches with upcoming hardware may include measuring and interpreting muscle activity or even analysing brain waves (see e.g. <https://www.mindmaze.com>).

## 5 Conclusion

In a previous study, A Case Study of Facial Emotion Classification Using Affdex (Magdin et al. 2019a) we pointed out the lack of Affdex. In the research, we focused on evoking emotional states through images from a standardized NAPS database (Marchewka et al. 2014), where software failed to classify emotional states and valence assessment did not correlate with participant assessment (SAM).

In this case, we can state that Affdex can be considered as a reference tool for experimental solutions to issues related to valence in evoked emotional states. Nonetheless, the use of this tool is strongly dependent on the type of medium deployed to evoke the emotional state.

The analysis of the obtained data showed that the answers of our respondents were highly correlated (0.92) with the valence reference values, which were obtained by standard Affdex methodologies. Based on this, we can consider that the respondents' answers are credible.

The article aimed to verify whether the respondents will perceive emotions more strongly when using virtual reality compared to using a standard computer monitor. A significant difference was not confirmed here. Although this difference was not confirmed overall, partly due to the emotion of anger, it turned out that this emotion was felt much more intensively through VR. In contrast, the emotion of disgust was more intensively felt when watching movie sections on a monitor.

We also verified whether there is a dependence between the valence and arousal values, where a negative dependence was shown. This means that the more negative the valence, the higher the arousal. Based on further analyses, this trend was generally confirmed, which shows that negative emotions were felt more intensively than positive ones. However, we further examined the data separately for each emotion, wherein the case of the emotion, joy, an opposite trend emerged. It turned out that the more the movie section was taken positively, the higher the intensity of the emotional feeling was. These conclusions were drawn on the basis of all data together, i.e. regardless of whether the measurements were obtained by the VR or LCD method.

We rejected hypotheses H1 and H2. Based on hypothesis H3, we verified whether there is a difference between the intensity of feeling negative and positive emotions. The results of our analysis show that there is a statistically significant difference, with negative emotions being felt more intensively than positive ones.

All participants during the implementation of the experiment were students of the Department of Informatics (with the focus of the Applied Informatics). We used Virtual reality for specific tasks (topics) during the teaching

process (in our case): for example, technical repair of computer (change of hardware components). However, due to lack of funding, it is not always possible to use VR for every student during the teaching process. We wanted to use experiments to verify how VR affects the emotional state of students. We found that depending on the input conditions, it is possible to achieve approximately the same results for the curved type of monitor as for the VR.

In further research, we can start from the obtained results of arousal values, which speak of the intensity of the emotion experienced. When using other sensors that will record different physiological characteristics (GSR, HR, Temperature, EEG) of the respondent while experiencing a given emotion, it is necessary to select movie sections focussing on sensations that evoke the most pronounced emotion, the ones that generate the greatest arousal. The reason is that the given emotion can affect the physiological manifestations only very slightly, and therefore the evoked stimuli must achieve the strongest possible emotion.

**Acknowledgement** This work was supported by the financial support of the project KEGA 036UKF-4/2019, Adaptation of the learning process using sensor networks and the Internet of Things.

**Funding** The authors received financial support for the research, authorship, and/or publication of this article. The financial support of the project was from project KEGA 036UKF-4/2019— "Adaptation of the learning process using sensor networks and the Internet of Things".

## Declarations

**Conflicts of interest** The authors declare that there is no conflict of interest.

### Ethical standards

The authors report that this paper involved human participants, their data or biological material. Authors declare that the study was approved by the appropriate institutional and/or national research ethics organisation KEGA (in English—Culture and Educational Grant Agency) and certify that the study was performed in accordance with the ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Alberdi A, Aztiria A, Basarab A (2016) Towards an automatic early stress recognition system for office environments based on multimodal measurements: a review. *J Biomed Inform* 59:49–75. <https://doi.org/10.1016/J.JBI.2015.11.007>
- Allcoat D, von Mühlenen A (2018) Learning in virtual reality: effects on performance, emotion and engagement. *Res Learn Technol* 26. <https://doi.org/10.25304/rlt.v26.2140>
- Bahreini K, Nadolski R, Westera W (2016) Towards multimodal emotion recognition in e-learning environments. *Interact Learn Environ* 24(3):590–605. <https://doi.org/10.1080/10494820.2014.908927>
- Bradley MM, Lang PJ (1994) Measuring emotion: the self-assessment manikin and the semantic differential. *J Behav Ther Exp Psychiatry* 25(1):49–59. [https://doi.org/10.1016/0005-7916\(94\)90063-9](https://doi.org/10.1016/0005-7916(94)90063-9)
- Chang SC, Hsu TC, Jong MSY (2020) Integration of the peer assessment approach with a virtual reality design system for learning earth science. *Comput Educ* 146:103758. <https://doi.org/10.1016/j.compedu.2019.103758>
- Cruz-Albarran IA, Benitez-Rangel JP, Osornio-Rios RA, Morales-Hernandez LA (2017) Human emotions detection based on a smart-thermal system of thermographic images. *Infrared Phys Technol* 81:250–261. <https://doi.org/10.1016/J.INFARED.2017.01.002>
- Demitriadou E, Stavroulia K, -, & Lanitis, A. (2020) Comparative evaluation of virtual and augmented reality for teaching mathematics in primary education. *Educ Inf Technol* 25(1):381–401. <https://doi.org/10.1007/s10639-019-09973-5>
- Diemer J, Alpers GW, Peperkorn HM et al (2015) The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Front Psychol* 6:1–9. <https://doi.org/10.3389/fpsyg.2015.00026>
- Ding N, Zhou W, Fung AYH (2018) Emotional effect of cinematic VR compared with traditional 2D film. *Telematics Inform* 35(6):1572–1579. <https://doi.org/10.1016/j.tele.2018.04.003>
- Ekman P, Friesen W (1978) *Facial Action Coding System: Investigator's Guide*. Consulting Psychologists Press, California
- Ekman, P., Friesen, W. ., & Hager, J. . (2002). *Facial action coding system*. ResearchNexus.
- Felnhöfer A, Kothgassner OD, Schmidt M, Heinzle AK, Beutl L, Hlavacs H, Kryspin-Exner I (2015) Is virtual reality emotionally arousing? Investigating five emotion inducing virtual park scenarios. *Int J Hum Comput Stud* 82:48–56. <https://doi.org/10.1016/j.ijhcs.2015.05.004>
- Francisti J, Balogh Z (2019a) An overview of solutions to the issue of exploring emotions using the internet of things. *Lect Notes Electric Eng* 574:59–67. [https://doi.org/10.1007/978-3-030-21507-1\\_9](https://doi.org/10.1007/978-3-030-21507-1_9)
- Francisti J, Balogh Z (2019b) Identification of emotional states and their potential. *Adv Intell Syst Comput* 924:687–696. [https://doi.org/10.1007/978-981-13-6861-5\\_58](https://doi.org/10.1007/978-981-13-6861-5_58)
- Garzotto F, Torelli E, Vona F, Aruanno B (2019). HoloLearn: Learning through mixed reality for people with cognitive disability. In Paper presented at the proceedings—2018 IEEE international conference on artificial intelligence and virtual reality, AIVR 2018, pp 189–190. doi:<https://doi.org/10.1109/AIVR.2018.00042>
- Gonçalves VP, Giancristofaro GT, Filho GPR, Johnson T, Carvalho V, Pessin G, Neris VP, d. A., & Ueyama, J. (2017) Assessing users' emotion at interaction time: a multimodal approach with multiple sensors. *Soft Comput* 21(18):5309–5323. <https://doi.org/10.1007/s00500-016-2115-0>
- Goshvarpour A, Goshvarpour A (2018) Poincaré's section analysis for PPG-based automatic emotion recognition. *Chaos Solitons Fractals* 114:400–407. <https://doi.org/10.1016/j.chaos.2018.07.035>
- Gross T (2015) Foreword. *Lect Notes Comput Sci*. <https://doi.org/10.1007/978-3-319-22701-6>
- Han K, Ku J, Kim K, Jeong Jang H, Park J, Kim JJ, Kim CH, Choi MH, Young Kim I, Kim SI (2009) Virtual reality prototype for measurement of expression characteristics in emotional situations. *Comput Biol Med* 39(2):173–179. <https://doi.org/10.1016/j.combiomed.2008.12.002>
- Hossain MS, Muhammad G (2019) Emotion recognition using secure edge and cloud computing. *Inf Sci* 504:589–601. <https://doi.org/10.1016/j.ins.2019.07.040>
- Imani M, Montazer GA (2019) A survey of emotion recognition methods with emphasis on E-Learning environments. *J Netw Comput Appl* 147(July):102423. <https://doi.org/10.1016/j.jnca.2019.102423>
- Isomursu M, Tähti M, Väinämö S, Kuutti K (2007) Experimental evaluation of five methods for collecting emotions in field settings with mobile applications. *Int J Hum Comput Stud* 65(4):404–418. <https://doi.org/10.1016/j.ijhcs.2006.11.007>
- Jang Y, Park E (2019) An adoption model for virtual reality games: The roles of presence and enjoyment. *Telemat Inform* 42(June):101239. <https://doi.org/10.1016/j.tele.2019.101239>
- Javaid M, Haleem A (2019) Virtual reality applications toward the medical field. *Clin Epidemiol Global Health* 8:600–605. <https://doi.org/10.1016/j.cegh.2019.12.010>
- Kaklauskas A, Zavadskas EK, Seniut M, Dzemyda G, Stankevicius V, Simkevicius C, Stankevicius T, Paliskiene R, Matuliuskaite A, Kildiene S, Bartkiene L, Ivanikovas S, Gribniak V (2011) Web-based biometric computer mouse advisory system to analyze a user's emotions and work productivity. *Eng Appl Artif Intell* 24(6):928–945. <https://doi.org/10.1016/J.ENGAPPAL.2011.04.006>
- Kerkeni L, Serrestou Y, Raoof K, Mbarki M, Mahjoub MA, Cleder C (2019) Automatic speech emotion recognition using an optimal combination of features based on EMD-TKEO. *Speech Commun* 114(September):22–35. <https://doi.org/10.1016/j.specom.2019.09.002>
- Leukhin A, Talanov M, Vallverdú J, Gafarov F (2018) Bio-plausible simulation of three monoamine systems to replicate emotional phenomena in a machine. *Biol Inspir Cognit Archit* 26(October):166–173. <https://doi.org/10.1016/j.bica.2018.10.007>
- Loureiro SMC, Guerreiro J, Ali F (2020) 20 years of research on virtual reality and augmented reality in tourism context: A text-mining approach. *Tour Manag* 77:104028. <https://doi.org/10.1016/j.tourman.2019.104028>
- Loureiro SMC, Guerreiro J, Eloy S, Langaro D, Panchapakesan P (2019) Understanding the use of virtual reality in marketing: a text mining-based review. *J Bus Res* 100:514–530. <https://doi.org/10.1016/j.jbusres.2018.10.055>
- Lövheim H (2012) A new three-dimensional model for emotions and monoamine neurotransmitters. *Med Hypotheses* 78(2):341–348. <https://doi.org/10.1016/j.mehy.2011.11.016>
- Magdin M, Benko L, Koprda Š (2019a) A case study of facial emotion classification using Affdex. *Sensors* 19(9):2140
- Magdin M, Kohútek M, Koprda Š, Balogh Z (2019b) EmoSens—the proposal of a system for recognition of emotion with SDK Affective and various sensors. *International Conference on Intelligent Computing*. Springer, Cham, pp 400–411
- Mahlke S, Minge M, and Thüring M (2006) Measuring multiple components of emotions in interactive contexts. In *Conference on human factors in computing systems—proceedings*, pp 1061–1066. <https://doi.org/10.1145/1125451.1125653>. April 2006
- Makki I, Alhalabi W, Adham RS (2019) Using emotion analysis to define human factors of virtual reality wearables. *Procedia Comput Sci* 163:154–164. <https://doi.org/10.1016/j.procs.2019.12.097>
- Marchewka A, Żurawski Ł, Jednoróg K, Grabowska A (2014) The Nencki Affective Picture System (NAPS): introduction to a novel,

- standardized, wide-range, high-quality, realistic picture database. *Behav Res Methods* 46(2):596–610. <https://doi.org/10.3758/s13428-013-0379-1>
- Marín-Morales J, Higuera-Trujillo JL, Greco A, Guixeres J, Llinares C, Scilingo EP, Alcañiz M, Valenza G (2018) Affective computing in virtual reality: Emotion recognition from brain and heartbeat dynamics using wearable sensors. *Sci Rep* 8(1):1–15
- McDuff D, Mahmoud A, Mavadati M, Amr M, Turcot J and Kalioubi RE (2016) AFFDEX SDK: a cross-platform real-time multi-face expression recognition toolkit. In Proceedings of the 2016 CHI conference extended abstracts on human factors in computing systems, pp 3723–3726
- Merta T (2019) ScienceDirect Emotion monitoring system for drivers. *Emotion Syst Drivers* 52(8):200–205. <https://doi.org/10.1016/j.ifacol.2019.08.071>
- Michelini Y, Acuña I, Guzmán JI, Godoy JC (2019) LATEMO-E: a film database to elicit discrete emotions and evaluate emotional dimensions in Latin-Americans. *Trends Psychol* 27(2):473–490
- Mohamed Elias Z, Batumalai UM, Azmi ANH (2019) Virtual reality games on accommodation and convergence. *Appl Ergon* 81(June):102879. <https://doi.org/10.1016/j.apergo.2019.102879>
- Molnár G, Szűts Z, Biró K (2018) Use of Augmented Reality in Learning *Acta Polytechnica Hungarica* 15(5):209–222
- Munk M, Drlík M (2011) Impact of different pre-processing tasks on effective identification of users' behavioral patterns in web-based educational system. Paper Present *Procedia Comput Sci* 4:1640–1649. <https://doi.org/10.1016/j.procs.2011.04.177>
- Nam J, Chung H, and Lee H (2019) A new terrain in HCI: Emotion Recognition Interface using biometric data for an immersive VR experience. *arXiv preprint*.
- Niu Y, Wang D, Wang Z, Sun F, Yue K, Zheng N (2019) User experience evaluation in Virtual Reality based on subjective feelings and physiological signals. *J Imag Sci Technol* 13:60413
- Pan Z, Cheok AD, Yang H, Zhu J, Shi J (2006) Virtual reality and mixed reality for virtual learning environments. *Comput Graph Pergamon* 30(1):20–28. <https://doi.org/10.1016/j.cag.2005.10.004>
- Picard RW (2009) Future affective technology for autism and emotion communication. *Philos Trans R Soc B Biol Sci* 364(1535):3575–3584. <https://doi.org/10.1098/rstb.2009.0143>
- Pizzi G, Scarpi D, Pichierri M, Vannucci V (2019) Virtual reality, real reactions?: Comparing consumers' perceptions and shopping orientation across physical and virtual-reality retail stores. *Comput Hum Behav* 96:1–12. <https://doi.org/10.1016/j.chb.2019.02.008>
- Plutchik R (1980) A general psychoevolutionary theory of emotion. In: Plutchik R, Kellerman H (eds) *Emotion: theory, research and experience, theories of emotion*, vol 1. Academic Press, New York, pp 3–33
- Poria S, Cambria E, Bajpai R, Hussain A (2017) A review of affective computing: From unimodal analysis to multimodal fusion. *Inf Fus* 37:98–125. <https://doi.org/10.1016/j.inffus.2017.02.003>
- Quazi MT, Mukhopadhyay SC (2011) Continuous monitoring of physiological parameters using smart sensors. *Fifth Int Conf Sens Technol* 2011:464–469. <https://doi.org/10.1109/ICSensT.2011.6137022>
- Russell JA (1979) Affective space is bipolar. *J Pers Soc Psychol* 37(3):345–356. <https://doi.org/10.1037/00223514.37.3.345>
- Seo YS, Huh JH (2019) Automatic emotion-based music classification for supporting intelligent IoT applications. *Electronics* 8(2):164
- Scherer KR (2005) What are emotions? And how can they be measured? *Soc Sci Inf* 44(4):695–729. <https://doi.org/10.1177/0539018405058216>
- Shalini TB, Vanitha L (2013) Emotion detection in human beings using ECG signals. *Int J Eng Trends Technol* 4(May):1337–1342
- Sharma R, Pachori RB, Sircar P (2020) Automated emotion recognition based on higher order statistics and deep learning algorithm. *Biomed Signal Process Control* 58:101867. <https://doi.org/10.1016/j.bspc.2020.101867>
- Tonguç G, Ozaydın Ozkara B (2020) Automatic recognition of student emotions from facial expressions during a lecture. *Comput Educ* 148:103797. <https://doi.org/10.1016/j.compedu.2019.103797>
- Tsonos D, and Kouroupetroglou G (2008) A methodology for the extraction of the reader's emotional state triggered from text typography. In *Tools in artificial intelligence*. IntechOpen
- Vanderlind WM, Millgram Y, Baskin-Sommers AR, Clark MS, Joormann J (2020) Understanding positive emotion deficits in depression: From emotion preferences to emotion regulation. *Clin Psychol Rev* 76:101826. <https://doi.org/10.1016/j.cpr.2020.101826>
- Vera L, Herrera G, Vived E (2005) Virtual reality school for children with learning difficulties. Paper Present *ACM Int Conf Proc Ser* 265:338–341. <https://doi.org/10.1145/1178477.1178541>
- Visch VT, Tan ES, Molenaar D (2010) The emotional and cognitive effect of immersion in film viewing. *Cogn Emot* 24:1439–1445. <https://doi.org/10.1080/02699930903498186>
- Waller BM, Julle-Daniere E, Micheletta J (2020) Measuring the evolution of facial 'expression' using multi-species FACS. *Neurosci Biobehav Rev* 113:1–11. <https://doi.org/10.1016/j.neubiorev.2020.02.031>
- Wang X, Tang L, Kim E (2019) More than words: Do emotional content and linguistic style matching matter on restaurant review helpfulness? *Int J Hospit Manag* 77:438–447. <https://doi.org/10.1016/j.ijhm.2018.08.007>
- Yang X, Lin L, Cheng PY, Yang X, Ren Y (2019) Which EEG feedback works better for creativity performance in immersive virtual reality: The reminder or encouraging feedback? *Comput Hum Behav* 99(June):345–351. <https://doi.org/10.1016/j.chb.2019.06.002>
- Zhang J, Yin Z, Chen P, Nichele S (2020) Emotion recognition using multi-modal data and machine learning techniques: A tutorial and review. *Inf Fus* 59:103–126. <https://doi.org/10.1016/j.inffus.2020.01.011>
- Zinchenko YP, Khoroshikh PP, Sergievich AA, Smirnov AS, Tummyalis AV, Kovalev AI, Gutnikov SA, Golokhvast KS (2020) Virtual reality is more efficient in learning human heart anatomy especially for subjects with low baseline knowledge. *New Ideas Psychol* 59(February):100786. <https://doi.org/10.1016/j.newideapsych.2020.100786>

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.