UNIVERZITA KONŠTANTÍNA FILOZOFA V NITRE
FAKULTA PRÍRODNÝCH VIED

ROZPOZNÁVANIE EMOCIONÁLNEHO STAVU POUŽÍVATEĽA POMOCOU INTELIGENTNÝCH RIADIACICH SYSTÉMOV

Diplomová práca

Študijný program: Aplikovaná informatika (magisterský II. St., denná forma)
Študijný odbor: AI15m - Aplikovaná informatika
Školiace pracovisko: Katedra Matematiky
Školiteľ: Mgr. Martin Magdin, PhD

Nitra 2018 **Bc. Michal Kohútek**

Abstrakt

Abstract

Abstract goes here

Dedication

To mum and dad

Declaration

I declare that..

Acknowledgements

I want to thank...

Contents

Introduction

”Smiles are probably the most underrated
facial expressions, much more complicated
than most people realize. There are dozens
of smiles, each differing in appearance
and in the message expressed.”
- Paul Ekman

Emotions are at the core of the human experience, albeit very hard to define, recognize and name, even in yourself. They are, by definition different from person to person, in diverse cultures and upbringings. Our perception of emotions and their classification has evolved in recent years. Various authors has tried to divide our emotional states into basic categories such as Ekman’s Anger, Fear, Disgust, Happiness, Sadness and Surprise. However, recent work by psychologists and historians alike show, that a more complex look at emotions might be needed.

Emotional recognition changes with time. In 12th century, bards looked at yawning not as a sign of boredom or tiredness, but as a sign of a hidden and deep love. Early Christians recognized an emotion called ”accidie”, a lethargy and despair brought about by flying demons. Boredom, as such was first really felt by the Victorians as a response to the new ideas of leisure time and self-improvement. Among the psychologist, there is a standing question whether some cultures feel some hard to define emotions more strongly, because they bothered to name them as separate kinds. For example the Russian ”toska”, a longing with nothing to long for, as coined by Vladimir Nabokov. Recent developments of cognitive science tell us, that emotions are not just simple reflexes, but inherently complex and elastic systems of response towards both the biologies that we’ve inherited and the cultures, that we live in now. They are not just simple chemistry, but a cognitive phenomena, not shaped only by our body functions, but also our thought process, concepts and language. The neuroscientist Lisa Feldman Barrett studies this dynamic relationship between words and emotions. She argues, that when a person learns a new word for an emotion, they also learn to feel and recognize it (Barrett, 2007). There is a historicity to emotions, they have changed in history, often times very dramatically, in response to new cultural expectations, religious beliefs, new ideas about gender, age, ethnicity, economical and political ideologies.

There is a push to increase our emotional intelligence. Emotions are so powerful, that in past, they were sometimes thought to be a cause of illness. In 17th century, there was a student attending the Swiss university in Basel. He came afflicted with fever, heart palpitations, skin sores on his body and was close to dying. When they sent him back home to die, he started getting better and by the time, he returned to his hometown, he almost entirely recovered. In 1688, Johannes Hoffer, medical doctor, learnt of this case and many like it and coined the term for a severe homesickness as ”nostalgia”. Last confirmed death by nostalgia was an American soldier fighting during the First World War in France. In early 20th century, this feeling has morphed more into a longing for lost time, instead of homesickness and downgraded in severity. Nowadays, our culture celebrates happiness, as it is said to make us a better workers, parents and partners. However, in 16th century, this position was filled by sadness, as is evident by self-help books from that period, which tried to encourage sadness in readers by giving them lists of reasons to be disappointed.

In order to study both basic emotions and their mixtures in the complex variety, we need appropriate detection and classification tools and techniques. Over the years, many psychologists and computer scientist collaborated to create a set of markers and techniques recognizing them. We will discuss several of them and compare their usefulness, reliability and practicality. These methods most often try to detect basic emotions such as the seven basic emotions (Ekman, 1992) or a 12-point Affect Circumplex model (Yik et al., 2011).

Figure 1: Six primary emotions

Affective computing has great many possible applications. It can allow for a better education process, providing a large amount of valid feedback to the tutors and students (Asteriadis et al., 2008). It can help broaden the knowledge in the study of psychology and psychiatry, improve the efficiency of counseling. Since emotions recognition and empathy is one of the traits of inter-personal communication, properly analyzing emotions can help in developing communicative technologies for use by people with autism. Growing interest in affective computing is shown by many industries. Automobile manufacturers are partnering with companies like Affectiva and integrating their solutions in their high-end cars (*Affectiva Automotive AI*). Other venue for emotion recognition is targeted marketing. In past, online e-commerce sites would often use "heat maps" to see, whether the potential customer can find products on the website, and whether the paid advertisement holds their attention. If the same system added emotion recognition, online marketers would gain greater insight into minds of customers. Another future purpose for affective computing may be found in robotic pets and social robots, especially those working in health care industry. Emotional awareness could allow robotic nurses to better judge users' and patient's emotional states and needs (Yonck, 2017). It is clear, that affective computing will grow to become an important aspect of computer science theory.

1. Emotion classification

Emotion classification is a contested issue in emotion research and in affective science. The two fundamental viewpoints of affective scientists' approach are:

1. Emotions as discrete and fundamentally different constructs

2. Emotions as fluid and characterized on a dimensional basis in groupings

Various categorizations of emotions also vary in description how emotions relate to each other.

1.1. Discrete models of emotions

Discrete emotion theory claims that there is a small number of core affects. This number can vary depending on the proponent, for example Silvan Tomkins considered nine basic emotions. Six, that came evolutionarily earlier, interest-excitement, enjoyment-joy, surprise-startle, distress-anguish, anger-rage, fear-terror, once that evolved later, shame-humiliation and finally disgust and dissmell, which he later took back. In the paired affects, the first of the pair is the mild manifestation and the second the more intense. (Tomkins, 1962), (Tomkins, 1963). This model is somewhat controversial nowadays among affective theorists, especially over Tomkins' firm insistence that there were nine and only nine, biologically based affects. He also argued, that these affects are quite discreet (in contrast to the more muddled and complex emotions) and that they shared a common biological heritage with what Darwin called emotions in animals (Darwin et al., 1998). They also differ from Freudian drives in lacking an object.

Similarly, Carroll Izard delineated 12 discrete emotions: interest, joy, surprise, sadness, anger, disgust, contempt, self-hostility, fear, shame, shyness and guild. He measured these via his Differential Emotion Scale (Boyle et al., 2015). Among other contributors to this theory, such as John Watson, Edwin Newman and Ross Buck, Paul Ekman performed a series of cross-cultural studies with Carrol Izard and reported that there are at least six emotions, that people across the world produce and are able to recognize. This was further evidenced when researchers approached the people of New Guinea with no previous exposure to Westerners nor their culture. When they showed them pictures of people expressing six core emotions, subjects could in fact point out the different emotions and distinguish between them (Ekman et al., 1971).

1.2. Dimensional models of emotions

There are various theoretical and practical reasons for which some researchers define emotions according to one or more dimensions. Dimensional models are an attempt to conceptualize human emotions by defining where they lie in two or three dimensions. Most incorporate valence and arousal or intensity dimensions. In contrast to theories of basic emotion, which propose that different emotions arise from separate neural systems, dimensional models suggest that a common and interconnected neurophysiological system is responsible for all affective states. (Posner et al., 2005)

In 1897, Wilhelm Max Wundt proposed, that emotions can be described by three dimensions: ”pleasurable versus unpleasurable”, ”arousing versus subduing” and ”strain versus relaxation” (Wundt, 2017). Later, Harold Schlosberg named three dimension, ”pleasantness–unpleasantness”, ”attention–rejection” and ”level of activation” (Schlosberg, 1954).

Another model, called the circumplex model, was developed by James Russell. This model suggest that emotions are distributed in a two-dimensional circular space, containing both arousal and valence dimensions. The vertical axis represents arousal, horizontal axis valence and the center of the circle represents a neutral valence and medium level of arousal.(Russell, 1980).

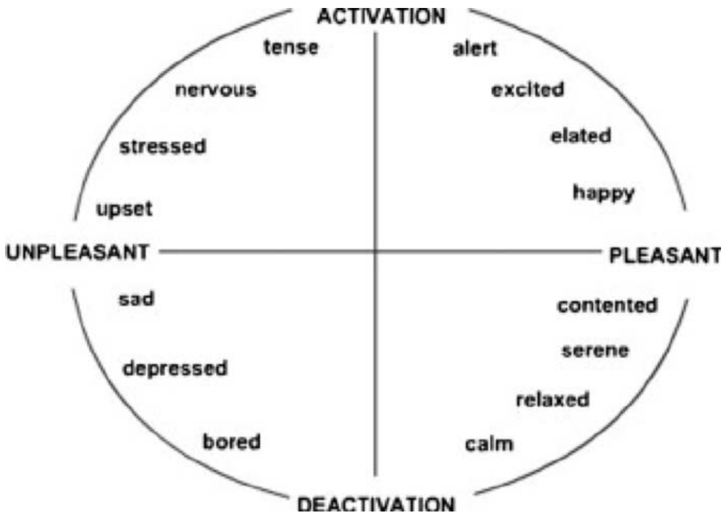
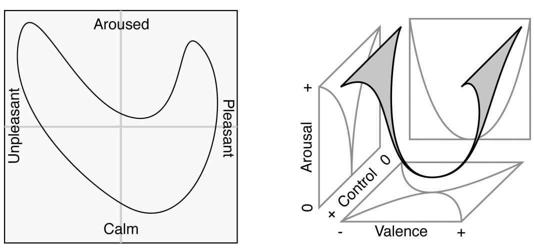


Figure 1.1: A. Russell’s (1980) circumplex model of affect

This model was later modified by Russel and Lisa Feldman Barret, which they described as representative of core affect, which are the most elementary feelings that need not be directed toward anything. Different prototypical emotional episodes, or clear emotions that are evoked or directed by specific objects, can be plotted on the circumplex, according to their levels of arousal and pleasure (Russell et al., 1999).

Another model of emotion appeared in 1992. This two-dimensional model consists of vectors, that point in two directions. The vector model assumes that there is always an underlying arousal dimension and that valence determines the direction in which a particular emotion lies.



(a) 2D representation (b) 3D representation

Figure 1.2: Vector model of emotions

The positive activation – negative activation (PANA) was originally created in 1985 by David Watson and Auke Tellegen. It suggests that positive and negative affect are two separate systems. Like in the vector model, states of higher arousal tend to be defined by their valence and states of lower arousal tend to be more neutral in term of valence. In the PANA model, the vertical axis represents low to high positive affect and the horizontal axis represents low to high negative affect. The dimensions of valence and arousal lay at a 45-degree rotation over these axes (Watson et al., 1985).

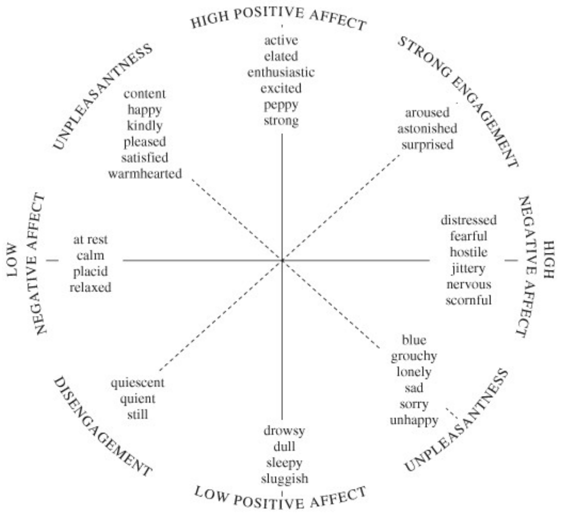


Figure 1.3: Consensual (PANA) model of emotion

In 1980, Robert Plutchik constructed a wheel of emotions. This model is a hybrid of both basic-complex categories and dimensional theories. Emotions are arranged in concentric circles, with the more basic emotions on the inner circles, while the outer circles are occupied by complex emotions. Notably, outer circles are also formed by blending the inner circle emotions. Plutchik suggested 8 primary contrasting pairs of emotions. Joy/sadness, anger/fear, trust/disgust and surprise/anticipation. Like colors, primary emotions can be expressed at different intensities and can mix with one another to form different emotions (Plutchik, 1988).

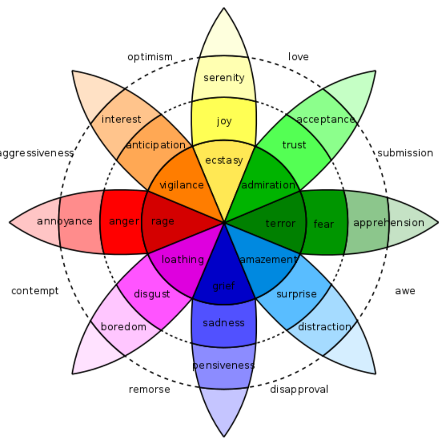


Figure 1.4: Plutchik's wheel of emotions

The PAD emotional state model was developed by Albert Mehrabian and James A. Russel. It uses three numerical dimensions, **P**leasure, **A**rousal, and **D**ominance to represent all emotions (Mehrabian, 1980). Initially it's use was in a theory of environmental psychology, the core idea being that physical environments influence people through their emotional impact (Mehrabian et al., 1974). Subsequently it was used by Peter Lang to propose a physiological theory of emotion (J. Lang et al., 1990). Furthermore, it was also used by Russel to develop a theory of emotional episodes (Russell, 2003). The Pleasure-Displeasure Scale measures how pleasant an emotion may be. Anger and fear are, for instance, unpleasant emotions and thus score high on the displeasure. Contrarily, joy is a pleasant emotion. The Arousal-Nonarousal Scale measures the intensity of emotion. For instance while both anger and rage are unpleasant emotions, rage is much more intense than anger. Boredom, while also an unpleasant state, has a low arousal volume. Lastly, the Dominance-Submissiveness Scale shows the dominant nature of the emotion. While both fear and anger are unpleasant emotions, anger is dominant, but fear is a submissive emotion (Mehrabian, 1980). A more

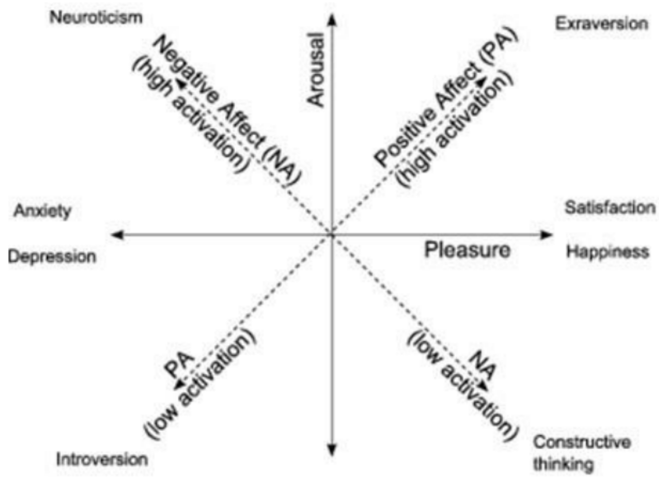


Figure 1.5: PAD emotional state model

abbreviated version of the PAD model has also been used in organizational studies where the emotions towards specific entities or products are measured. It uses just 4 values for each dimension, providing only 64 values for emotions (Ashkanasy et al., 2008).

An example of three-dimensional models, the Lövheim cube of emotion was presented where the signal substances (dopamine, noradrenaline and serotonin) form the axes of a coordinate system, and the eight basic emotions according to Tomkins are placed in the eight corners. As shown on the figure below, anger is produced by the combination of low serotonin, high dopamine and high noradrenaline. Conversely joy is a product of high serotonin, high dopamine and low noradrenaline. Since none of the axis is identical to valence [1], the cube seems somewhat rotated when compared to other models (Lövheim, 2012).

Most recently Cowen and Kelter, researchers from University of California, Berkeley introduced a statistically derived taxonomy of emotion. (Cowen et al., 2017)

[1] pleasantness

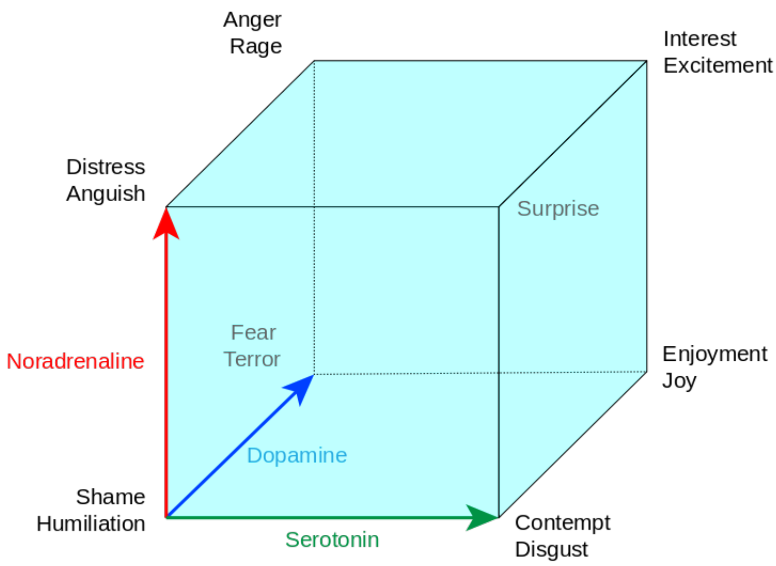


Figure 1.6: Lövheim cube of emotion

> Across self-report methods, we find that the [2185] videos [selected and shown to volunteer subjects] reliably elicit 27 distinct varieties of reported emotional experience. Further analyses revealed that categorical labels such as amusement better capture reports of subjective experience than commonly measured affective dimensions (e.g., valence and arousal). Although reported emotional experiences are represented within a semantic space best captured by categorical labels, the boundaries between categories of emotion are fuzzy rather than discrete. By analyzing the distribution of reported emotional states we uncover gradients of emotion—from anxiety to fear to horror to disgust, calmness to aesthetic appreciation to awe, and others—that correspond to smooth variation in affective dimensions such as valence and dominance. Reported emotional states occupy a complex, high-dimensional categorical space

More dimensional models of emotion have been developed, though there are just a few that remain as the dominant models currently accepted by most (Rubin et al., 2009). There have been observed great cultural differences in the way in which emotions are valued, expressed and regulated. The social norms for emotions, like the frequency with or circumstances in which they are expressed also vary drastically in diverse cultures. An important piece of evidence that disputes the universality of emotions is language. Emotions such as the schadenfreude [2] in German and saudade[3] in Portuguese are commonly expressed in emotions in their respective languages, but lack an English equivalent. Thus it is reasonable in our research to scale back on the complex, culturally influenced emotions and focus on the more primal, basic emotions, that may be more quantifiable by studying the physiological markers and responses in subjects.

[2]The experience of pleasure, joy, or self-satisfaction that comes from learning of or witnessing the troubles, failures, or humiliation of another.

[3]Deep emotional state of nostalgic or profound melancholic longing for an absent something or someone that one loves.

1.3. Physiological responses of emotions

When dealing with physiological effects of emotion, we come across a few prevalent theories. The James-Lange theory is one of the earliest theories of emotion within modern psychology. It was developed independently by William James and Carl Lange in 19th century. The basic premise [4] is that physiological arousal instigates the experience of emotion, i.e., the arousal precedes and causes the emotion (Cannon, 1927). According to the Cannon-Bard theory of emotions, the emotion is accompanied by physiological arousal.

In order for us to be able to detect and recognize emotions from the data samples, we need to asses which physiological effects of emotions we want to capture. Therefore, we need to divide our standard basic emotions depending on the combination of kinds and intensity of effects they have on human body. There is a difficulty with dealing with physiological markers of emotions in that not all emotions do perceivably alter the physiology of the subjects. That is the reason, why we need to supplement this data with the facial expressions or, as we have tried to confirm, ocular movements.

Anger is one of the emotions with more pronounced physiological signs. It comes along with faster and deeper breathing (Philippot et al., 2010), increased heart rate, blood pressure, perspiration and tensing of muscles. Fear, on the other hand, shares most of the surface physiological marks with anger, such as the accelerating breathing rate, heart rate and muscle tension, but facial expression and body language are vastly different. Even deeper markers, like the cortisol levels differ after subjecting a person to these stressors (Moons et al., 2010). Joy, sadness, disgust and surprise are similar to each other in their difficulty to asses using pure physiological markers and therefore requiring additional information, usually in form of facial expression.

1.4. Current approaches in detecting and classifying emotions

As we've described motion recognition is an important object of studies in today's psychology, with many potential uses and applications. Correctly assessing and recognizing subject's emotion can lead to better understanding it's motivation and inner working. Data gained through methods described below can be used to assess the effectiveness of marketing, comprehensibility of lectures, usability of user interfaces, measure of impact of psychological therapy, etc. Previous implementations of emotion recognition technology often have had a multitude of disadvantages, which prohibited it's daily and widespread usage. Therefore, we have set upon creating a solution, that is modular, reasonable to wear for prolonged durations of time and still maintains a degree of reliability in captured data. We can do so by using an array of data resources, that complement each other, diminishing their disadvantages and reinforcing confidence.

One particularly rich resource for data on human emotions is the brain. In 2015 a group of researchers in India published a paper on a system using EEG signals as input,

[4] Which is also the point of most criticism

Independent Component Analysis [5], Kernel Destiny Estimation [6] and an Artificial Neural Network to transform the inputs into meaningful outputs. They observed better results for clustering of EEG and ECG data stream (Lahane et al., 2015). Similar approach was taken by researchers in China. Main difference in their approach is that they first applied EMD [7] strategy to split EEG signals into a series of intrinsic mode functions, which were then fed as sample entropies and as feature vectors into SVN[8] classifier for testing and training. With this approach, they claim to have reached accuracy levels of 94.98% for binary-class tasks and 93.20% for the multi-class task on DEAP database (Zhang et al., 2016). Also in 2016, researchers from the Duke University, North Carolina demonstrated an emotional recognition technique using functional MRI with results, that brain-based models may, in future, allow us deeper understanding and assessing emotional status in clinical settings, particularly in individuals incapable of providing self-report of their own emotional experience (Kragel et al., 2016). These techniques are, however, impossible to replicate on greater scale and in the context of a classroom, or other commonplace environment.

Another indicator of the subject's emotional state is heartbeat. Research published in May 2013 in International Journal of Engineering Trends and Technology shown compelling data gathering by means of ECG[9] and shown differences in ECG signal in subjects in chosen emotional states. [10] (Shalini et al., 2013). More complex approach was taken by researchers at University of Calabria in collaboration with Washington State University. Instead of applying a single or a few sensors to study physiological states, they used a whole BSN[11] Such networks can include accelerometers, gyroscopes, pressure sensors for body movements and applied forces, skin/chest electrodes (for electrocardiogram (ECG), electromyogram (EMG), galvanic skin response (GSR), and electrical impedance plethysmography (EIP)), (PPG) sensors, microphones (for voice, ambient, and heart sounds), scalp-placed electrodes for electroencephalogram (EEG) (Gravina et al., 2016). Albeit they were not focusing primarily on emotion recognition, this survey shows consequential advancement in the state-of-the-art body data collection, especially in the data fusion techniques. Similar approach using a variety of sensors and data sources was also taken by Mosciano et al., 2017 to reasonably accurately classify the two dimensions of affect both normal and simulated critical working conditions.

An interesting approach to emotion detection and classification is based upon speech patterns and variations and are better suited for emotions, which are otherwise hard to physiologically measure, such as sadness and joy. One study shows, that speech signal and feature distances of letters and words vary depending on the mood and emotional state of the subject. The study was done upon the sample size of 30 people, however, researchers point out, that for getting better accuracy, one should consider the data collected from one person rather than considering the data from a group of people (Davletcharova et al., 2015).

[5] A statistical procedure used for splitting up a set of mixed signals into its sources

[6] A method used for feature extraction of signals by computing density estimate using kernel-smoothing method

[7] Empirical Mode Decomposition

[8] Support Vector Machine

[9] Electrocardiography

[10] This study was limited to joy, sadness, fear and anger

[11] Body Sensor Network, specialized Wireless Sensor Network applied to the whole human body.

In terms of evaluating arousal, respiration-based emotion recognition shows promise. A study in China showed, that using respiration data to evaluate valence and arousal levels of Russel theory, they reached classification accuracy of valence and arousal at 73.06% and 80.78%, respectively (Zhang et al., 2017). When compared to other studies using ECG or EEG data, the accuracy of valence levels are not exceptional, however the classification accuracy of arousal better than most other approaches.

A study on thermal behavior of anger, disgust, fear, joy and sadness was carried out in 2016. When an emotion occurs a change in facial temperature appears due to the blood flow that the body emits through blood vessels in the subcutaneous area (Ioannou et al., 2014). For example, research focused on the emotion of joy, in other words, when a subject is smiling, it has been found that the temperature of the nose and forehead decreases during this event (Salazar-Lopez et al., 2015). Biomedical thermal images of the facial expressions of 44 subjects were captured experiencing the five studied emotions, with results of this test at 89.9% success rate (Cruz-Albarran et al., 2017).

Another novel approach was experimented with by researchers from American University of Sharjah, United Arab Emirates. They have created a software touch keyboard, that was installed on Android smartphones, which then have been collecting sensor data while users were typing on the keyboard. As they have been typing, he or she were prompted to indicate their current emotional state, which then tagged the sensor data collected for the particular user. Afterwards the data was classified by multiple machine learning algorithms to find the best classification method. Based on ROC[12] and Precision-Recall curves, it was concluded that both J48 and Multi-response linear regression performed well. This system demonstrates that it is possible to enable emotion recognition on mobile phones using built-in sensors (Zualkernan et al., 2017). More research is, nevertheless, needed to precisely ascertain the accuracy and applicability of such solution in common practice.

One approach seems to be very promising, especially due to it's implementation in commercial solutions and services. That approach is based on facial expression evaluation and it is used by a wide variety of software vendors and research institutions. One such research has used Microsoft Kinect for 3D face modeling with a goal to computationally recognize facial expressions of seven basic emotional states: neutral, joy, surprise, anger, sadness, fear and disgust. The subjects of the experiment were six men aged 26-50 years, told to mimic expressions shown on the screen. Researchers used nearest neighbor classifier (3-NN) and two-layer neural network classifier (MLP) with 7 neurons in the hidden layer with output accuracy rate of 96% for random division of data and 73% for "natural" division of data (Tarnowski et al., 2017).

Last but not least, one of the more recent[13] approaches taken in affective computing is gaze and pupil tracking. This builds on classic theories (H. Hess, 1975 ; Beatty et al., 2012) and tries to implement eye-tracking techniques to investigate emotional state and behavior of subjects. They were successful in distinguishing between emotions, albeit having to limit

[12]Receiver Operating Characteristic
[13]And less explored.

them from the core seven to four. Accuracy of their recognition ranged in the interval from 80 to 90% (Maskeliunas et al., 2016)

There are two questions, that remains unanswered. Is our current method of selecting impulses the correct one? Are we rating approaches the right way? In research mentioned above, there are mainly two types of referential data. One is based on self-identifying emotions, the other on using previously assorted categories of pictures or impulses. As we've mentioned before, self-identifying emotions is a hard task for most people and it is the task, that we aim to solve. Also, is the methodology we use, to personally evaluate emotions the correct one? We usually refer to the previous evaluations, that trace back to psychologist and we take their emotion assessments at face value. It might be wise, that we should ignore previous results and rather try to classify and categorize emotions from ground up and later compare them with historical data. Finally, most of the researches delving into face expression-based emotion recognition declare their statistics based upon the success rate of recognizing simple expressions, not emotions themselves.

1.4.1 Emotion recognition used in commercial environment

With plethora of possible applications and use-cases, from marketing, automotive industry, health care, education and psychology research, it is no wonder that a lot of industries are pushing towards better and faster classification of emotions. Consequently the number of dependable SDK[14] and API[15] grows rapidly. A few of them also merge and/or work together towards better integration.

Emotient was one such company. Their FACET SDK allows application to track and analyze the emotional responses of users in real-time, detecting and tracking expressions of primary emotion, as well as overall positive, negative and neutral sentiments and blended composites of two or more emotions. In 2013 they partnered with iMotions to provide a facial expression recognition, EEG and GSR analysis platform to Procter and Gamble, the United States Air Force and Yale University. As part of the agreement between the two companies, Emotient's industry-leading emotion recognition solution was fully integrated into iMotions Attention Tool, an eye tracking and biosensor software platform for research and usability. In 2016 Emotient got acquired by Apple and their website is no longer online.

Next in consideration was Eyeris EmoVu. It's a closed-source, multi-platform SDK and API, that supports all three major desktop operating systems: Microsoft Windows, macOS and GNU/Linux. It has language hooks for C++, Java, VB.NET, Python, Objective-C, Node.js and more. Their pricing starts with a free license, that covers the analysis of 500 frames per month. It's declared features are

- Emotion Recognition
- Gender Recognition,

[14]Software development kit - A set of software development tools that allows the creation of application for a certain software package, framework, hardware platform, etc.

[15]Application programming interface - A set of functions and procedures that allow the creation of applications which access the features or data of an operating system, application, or other service.

[16]Although not in numerical values, but in labels such as Child, Young Adult, Adult and Senior

- Age Recognition[16],
- Face Recognition,
- Facial Tracker,
- Engagement and Mood metrics.

EmoVu's main drawback is lack of clear information about the SDK availability, since they only mention it on their website in unclear terms and the pricing only mentions API.

An interesting approach to confronting racial bias in face recognition algorithms is shown by Kairos. As several studies shown, there is a great difference between the error rates in human races. The extent of these biases are reflected in an error rate of 0.8 percent for light-skinned men, and as high as 34.7% for dark-skinned women. Since this problem can have very far reaching implications, it's important to try to mitigate it. The solution advocated by Kairos is to:

1. Trust the process - as it's a new technology with room to grow,
2. Improve the data sets - gathering, training, and testing data from a population that is truly representative of global diversity
3. Seek constant feedback - trying to see what the users and communities see.

(*Face Off: Confronting Bias in Face Recognition AI*)

Microsoft, the epitome of a immense corporation, has entered this technological race with the launch of their Project Oxford. It's a collection of artificial intelligence APIs with focus on computer vision, speech and language analysis. It got it's own fair share of controversy with it's mediocre launch, but continued to improve steadily. Project Oxford's tools enable face detection and emotion detection [17] Unfortunately, the API works only with photos. The pricing start with a free plan, that covers 30,000 transactions per month.

InSight SDK is facial recognition C++ toolbox developed by Sightcorp in collaboration with the University of Amsterdam. It boasts a large set of features, apart from face detection and emotion recognition, it uncovers gaze estimation, head pose estimation and motion tracking features. InSight does not provide academic trials, therefore we could not evaluate their performance.

Another company that partnered with iMotions is Affectiva. The deal came with Affectiva making their Affdex facial coding and emotion analytics software available on the iMotions Biometric Research Platform. With this joint solution human behavior researchers can combine facial coding and emotion analytics with eye tracking, brainwave measurement (EEG), as well as physiological sensors (GSR, ECG, EMG) on top of traditional surveys and questionnaires, which are also fully integrated. The main Affectiva's achievement was the amassment of the world's largest emotion data repository. By 2018, they analyzed over 6 million faces from 87 countries The offer both API [18] and SDK Which allows applications to run locally and without needing Internet connection. for the major platform like Android,

[17] It uses a model of core seven emotions plus neutral.
[18] Which can be used online on a thin client.

iOS, GNU/Linux, Windows and macOS. They have built their massive data set of facial expressions by analyzing millions of face videos, of people engaged in various activities, such as watching media content[19], driving cars, people in conversational interactions and animated gifs[20]. To get this impressive amount of data, Affectiva partnered with their market research partners such as Millward Brown, Unruly, Lightspeed, Added Value, Voxpopme and LRW and also with partners in the automotive industry, robotics and Human Resources space. Affectiva's Affdex SDK was tested on 10 000 images to verify the generalizability of algorithms. Those images cover different lightning conditions, both genders, various poses of participants, etc. Via the SDK, application can access and read multiple metrics. Those include 7 emotion metrics, 20 facial expression metrics, 13 emojis and 4 appearance metrics. Apart from the 7 emotional states, the SDK surfaces also the Engagement[21] and Valence [22] metric. The SDK offers four main features.

1. Face and facial landmark detection,
2. Face texture feature extraction,
3. Facial action classification,
4. Emotion expression modelling.

(McDuff et al., 2016) The SDK are offered under various licensing deals, with a free Commercial Evaluation License valid for 60 days, or Student Evaluation License valid for 6 months. The API is paid by the processing time, with pricing starting at $1/minute of video processed. Affectiva's mapping of expressions onto emotions builds on EMFACS mapping, developed by Friesen and Ekman. EMFACS is a variant of FACS system, that considers only emotion-related facial actions. (Ekman et al., 1997) The table 1.1 shows the relationship between the facial expressions and the emotions predictors as used in the Affdex SDK.

Emotion	Increase Likelihood	Decrease Likelihood
Joy	Smile	Brow Raise, Brow Furrow
Anger	Brow furrow, Lid Tighten, Eye Widen, Chin Raise, Mouth Open, Lip Suck	Inner Brow Raise, Brow Raise, Smile
Disgust	Nose Wrinkle, Upper Lip Raise	Lip Suck, Smile
Surprise	Inner Brow Raise, Brow Raise Eye Widen, Jaw Drop	Brow Furrow
Fear	Inner Brow Raise, Brow Furrow Eye Widen, Lip Stretch	Brow Raise, Lip Corner Depressor Jaw Drop Smile
Sadness	Inner Brow Raise, Brow Furrow Lip Corner Depressor	Brow Raise, Eye Widen, Lip Press Mouth Open, Lip Suck, Smile
Contempt	Brow Furrow, Smirk	Smile

Table 1.1: Relationship between the facial expressions and the emotions predictors

[19]i.e, ads, movie trailers, television shows and online viral campaigns

[20]Via the partnership with giphy

[21]A measure of facial muscle activation that illustrates the subject's expressiveness. The range of values is from 0 to 100.

[22]A measure of the positive or negative nature of the recorded person's experience. The range of values is from -100 to 100.

Increase Positive Valence Likelihood	Increase Negative Valence Likelihood
Smile Cheek Raise	Inner Brow Raise Brow Furrow Nose Wrinkle Upper Lip Raise Lip Corner Depressor Chin Raise Lip Press Lip Suck

Table 1.2: Markers increasing and decreasing predicted valence metric. likelihood.

Lastly, the engagement of the subject is measured as a weighted sum of the following facial expressions:

- Brow raise,
- Brow furrow,
- Nose wrinkle,
- Lip corner depressor,
- Chin raise,
- Lip pucker,
- Lip press,
- Mouth open,
- Lip suck,
- Smile.

Due to the ease of use, clear documentation, offline capabilities and free 6 month trial, we have decided to proceed using Affectiva's Affdex SDK.

2. Data gathering methodology

As we've demonstrated, there are great many approaches to take and physiological markers to explore. Due to the nature of our desired use-case, we've decided to omit those physiological signs and pertaining sensors, that would excessively constrain and disturb the user. Similarly, we did not employ techniques, that would restrict our application from workplace or school environment. Therefore, we did not use EEG, fMRI scanning, neither Thermal imagining, nor speech recognition. For our body sensors, we have chosen a simple, inexpensive optical heart rate sensor, GSR[1] sensor and Arduino to read and resend collected data. Furthermore, we used a headset with two high-speed infrared cameras with infrared illumination and one 1920x1080 "world" camera. To provide referential emotion values, we used a regular 720p web camera, and processed the camera feed with modified Affdex sample OpenCV application.

2.1. Hardware

The research has been done using two different computers.

- Desktop workstation with hexa-core AMD FX-6300 clocked at 4.0GHz, 8GB of RAM and Radeon 7700 GPU,
- Laptop with quad-core Intel Core i7-4770HQ, 16 GB of RAM and NVidia 860M GPU.

These two computers proved able to fully satisfy hardware needs of the research. We have also tried to run the modifiet Affdex sample application on Raspberry Pi 3B with official IR camera, however we found out, that the single board computer was greatly lacking computing power and would freeze, hang and drop frames with even one face detected on video feed. Powerful hardware was needed not only for the use of Affectiva SDK, but also for the Pupil Capture software, supplied by the manufacturer of our headset. These two applications taxed both our computers close to peak load, although since our EmoSens data collection application requires very few resources, it did not negatively impact our experiments.

The heart rate data was gathered by Pulse Sensor Amped SEN-11574, a plug-and-play heart rate sensor for Arduino and Raspberry Pi. It's a simple, inexpensive optical sensor, that incorporates amplification and noise cancellation circuitry on a single board. We did find out, that it requires tight strapping to subject's skin, otherwise the data collected has been inconsistent with reality.

[1] Galvanic Skin Response

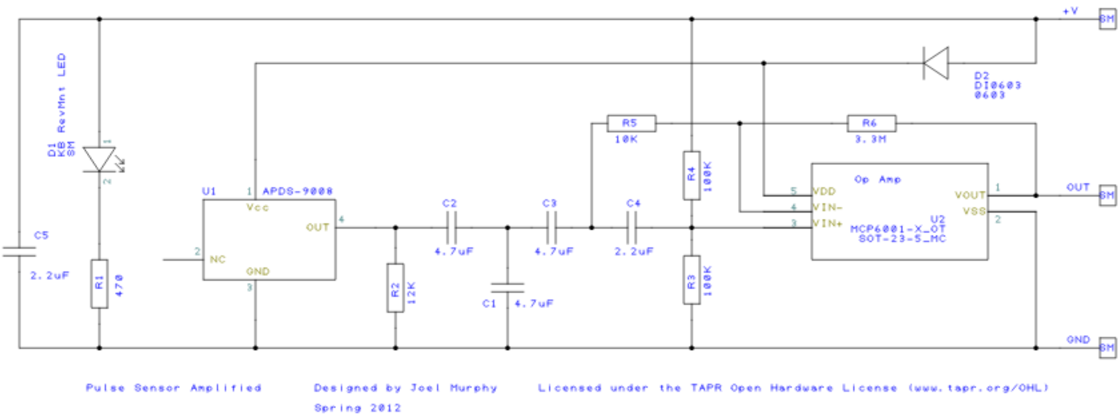


Figure 2.1: Pulse Sensor Amped SEN-11574 circuitry

The Galvanic Skin Response data was collected using Groove GSR sensor. Groove is an Arduino shield maker, with a plethora of different sensors available to buy. Since we only needed GSR and it is a relatively trivial task to wire the Sensor directly to Arduino while circumventing the need for the Groove shield itself, we connected the Groove GSR sensor straight to Arduino's GPIO pins.

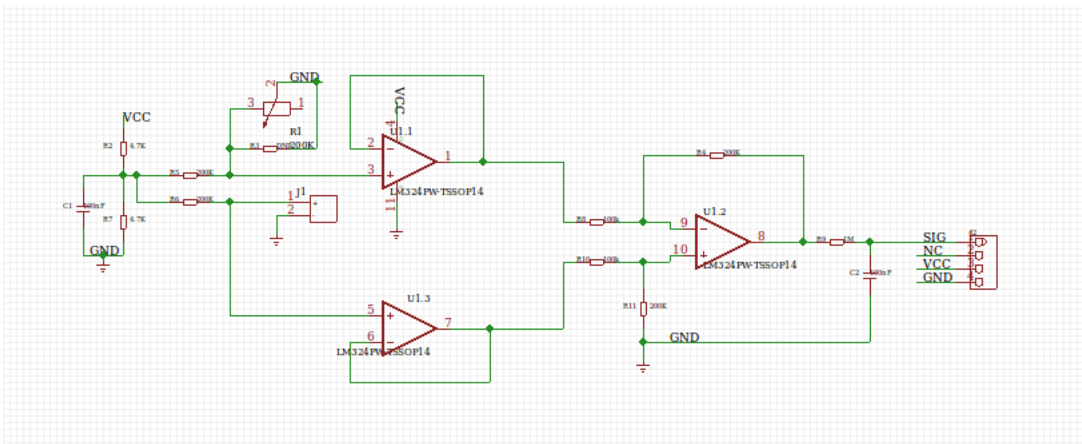


Figure 2.2: Groove GSR Sensor circuitry

Both of these sensors were sewn into a fingerless glove and wired to Arduino Due connected via USB cable to computer.

(a) Pulse Sensor Amped SEN-11574

(b) Groove GSR Sensor

Figure 2.3: Sensors

The Pupil headset consists of two high-speed 200Hz IR eye cameras with illumination and a single FullHD World camera connected together to a computer using USB-A to USB-C connector.

Figure 2.4: Pupil headset

2.2. Software

Both development and experiments took place on computers running Ubuntu 16.04. We have chosen this operating system, as it has large community of developers, computer scientists and because most innovative frameworks and SDKs target this platform. To create the data collection software, we have used the Qt5 application framework. It's main advantages in our use-case is it's intuitiveness, open-source and cross-platform compatibility. It has bindings with Python[2] and C++. It also adds extensions to C++, including signals and slots, which greatly simplify handling of events, thus helping in development of GUI. It was compiled by GNU GCC 5.4.0. The Arduino is programmable in a it's own variant of simplified C++, that forgoes features of the language such as inheritance. Last but not least, the modified Affdex webcam application is written in C++, albeit due to the incompatibilities with newer compilers was compiled in GNU GCC 4.8.

This leads us to the unexpected setback that we came across. Both Affectiva and Pupil use OpenCV 2.4, but Affectiva forked it's version and when building with CMake on a recent operating system, it will either fail, or overwrite shared libraries, thus breaking QT5 and Pupil. One solution to this problem might be to install Affectiva sample application on a separate computer, but since we aimed to create a compact solution, we opted for installing it in a virtual machine running stock Ubuntu 16.04 with only the libraries, that Affectiva requires and nothing else.

[2] PyQt

Since Affectiva's licence agreement prohibits it's integration with a copyleft open source licensed projects [3] we have modified the sample OpenCV application to send out data serialized in messagepack over zeromq, thus eliminating the need to include it in our project. This has added benefit in that our project is therefore modular and can use any and all face or other emotion recognition SDK with little to no changes to our source code.

ZeroMQ is a large community of projects with focus on decentralized messaging and computing engines. Both software and protocols are open source and maintained by community of experts. The original core engine is libzmq, available in repositories of most GNU/Linux distributions, but also available to install on all the other major operating systems. On top of libzmq, there is a large number of language bindings, such as PyZMQ, CZMQ, CPPZMQ, NZMQT [4]. There are also native languages written for Java, .NET, JavaScript and many others. Most importantly, it's really lightweight and fast. The core library is only 20K lines of C++ code and the message throughput is 8 million messages per second with latency in microseconds. There are three main patterns used in ZeroMQ. Request-Reply, Publish-Subscribe and Pipeline pattern. (Hintjens, 2013 ; Akgul, 2013)

MessagePack is a low-overhead binary serialization format. Compared to the better know binary serialization format BSON, it's better in every possible technical aspect. It's faster, smaller, and even more compatible to JSON than BSON is. It's got support for dozens of languages, notably for Python, Java, C, C# and multiple standards of C++. MessagePack by design solves the problem of interprocess communication, especially if the processes are written in different languages, or are even running on different platforms.

For our project, we opted to use the Publish-Subscribe pattern, with Pupil Capture and Affectiva being the Publishers and our application serving as a Subscriber to their messages serialized with MessagePack. That way we ensure, that the communication is fast and reliable and also that the peculiarities of different programming languages are covered for ease of use.

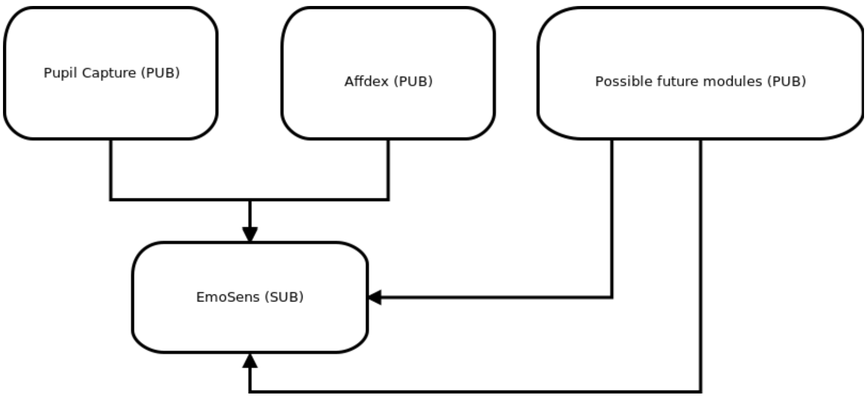


Figure 2.5: Simplified diagram of the data flow.

For visualization of collected data, we have used two QT libraries. The QtCharts library, which is included in the open-source Qt 5.10 installation, and also the QCustomPlot

[3] In our case GNU GPL v2

[4] Our project uses NZMQT due to it's great integration with QT5's signals and slots

for more complex and granular control. QCustomPlot is a free and open-source Qt C++ widget for plotting and data visualization. It has no dependencies and is well documented. It even supports exporting to various format such as vectorized PDF files and rasterized pictures in PNG, JPG and BMP.

2.3. EmoSens application

2.4. The experiments

3. Data evaluation

4. Results

5. Discussion

Bibliography

Affectiva Automotive AI. Available also from: https://www.affectiva.com/product/affectiva-automotive-ai/.

AKGUL, Faruk, 2013. *ZeroMQ*. Packt Publishing. ISBN 178216104X, 9781782161042.

ASHKANASY, Neal; COOPER, Cary, 2008. Research companion to emotion in organizations.

ASTERIADIS, Stylianos; TZOUVELI, Paraskevi K.; KARPOUZIS, Kostas; KOLLIAS, Stefanos D., 2008. Estimation of behavioral user state based on eye gaze and head pose - application in an e-learning environment. *Multimedia Tools and Applications*. Vol. 41, pp. 469–493.

BARRETT, L. F., 2007. Language as context for the perception of emotion. *Trends in Cognitive Sciences*. Vol. 11, no. 8. ISSN 1364-6613.

BEATTY, Jackson; LUCERO-WAGONER, Brennis, 2012. The pupillary system.

BOYLE, Gregory J.; HELMES, Edward; MATTHEWS, Gerald; IZARD, Carroll E., 2015. Chapter 8 - Measures of Affect Dimensions. In: BOYLE, Gregory J.; SAKLOFSKE, Donald H.; MATTHEWS, Gerald (eds.). *Measures of Personality and Social Psychological Constructs*. San Diego: Academic Press, pp. 190 –224. ISBN 978-0-12-386915-9. Available from DOI: https://doi.org/10.1016/B978-0-12-386915-9.00008-5.

BRACKEEN, Brian. *Face Off: Confronting Bias in Face Recognition AI*. Available also from: https://www.kairos.com/blog/face-off-confronting-bias-in-face-recognition-ai.

CANNON, W.B., 1927. *The James-Lange Theory of Emotions: A Critical Examination and an Alternative Theory*. University of Illinois Press. Available also from: https://books.google.sk/books?id=DE3BnQEACAAJ.

COWEN, Alan S.; KELTNER, Dacher, 2017. Self-report captures 27 distinct categories of emotion bridged by continuous gradients. *Proceedings of the National Academy of Sciences*. ISSN 0027-8424. Available from DOI: 10.1073/pnas.1702247114.

CRUZ-ALBARRAN, Irving A.; P. BENĂTEZ-RANGEL, Juan; OSORNIO-RIOS, Roque; MORALES-HERNĂˇNDEZ, Luis, 2017. Human emotions detection based on a smart-thermal system of thermographic images. Vol. 81.

DARWIN, C.; EKMAN, P.; PRODGER, P., 1998. *The Expression of the Emotions in Man and Animals*. Oxford University Press. ISBN 9780195158069. Available also from: https://books.google.sk/books?id=TFRtLZSHMcYC.

DAVLETCHAROVA, Assel; SUGATHAN, Sherin; ABRAHAM, Bibia; JAMES, Alex Pappachen, 2015. Detection and Analysis of Emotion from Speech Signals. *Procedia Computer Science*. Vol. 58, pp. 91–96. ISSN 1877-0509. Available from DOI: https://doi.org/10.1016/j.procs.2015.08.032.

EKMAN, P., 1992. An Argument for Basic Emotions. *Cognition and Emotion*. Vol. 6, no. 3-4, pp. 169–200. Available from DOI: 10.1080/02699939208411068. cited By 2387.

EKMAN, P.; ROSENBERG, E.L., 1997. *What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)*. Oxford University Press. Series in affective science. ISBN 9780195104462. Available also from: https://books.google.sk/books?id=fFGYs079-7YC.

EKMAN, Paul; V. FRIESEN, Wallace, 1971. Constants across cultures in the face and emotion. Vol. 17, pp. 124–9.

GRAVINA, Raffaele; ALINIA, Parastoo; GHASEMZADEH, Hassan; FORTINO, Giancarlo, 2016. Multi-Sensor Fusion in Body Sensor Networks: State-of-the-art and research challenges. Vol. 35.

H. HESS, Eckhard, 1975. The tell-tale eye: How your eyes reveal hidden thoughts and emotions.

HINTJENS, P., 2013. *ZeroMQ: Messaging for Many Applications*. O'Reilly Media, Incorporated. Oreilly and Associate Series. ISBN 9781449334062. Available also from: https://books.google.sk/books?id=KWtT5CJc6FYC.

IOANNOU, Stephanos; GALLESE, Vittorio; MERLA, Arcangelo, 2014. Thermal infrared imaging in psychophysiology: Potentialities and limits. In: *Thermal infrared imaging in psychophysiology: Potentialities and limits. Psychophysiology.*

J. LANG, Peter; M. BRADLEY, Margaret; N. CUTHBERT, Bruce, 1990. Emotion, Attention, and the Startle Reflex. Vol. 97, pp. 377–95.

KRAGEL, Philip A.; KNODT, Annchen R.; HARIRI, Ahmad R.; LABAR, Kevin S., 2016. Decoding Spontaneous Emotional States in the Human Brain. *PLOS Biology*. Vol. 14, no. 9, pp. 1–19. Available from DOI: 10.1371/journal.pbio.2000106.

LAHANE, Prashant; SANGAIAH, Arun Kumar, 2015. An Approach to EEG Based Emotion Recognition and Classification Using Kernel Density Estimation. *Procedia Computer Science*. Vol. 48, pp. 574 –581. ISSN 1877-0509. Available from DOI: https://doi.org/10.1016/j.procs.2015.04.138. International Conference on Computer, Communication and Convergence (ICCC 2015).

LÖVHEIM, Hugo, 2012. A new three-dimensional model for emotions and monoamine neurotransmitters. *Medical hypotheses*. Vol. 78 2, pp. 341–8.

MASKELIUNAS, Rytis; RAUDONIS, Vidas, 2016. Are you ashamed? Can a gaze tracker tell? Vol. 2, pp. e75.

MCDUFF, Daniel; MAHMOUD, Abdelrahman; MAVADATI, Mohammad; AMR, May; TURCOT, Jay; KALIOUBY, Rana el, 2016. AFFDEX SDK: A Cross-Platform Real-Time Multi-Face Expression Recognition Toolkit, pp. 3723–3726. Available also from: https://www.affectiva.com/wp-content/uploads/2017/03/McDuff_2016_Affdex.pdf.

MEHRABIAN, A., 1980. *Basic Dimensions for a General Psychological Theory: Implications for Personality, Social, Environmental, and Developmental Studies*. Oelgeschlager, Gunn & Hain. Social Environmental and Developmental Studies. ISBN 9780899460048. Available also from: https://books.google.sk/books?id=1RR9AAAAMAAJ.

MEHRABIAN, A.; RUSSELL, J.A., 1974. *An approach to environmental psychology*. M.I.T. Press. ISBN 9780262130905. Available also from: `https://books.google.sk/books?id=Eth0AAAAMAAJ`.

MOONS, Wesley G.; EISENBERGER, Naomi I.; TAYLOR, Shelley E., 2010. Anger and fear responses to stress have different biological profiles. *Brain, behavior, and immunity*. Vol. 24 2, pp. 215–9.

MOSCIANO, Francesco; MENCATTINI, Arianna; RINGEVAL, Fabien; SCHULLER, BjĂ¶rn; MARTINELLI, Eugenio; NATALE, Corrado Di, 2017. An array of physical sensors and an adaptive regression strategy for emotion recognition in a noisy scenario. *Sensors and Actuators A: Physical*. Vol. 267, pp. 48 –59. ISSN 0924-4247. Available from DOI: `https://doi.org/10.1016/j.sna.2017.09.056`.

PHILIPPOT, Pierre; CHAPELLE, GaĂ«tane; BLAIRY, Sylvie, 2010. Respiratory feedback in the generation of emotion. Vol. August 01, pp. 605–627.

PLUTCHIK, Robert, 1988. The Nature of Emotions: Clinical Implications. In: *Emotions and Psychopathology*. Ed. by CLYNES, Manfred; PANKSEPP, Jaak. Boston, MA: Springer US, pp. 1–20. ISBN 978-1-4757-1987-1. Available from DOI: `10.1007/978-1-4757-1987-1_1`.

POSNER, Jonathan; RUSSEL, James A.; PETERSON, Bradley S., 2005. The circumplex model of affect: An integrative approach to affective neuroscience, cognitive development, and psychopathology. *Development and Psychopathology*. Vol. 17, no. 3, pp. 715–734.

RUBIN, David C.; TALARICO, Jennifer M., 2009. A comparison of dimensional models of emotion: Evidence from emotions, prototypical events, autobiographical memories, and words. *Memory*. Vol. 17, no. 8, pp. 802–808. Available from DOI: `10.1080/09658210903130764`. PMID: 19691001.

RUSSELL, J. A.; BARRETT, Lisa Feldman, 1999. Core affect, prototypical emotional episodes, and other things called emotion: dissecting the elephant. *Journal of personality and social psychology*. Vol. 76 5, pp. 805–19.

RUSSELL, J.A., 2003. Core Affect and the Psychological Construction of Emotion. Vol. 110, pp. 145–72.

RUSSELL, James, 1980. A Circumplex Model of Affect. Vol. 39, pp. 1161–1178.

SALAZAR-LOPEZ, E.; DOMĂNGUEZ, E.; RAMOS, V. JuĂˇrez; FUENTE, J. de la; MEINS, A.; IBORRA, O.; GĂˇLVEZ, G.; RODRĂGUEZ-ARTACHO, M.A.; GĂLMEZ-MILĂˇN, E., 2015. The mental and subjective skin: Emotion, empathy, feelings and thermography. *Consciousness and Cognition*. Vol. 34, pp. 149 –162. ISSN 1053-8100. Available from DOI: `https://doi.org/10.1016/j.concog.2015.04.003`.

SCHLOSBERG, Harold, 1954. Three dimensions of emotion. *Psychological review*. Vol. 61, no. 2, pp. 81.

SHALINI, T Baby; VANITHA, L, 2013. Emotion detection in human beings using ECG signals. *International Journal of Engineering Trends and Technology (IJETT)*. Vol. 4, no. 5, pp. 1337–1342.

TARNOWSKI, PaweĹ,; KOĹ,ODZIEJ, Marcin; MAJKOWSKI, Andrzej; RAK, Remigiusz J., 2017. Emotion recognition using facial expressions. *Procedia Computer Science*. Vol. 108, pp. 1175 –1184. ISSN 1877-0509. Available from DOI: `https://doi.org/10.1016/j.procs.2017.05.025`. International Conference on Computational Science, ICCS 2017, 12-14 June 2017, Zurich, Switzerland.

TOMKINS, S., 1962. *Affect Imagery Consciousness: Volume I: The Positive Affects*. Springer Publishing Company. Springer Series. ISBN 9780826104427. Available also from: `https://books.google.sk/books?id=WIpgNerqaIkC`.

TOMKINS, S., 1963. *Affect Imagery Consciousness: Volume II: The Negative Affects*. Springer Publishing Company. Springer Series. ISBN 9780826104434. Available also from: `https://books.google.sk/books?id=1xL\_TT04OxsC`.

WATSON, David; TELLEGEN, Auke, 1985. Toward a Consensual Structure of Mood. Vol. 98, pp. 219–35.

WUNDT, W.M., 2017. *Outlines of Psychology (Classic Reprint)*. Fb&c Limited. ISBN 9780331795738. Available also from: `https://books.google.sk/books?id=zj6atAEACAAJ`.

YIK, M.; RUSSELL, J. A.; STEIGER, J. H., 2011. A 12-Point Circumplex Structure of Core Affect. *Emotion*. Vol. 11, no. 4, pp. 705–731. Available also from: `www.scopus.com`.

YONCK, R., 2017. *Heart of the Machine: Our Future in a World of Artificial Emotional Intelligence*. Arcade Publishing. ISBN 9781628727340. Available also from: `https://books.google.sk/books?id=HquvDQAAQBAJ`.

ZHANG, Qiang; CHEN, Xianxiang; ZHAN, Qingyuan; YANG, Ting; XIA, Shanhong, 2017. Respiration-based emotion recognition with deep learning. *Computers in Industry*. Vol. 92-93, pp. 84–90. ISSN 0166-3615. Available from DOI: `https://doi.org/10.1016/j.compind.2017.04.005`.

ZHANG, Yong; JI, Xiaomin; ZHANG, Suhua, 2016. An approach to EEG-based emotion recognition using combined feature extraction method. Vol. 633.

ZUALKERNAN, I.; ALOUL, F.; SHAPSOUGH, S.; HESHAM, A.; EL-KHORZATY, Y., 2017. Emotion recognition using mobile phones. *Computers & Electrical Engineering*. Vol. 60, pp. 1–13. ISSN 0045-7906. Available from DOI: `https://doi.org/10.1016/j.compeleceng.2017.05.004`.