

Are Instructed Emotional States Suitable for Classification? Demonstration of How They Can Significantly Influence the Classification Result in An Automated Recognition System

M. Magdin*, F. Prikler

Constantine the Philosopher University in Nitra, Faculty of Natural Sciences , Department of Computer Science, Tr. A. Hlinku 1, 949 74 Nitra, Slovakia

Received 7 February 2018 | Accepted 7 March 2018 | Published 16 March 2018



ABSTRACT

At the present time, various freely available or commercial solutions are used to classify the subject's emotional state. Classification of the emotional state helps us to understand how the subject feels and what he is experiencing in a particular situation. Classification of the emotional state can thus be used in various areas of our life from neuromarketing, through the automotive industry (determining how emotions affect driving), to implementing such a system into the learning process. The learning process, which is the (mutual) interaction between the teacher and the learner, is an interesting area in which individual emotional states can be explored. In this pedagogical-psychological area several research studies were realized. These studies in some cases demonstrated the important impact of the emotional state on the results of the students. However, for comparison and unambiguous classification of the emotional state most of these studies used the instructed (even constructed) stereotypical facial expressions of the most well-known test databases (Jaffe is a typical example). Such facial expressions are highly standardized, and the software can recognize them with a fairly big percentage, but this does not necessarily point to the actual success rate of the subject's emotional classification in such a test because the similarity to real emotional expression remains unknown. Therefore, we examined facial expressions in real situations. We have subsequently compared these examined facial expressions with the instructed expressions of the same emotions (the Jaffe database). The overall average classification score in real facial expressions was 94.58%.

KEYWORDS

Instructed Stimuli,
Real-life Stimuli, Facial
Expression, Emotion,
Face Detection.

DOI: 10.9781/ijimai.2018.03.002

I. INTRODUCTION

FACE recognition and subsequently classification of emotional state is a very important research area of learning process in recent days. This fact was confirmed in study [2], where result of this study shows that emotions may have a significant effect on human trust because the necessary part of social communication between people is the skill to correctly express their emotions. Expressions in face are for all people not only easy movements of face parts but also are the most efficient means to express emotion [10], [16]. In the mutual communication we need correct understanding of emotional state. Therefore we need not only perceive correctly change of muscles in face. We must also know how decode this ostensible movement of muscle in the right way for the following classification of the emotion that is represented by the facial expression. [17], [20], [42]. In present face recognition is an important area not only in research but also in various applications using human-computers interface. For normal human is not a problem to recognize various expressions in face (without delay), however recognition expression of face by computer in real time is still a big problem [32], [28]. The past decade has witnessed many new developments in facial

expression analysis due to its wide application in robotic vision [37], [24], forensics [5], affective computing, man-machine interaction [29], [18] and even medical diagnosis [24]. Therefore we can say that the correct classification of emotional state is an important role in various applications and industries of our life, for example: robust tools for behavioral people research, speech processing and speech recognition or access control of people by building monitoring, etc. [25]. Expression of emotions in face is the most natural way for people to express not only their emotions but also their intentions. This is a relatively simple task for a human from view point of detection in real time. However, for a computer, as automatic recognition system, this task is not simple [36]. We can identify the emotional state in several ways, also on the bases of voice analysis, gesture analysis, analysis of handwritten texts, etc. The different studies in last years have shown that many characteristic features needed to classify the emotional state are expressed especially by the expressions of the face [28] in real time or emotion recognition from facial images, e.g. [6], [34] and [41]. In these research studies researchers reported on significant differences of examined subjects. Recently "EMOTHAW" (EMOTION recognition from HAndWriting and drAWing) was designed and created. EMOTHAW is a first publicly available database (<https://sites.google.com/site/becogsys/emothaw>) for recognition of emotional states on the basis handwriting [27]. Recognition and classification of emotional state by using multiple ways (so-called as multimodal recognition of

* Corresponding author.

E-mail address: mmagdin@ukf.sk

emotions) is currently considered in the research area as a perspective area, especially because of the big amount of data that can be acquired concurrently on the basis of different physiological signals. These physiological signals are a natural expression of the human body and therefore they can be successfully used to classify the emotional state. Such physiological signals may be electroencephalogram (EEG), temperature or electrocardiogram (ECG). Through them, we can classify an emotional state such as happiness, sadness or anger [39]. On the basis of different sensors we can capture these signals and we can identify individual difference factors. These factors significantly affect the ability to classify emotions (see e.g. [21] for an overview).

Multimodal face recognition is based on capture of individual facial features and observing of each difference. The parts of the face that can be captured are: facial features, eyelids, eyebrows, lips, nose, chin, etc. For this purpose, we can use the so-called “multi-point masks”. Individual points on this mask represent the extracted areas from which the classification of the emotional state is performed [7]. Because of the amount of individual parts on this mask that can be scanned simultaneously and the amount of output data thus produced, there have been developed various methods for detecting, extracting and subsequently classifying the emotional state either in real time or using stored face images. The effectiveness of these methods is evaluated using facial image databases with different expressions [25]. The first step to be able to respond to emotions is affect recognition that focuses on identifying emotions and other affective phenomena on the subject. The evaluation of the affective state is usually done according to an emotional model that suits the particular application. One of the simplest models is the one described by Ekman, which is composed of six discrete primitive emotions, namely anger, fear, sadness, surprise, disgust and happiness [12]. Other alternative models include Plutchik’s Wheel of Emotion [35] and Russell’s Circumplex Model [38], which locates emotions in a 2D space defined by the arousal (or activation), and valence (or positiveness). The latter was extended in [31] by adding a third dimension (dominance) to avoid overlapping of certain emotions [3].

From academic year 2016/2017 in our department of computer science a project has been realized: Modeling the behavior of users based on data mining with support from IBM Bluemix. Our aim in this project is the design and creation of a complex automatic system for detection, face features extraction and classification of the emotional state. In this moment our final application is capable of detecting face in real time, to extract various face features on the basis of the use of multi-point mask and the classification of the emotional state by Ekman scale. The data obtained from the web camera (using data mining techniques) is used as a basis for user behavior modeling in Moodle environment. According to the literature analysis of the evaluation of the emotional state, there are various solutions to this problem; however, all solutions allude to the fact - evaluating the amount of mined data of various participants in real time. In this publication, we point out the results of the solution that has significantly influenced our further research direction.

In the section Material and methods, we are focusing on researches dealing with face recognition using real-time webcam, followed by extraction of individual facial features and classification of the emotional states of students. In the Calculation chapter, we describe the methodology of the two experiments we used. In Experiment 1, we focused primarily on determining the degree of classification success using constructed expressions found in the Japanese models database called Jaffe. This is the most used database for determining emotional states due to precisely tagged images with facial expressions of individual subjects. The results of Experiment 1 are listed in the Results and Discussion section of Experiment 1. Based on the results obtained and their evaluation, we have come to the realization of

Experiment 2, which evaluates real-life expressions of emotion. Results and recommendations are listed in the Results and Discussion section.

II. METHODOLOGY

In our daily life, during performing various activities, we create and generate different types of data in a large quantity. This data contains an important information regarding our activities, style of living, our behavior in different situations and our emotions. Psychologists have gradually classified different types of emotions, such as love, joy, surprise, anger, sadness and fear. The classification model of Paul Ekman is considered a reference model. It contains emotional states as happiness, anger, disgust, fear, sadness, surprise. Later on, the neutral state was included in these six emotional states [23]. Certain facial expressions are recognizable signals of emotion: a smile signals happiness, crying signals sadness, a nose scrunch signals disgust [14], [15], [9], [26] is the cornerstone of a dominant theory of emotion. These facts are an important prerequisite for various emotion research, or in affective neuroscience, and in a range of applications. A survey of emotion researchers found that a large majority accepted the thesis that certain faces signal specific emotions [13], [11].

In 1975 a method of coding for the face using the action units (FACS) was proposed [16]. FACS was containing 46 points within the geometric face model. These facial parts were the basis for the design and implementation of multi-point masks. These facial parts and facial points have been designed not only for the detection and extraction of the necessary areas of interest, but with their changes, it was possible to classify the emotional state. Each emotion was defined on the basis of a facial point change. Over time, this extraction-classification method was enhanced by the Gaussian classifier, by the Hidden Markov Model (HMM) or the Bayes classifier [8]. The detection phase has been progressively improved. Instead of the robust Viola-Jones detector, the use of the Support Vector Machine (SVM) method has begun. This method completely replaced all three phases of the recognition process: detection, extraction, and classification [40], [43].

These three phases of the recognition process are important for various area researches. The results from recognition process can be used in various fields of our life as investigation of impact of some products on human behavior (neuromarketing), impact of advertisements on human choosing, etc. Emotion recognition accuracy (ERA) from faces has been conceptualized as a performance measure of emotional intelligence [29]. Currently for classification of emotional state, several research groups use in most cases various databases of human faces. However these databases contain only instructed emotional state [22]. These images are organized to individual databases and contain people’s faces, which differ depending on the lighting, the angle of the face rotation, the noise in the images and others. These differences are the basis for determining the percentage of success rate of face recognition and subsequent classification of the emotional state [33]. The problem is often the low count of images in the database. Therefore, it is questionable whether the standard algorithms used would work in a realistic situation and then emotional state would be classified with the same percentage of success rate. We can claim that, basing on the fact that a lot of research uses a method in which users mimic the emotional state of a person in the picture from the database, certain face expressions are highly consistent with images that contain databases and achieve a very high percentage of success classification [11]. The software then does not recognize the real states, but the states that are artificially induced. Several experts have pointed out this problem in their work [1], [11].

Based on their opinion, we are inclined to solve the proposed pilot project by submitting the realized evaluation solution, both from the point of proper facial detection under different lighting conditions, the

distance of the subject or the rotation of the subject's face, as well as determining the rate of success of the emotional state classification of the subject.

The success rate of the face detection in the image was discussed in the paper titled "Real Time Facial Expression Recognition Using Webcam and SDK Affectiva" [44], where the proposed solution was able to detect the face in real time with 100% success rate in frontal face situations. We are having problems only in the case when face is rotated in left or right, having a total average success rate of 84,27%. Detection was tested using 6 robust databases on a total of 9252 frames:

- Bao Face Database,
- CMU/VASC Image Database,
- Caltech Faces 1999 Database,
- NIST Mug-shot Images Database,
- Yale Face Database,
- Vision Group of Essex University Face Database.

In the following section, we present an experiment that determined the success rate of the emotional state classification of the subject under study (Jaffe database).

III. MEASURES AND PROCEDURE OF EXPERIMENT 1 – DETERMINING THE DEGREE OF SUCCESS OF THE EMOTIONAL STATUS CLASSIFICATION USING THE JAFFE DATABASE

The methods that we can use to analyze emotional states have a same basis as methods used for face recognition. Analysis of emotional state is standard divided into three basic phases [18]: detecting the face or face part on the image, extracting the area of interest, and classifying the emotional state. The use of a suitable classifier always depends on what the desired output will be. Classifiers may not be priority designed for phase of classification of the emotional state. We can use these classifiers appropriately also to filter the properties we are interested. At present there are a lot of different methods for detection, extraction and classification (about 200) and for the experienced programmer their use is not an easy process. On the other hand, however, we can use available libraries, apps and SDKs that are already verified, when we are designing and creating an automatic recognition system. Such a SDK solution is also a library from Affectiva. Face detection is performed using the Viola-Jones detector. Histogram of Oriented Gradient (HOG) is extracted from the image area of interest defined by the face orientation points. Support Vector Machine Algorithm (SVM), trained on 10,000 manually encoded face images collected from around the world is used to generate a score from 0 to 100 for each face action. For details on the training and testing system, see [44]. The use of such a solution brings the advantage that in the final phase of the recognition process (classification) we do not have to deal with the so-called "training set". A variety of methods are used to verify the overall functionality of the systems and also to determine the degree of classification success. The most commonly used method, however, is when the subject has to emulate the emotional expression and then the system must correctly classify it. This is done using images from available facial databases with corresponding emotional expressions (according to Ekman's classification). The entire testing process is performed under the supervision of a psychologist to rule out the possibility of incorrect classification or incorrectly instructed expression. More about this classification can be found in [4].

The biggest problem of different solutions is in the listed databases. Researchers often use only one database to determine the percentage of success rate of the recognition system, which ultimately results comparable to other top-level systems. Such testing is quite trivial and inconsistent [6].

We realized the solution may be evaluating using individual databases. These databases contain various images with different light conditions, face rotation, or with several captured faces in the image. This result is for us a comprehensive assessment of the rate of face detection success [44]. The percentage of success rate of the proposed solution was verified as follows (Fig. 1): on one side of a table, we placed a monitor on which the individual images from the respective databases were gradually displayed; on the other side of the table, we placed a laptop with an application for face recognition and classification of emotional state at a distance of 1m; the laptop contained a webcam that captured the image of the monitor; the distance of 1m is the standard distance to capture the object by a webcam.



Fig. 1. Verification of the percentage of success rate of the proposed solution.

To simplify the entire detection process and subsequent classification, we removed the step in which the subject would have to emulate emotions. Because these databases contained only the different faces of the subjects without the precisely defined type of emotional state, we were looking for a database that would contain the rated images. Such database is Jaffe, which contains 214 Japanese model photos (marked KA, KL, KM, KR, MK, NA, NM, TM, UY, YM). It is important to note that these models were asked to reproduce the desired emotion: anger, disgust, fear, happy (joy), neutral, sadness and surprise (Fig. 2). For this reason, we can say that those were forced expressions of the face where the monitored subject (in this case the Japanese model) was fully aware of the emotion she was trying to imitate according to instructions. The test set of images includes 10 people who were photographed during a particular emotion expression 3 to 4 times.



Fig. 2. Faces selected from the tested images with an instructed emotional state (Left to Right - Anger, Disgust, Fear, Happiness, Neutral, Sadness and Surprise)

We've continuously recorded the output from our classification software, which expressed the intensity of the classified emotion in percentage (the classifier is considered successful if it reaches a recognition rate of more than 50%). Testing for each of the Japanese models (Fig. 3) was repeated 5 times (but we have always achieved the same result, confirming that the software correctly detected the face in the same way and also classified the emotional state with the same result).



Fig. 3. Demonstration of facial expression classification (emotional state - fear).

IV. RESULTS AND DISCUSSION – EXPERIMENT 1

As we can see from Tables I and II, the software often classified a completely different emotional state with high percentage, as was expected.

TABLE I. THE RATE OF SUCCESS OF CLASSIFICATION (IN PERCENTAGE) OF THE EMOTIONAL STATE OF THE KM MODEL

		Recognized Emotion by the Software						
		An	Di	Fe	Ha	Ne	Sa	Su
Requested Emotions (Model name: KM)	An	0,00	1,75	1,34	0,00	94,32	0,00	2,59
	Di	0,00	61,78	0,00	0,00	34,46	0,00	3,76
	Fe	0,00	16,00	33,61	0,00	32,97	5,43	11,99
	Ha	0,00	0,14	0,38	92,16	0,07	0,23	7,02
	Ne	0,00	1,09	31,12	0,00	50,84	0,00	16,95
	Sa	0,00	1,47	11,72	0,00	72,63	1,40	12,78
	Su	0,00	0,00	0,00	0,00	100,00	0,00	0,00

Explanation of Table I:

- An – Anger
- Di – Disgust,
- Fe – Fear,
- Ha – Happy (Joy),
- Ne – Neutral,
- Sa – Sadness,
- Su – Surprise.

When classifying the model labelled as KM, the software rated an emotional state anger as a neutral expression (94.32%), also incorrectly determined the emotional state of sadness (72.63%) and surprise (100%) as neutral. However, after consulting with a psychology specialist, we have come to the conclusion that these expressions are problematic to interpret by the model KM, and that's why the software classified them as another type of emotional state (Fig. 4).



Fig. 4. Example of a problematically instructed emotional state (left Anger, right Sadness).

The situation was similar with the model labelled MK. The software classified anger as very neutral (97.48%), also misinterpreted disgust

(91.14%), sadness (97.93%) and surprise (100%) as neutral expressions (Table II).

TABLE II. THE RATE OF SUCCESS OF THE CLASSIFICATION (IN PERCENTAGE) OF THE EMOTIONAL STATE OF THE MK MODEL

		Recognized Emotion by the Software						
		An	Di	Fe	Ha	Ne	Sa	Su
Requested Emotions (Model name: MK)	An	0,00	1,45	0,00	0,00	97,48	0,00	1,07
	Di	0,00	2,66	0,00	0,00	91,14	4,03	2,17
	Fe	0,00	0,53	50,29	0,00	42,05	0,36	6,77
	Ha	0,00	0,00	0,00	99,84	0,05	0,00	0,11
	Ne	0,00	1,56	0,00	0,00	97,03	0,00	1,41
	Sa	0,00	1,19	0,00	0,00	97,93	0,00	0,88
	Su	0,00	0,00	0,00	0,00	100,00	0,00	0,00

After consulting a specialist in the psychology department, we also came to the conclusion that these expressions are misinterpreted by MK models and therefore classified by the software as another type of emotional state (Fig. 5).



Fig. 5. Example of a problematically instructed emotional state - Sadness.

For the model labeled UY (See Table III), the software has always evaluated (albeit to a minimal extent) each of the expressions as disgust (0.43-0.45%). Anger was classified as neutral state (87.31%), disgust as neutral state (80.16%), sadness as neutral state (87.76%) and surprise as neutral state (51.82%).

TABLE III. THE RATE OF SUCCESS OF CLASSIFICATION (IN PERCENTAGE) OF THE EMOTIONAL STATE OF THE UY MODEL

		Recognized Emotion by the Software						
		An	Di	Fe	Ha	Ne	Sa	Su
Requested Emotions (Model name: UY)	An	0,00	0,43	11,84	0,00	87,31	0,00	0,42
	Di	0,00	0,44	18,97	0,00	80,16	0,00	0,43
	Fe	0,00	0,43	49,73	0,00	49,28	0,00	0,56
	Ha	0,00	0,43	33,60	64,33	0,11	0,98	0,55
	Ne	0,00	0,43	35,76	0,00	63,27	0,00	0,54
	Sa	0,00	0,43	11,44	0,00	87,76	0,00	0,37
	Su	0,00	0,45	47,10	0,00	51,82	0,00	0,63

After consulting a psychology specialist, we also came to the conclusion that these expressions are misinterpreted by the model labelled as UY, and therefore classified by the software as another type of emotional state. As an example, we are showing the type of emotional state –anger, which was supposed to be expressed by the model (Fig. 6). However, according to the psychologist, this is a neutral expression, which is also confirmed by the recognition results.



Fig. 6. An example of a problematically induced emotional state - Anger.

This incorrect classification of the emotional states in these 3 models caused the average rate of classification to fall to 30.01% (Table IV). In other cases, we did not detect any errors or deviations and the software correctly detected the emotional states.

TABLE IV. AVERAGE RATE OF EMOTIONAL STATUS CLASSIFICATION (IN PERCENTAGE)

		Recognized Emotion by the Software						
		An	Di	Fe	Ha	Ne	Sa	Su
Requested Emotions	An	0,00	1,42	1,32	0,00	95,37	0,07	1,82
	Di	0,00	27,42	1,90	0,00	66,59	0,50	3,59
	Fe	0,00	19,43	33,23	0,00	37,68	2,52	7,14
	Ha	0,03	2,12	11,54	59,03	23,37	1,47	2,44
	Ne	0,00	1,53	17,33	0,00	76,97	0,02	4,14
	Sa	1,79	3,04	6,09	0,00	82,49	2,58	4,00
	Su	0,53	17,70	22,20	8,74	38,27	1,72	10,83

The average value was achieved by adding values of the diagonal of the table and we divided this value with the number of emotional expressions (number 7). The emotional state of Anger was not classified. It had a very low classification value of less than 0.1%. Low values were also classified for sadness (2.58%), surprise (10.83), disgust (27.42%) and fear (33.23%). The classifier was successful in only two cases: happiness (59.03%) and neutral expression (76.97%). This result is very interesting, because even after repeating 5 times the measurement (to remove any classifier error) we have always achieved the same results. There are two explanations for us:

E1: While this test method allows for real-time trouble-free face detection, it does not provide an adequate evaluation option for determining the degree of success of the emotional classification. Such a method of evaluation results in the loss of information essential for determining the degree of success of a classification, in particular due to the ambient light conditions.

E2: Every person has the emotions written in the face (typical example is UY model testing, where disgust reaches 0.43-0.45%) and those, whether to a greater or lesser extent, cause a possible classification error and also act as a measurement deviation.

As with Japanese models we are talking about instructed expressions, in order to get a clear answer, we have set the following hypothesis.

H1: Between the instructed expressions and emotional states (naturally evoked) there are differences in the degree of classification success.

We conducted an experiment to accept or reject the hypothesis.

V. MEASURES AND PROCEDURE OF EXPERIMENT 2 – DETERMINING THE DEGREE OF SUCCESS OF CLASSIFYING THE EMOTIONAL STATE WITH THE HELP OF STUDENTS

The experimental sample consisted of 10 students (to preserve consistency with the original Jaffe sample), both men and women aged 20-25. Students have given us written permission so that we can capture them in different situations, thus enabling us to classify the natural emotional state. As a result, during the summer semester of the academic year 2016/2017, students were doing their natural activities while we could capture them from different camera distances (maximum distance when the system is capable of recognizing face of student and realizing classification of the emotional state is 7.5 meters) without the students having a clue about it. So they were being monitored for example during test writing, test answering (Fig. 7), announcement of exam results etc. In this way, we were able to classify all 6 (or 7 - neutral states) emotional states several times, thus achieving the total number of 210 classified emotional states (explanation: 3 recurrent classification of 7 states with 10 students = 210 classification).



Fig. 7. Selection of student responses and classified emotional states during the test response.

We have chosen this method to be able to confront unequivocally the results obtained with the Jaffe classification results. The entire course of the experiment was under the supervision of a psychologist who subsequently determined whether or not it was actually a classified emotional state. Table V shows the desired emotions that we received by observing student emotions and classification by the software.

TABLE V. REQUIRED EMOTIONS AND EMOTIONS RECOGNIZED BY SOFTWARE

		Recognized Emotion by the Software						
		An	Di	Fe	Ha	Ne	Sa	Su
Requested Emotions	An	96,30	5,30	6,30	1,90	0,60	0,30	1,20
	Di	0,00	89,63	6,70	5,20	10,40	0,00	2,80
	Fe	2,90	5,00	98,78	2,90	5,70	8,60	11,40
	Ha	0,00	0,00	6,10	99,60	0,00	11,10	10,10
	Ne	7,10	0,00	1,10	1,10	99,36	5,10	3,10
	Sa	0,00	2,00	1,10	1,40	10,00	81,25	0,80
	Su	3,00	3,40	8,10	1,10	1,50	0,10	97,12

The overall average classification score was 94.58%. We determined it as follows: We added all the percentage values of the results for the individual expressions on the diagonal in table V and divided the number of emotional expressions (number 7).

VI. RESULTS AND DISCUSSION – EXPERIMENT 2

As we can see from Table V, we have achieved completely different results than in the first experiment, where we subjected the classification of the emotional state of the subjects' photographs from the Jaffe database to the instructed expressions. Experiment number

2 has shown that the software has been able to classify the faces of students under different lighting conditions (a total of 3 recurring classification) and therefore, if real-time face detection conditions are appropriate, the software can seamlessly classify the subject's emotions as well. At the same time, we can accept the H1 hypothesis, because there is a demonstrable difference in classification between the instructed expressions and emotional states that are naturally induced.

As it has already been highlighted in the previous sections of our contribution, to verify the percentage of success rate of the proposed automatic emotional classification system a dataset of instructed expressions is often used. The question, however, is how really these instructed expressions are similar to the emotions of human that are expressed in real life (to the best knowledge of the authors, this assumption has never been tested). For this reason, Abramson [1] as well as we in this study examined the faces of the students in real situations. Our aim was classification and evaluation of the emotional state of fear and anger in real situations, and following, to compare them with the instructed expressions from dataset. From the results we can see the significant deviations between the requested (instructed) expressions and the expressions that were captured in the real situation and subsequently classified. In the case of classification of emotional states in real situations, it has been not only detected the faces of the students but also their emotional states, being classified with a high percentage of success. These results suggest that there are significant differences in classification between the instructed and emotional states in real time situation. Therefore, we propose searching for other (more sophisticated) options to determine the percentage of success rate of systems for recognition and classification of the emotional state and not to rely only on the traditional methods of comparing the instructed expressions. Our findings also point to the fact that in real life the classification of the emotional state is strongly dependent on the information that the human individual processes from the surrounding environment. From the results, it is clear that the most striking qualitative difference between the instructed and real emotional states is in the case of emotional state - anger. An emotional state of anger can be detected in a real facial situation with a 96.30% of success rate, while in the case of instructed expressions it obtained a 0% of success rate in classification from database Jaffe. Similarly, it was the case also of other classified emotions. This novel finding suggests that instructed stimuli (anger, disgust, fear, sadness or surprise) can convey substantial ambiguity in classification process from database pictures.

VII. CONCLUSION

At present, recognition problematic (detection and extraction) is one of the fundamental areas of scientific research in the intelligent systems. The use is really wide - from access security (laptops, controlled entry to the room), through character recognition for document scanning (OCR) to virtual reality (Second Life game). However, the implementation of a recognition system, for example, in the education process, opens up a new area of pedagogical-scientific research - the classification of the emotional state of the user. Classification thus became the last (third) phase of the recognition system. Classifying a specific emotional state of the user brings us new opportunities, a new perspective on the issue, a better understanding of what is happening to the user within a particular stage of the learning process.

In the publication, we pointed out the determination of classification success degree with our proposed solution. Normally, the method of determining the classification level as described in Experiment 1 is used as a standard, as is the detection of the subject's face and the subsequent classification of the instructed face expression. However, due to the negative results from Experiment 1, we were looking for another method of determining the classification level to confirm or

reject the hypothesis. The hypothesis has not been confirmed, so we can conclude, based on the results of Experiment 2, that there really are large differences in the degree of classification between instructed (artificially induced) emotional expressions and natural (uncontrolled) emotional states. This is a positive finding for us because we intend to permanently link the system to the LMS Moodle e-learning environment and determine the emotional status of the students. In this way, we could then understand how students feel during the test period, determine what they are experiencing and look for a common solution to their problems together with them.

ACKNOWLEDGMENT

This work was supported by the IBM Slovakia [C-16-103/0001-01, Pilot project with Constantine the Philosopher University in Nitra: Modelling the behaviour of users based on data mining with support of IBM Bluemix].

REFERENCES

- [1] L. Abramson, I. Marom, R. Petranker & H. Aviezer, "Is fear in your head? A comparison of instructed and real-life expressions of emotion in the face and body," in *Emotion*, 17(3), 2017, 557-565. doi:10.1037/emo0000252
- [2] W.N.W. Ahmad, N.M. Ali, "A Study on Persuasive Technologies: The Relationship between User Emotions, Trust and Persuasion", in *International Journal of Interactive Multimedia and Artificial Intelligence*, Volume in Press, Issue in Press, 2018, 1-5. doi: 10.9781/ijimai.2018.02.010
- [3] P. Arnau-González, M. Arevalillo-Herráez & N. Ramzan, "Fusing highly dimensional energy and connectivity features to identify affective states from EEG signals," in *Neurocomputing*, 244, 2017, 81-89. doi:10.1016/j.neucom.2017.03.027
- [4] K. Bahreini, R. Nadolski & W. Westera, "Towards multimodal emotion recognition in e-learning environments," in *Interactive Learning Environments*, 24(3), 2016, 590-605. doi:10.1080/10494820.2014.908927
- [5] Y. Balasubramanian, K. Sivasankaran & S.P. Krishraj, "Forensic video solution using facial featurebased synoptic video footage record," in *IET Computer Vision*, 10(4), 2016, 315-320. doi:10.1049/iet-cvi.2015.0238
- [6] V. Bettadapura, G. Schindler, T. Ploetz & I. Essa, "Augmenting bag-of-words: Data-driven discovery of temporal and structural information for activity recognition," Paper presented at the *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2013, 2619-2626. doi:10.1109/CVPR.2013.338
- [7] A. Bhattacharya, D. Choudhury & D. Dey, "Edge-enhanced bi-dimensional empirical mode decomposition-based emotion recognition using fusion of feature set," in *Soft Computing*, 2016, 1-15. doi:10.1007/s00500-016-2395-4
- [8] I. Cohen, N. Sebe, F.G. Cozman, M.C. Cirelo & T.S. Huang, "Learning bayesian network classifiers for facial expression recognition using both labeled and unlabeled data," Paper presented at the *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1, 2003, 1/595-1/601.
- [9] D.T. Cordaro, D. Keltner, S. Tshering, D. Wangchuk & L.M. Flynn, "The voice conveys emotion in ten globalized cultures and one remote village in bhutan," in *Emotion*, 16(1), 2016, 117-128. doi:10.1037/emo0000100
- [10] C. Darwin & F. Darwin, "The expression of the emotions in man and animals. *The expression of the emotions in man and animals*, 2009, 1-401. doi:10.1017/CBO9780511694110
- [11] M.A. DiGirolamo & J.A. Russell, "The emotion seen in a face can be a methodological artifact: The process of elimination hypothesis," in *Emotion*, 17(3), 2017, 538-546. doi:10.1037/emo000024
- [12] P. Ekman, "An argument for basic emotions," in *Cognition and Emotion*, 6(3-4), 1992, 169-200. doi:10.1080/02699939208411068
- [13] P. Ekman, "What scientists who study emotion agree about," in *Perspectives on Psychological Science*, 11(1), 2016, 31-34. doi:10.1177/1745691615596992
- [14] P. Ekman & D. Cordaro, "What is meant by calling emotions basic," in *Emotion Review*, 3(4), 2011, 364-370. doi:10.1177/1754073911410740
- [15] P. Ekman & W.V. Friesen, "Felt, false, and miserable smiles," in *Journal of*

- Nonverbal Behavior*, 6(4), 1982, 238-252. doi:10.1007/BF00987191
- [16] P. Ekman & W.V. Friesen, *Unmasking the face: A guide to recognizing emotions from facial clues*. Prentice-Hall, Englewood Cliffs, NJ (1975)
- [17] N.L. Etcoff & J.J. Magee, "Categorical perception of facial expressions," in *Cognition*, 44(3), 1992, 227-240. doi:10.1016/0010-0277(92)90002-Y
- [18] B. Fasel & J. Luetttin, "Automatic facial expression analysis: A survey," in *Pattern Recognition*, 36(1), 2003, 259-275.
- [19] S. Fleck & W. Straßer, "Smart camera based monitoring system and its application to assisted living," Paper presented at the *Proceedings of the IEEE*, 96(10), 2008, 1698-1714. doi:10.1109/JPROC.2008.928765
- [20] D. Goren & H.R. Wilson, "Quantifying facial expression recognition across viewing conditions," in *Vision Research*, 46(8-9), 2006, 1253-1262. doi:10.1016/j.visres.2005.10.028
- [21] U. Hess & P. Thibault, "Darwin and emotion expression," in *American Psychologist*, 64(2), 2009, 120-128. doi:10.1037/a0013386
- [22] M.S. Hossain & G. Muhammad, "Audio-visual emotion recognition using multi-directional regression and ridgelet transform," in *Journal on Multimodal User Interfaces*, 10(4), 2016, 325-333. doi:10.1007/s12193-015-0207-2
- [23] M.A. Jarwar & I. Chong, "Exploiting IoT services by integrating emotion recognition in web of objects," Paper presented at the *International Conference on Information Networking*, 2017, 54-56. doi:10.1109/ICOIN.2017.7899474
- [24] D. Jayatilake, T. Isezaki, Y. Teramoto, K. Eguchi & K. Suzuki, "Robot assisted physiotherapy to support rehabilitation of facial paralysis," Paper presented at the *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 22(3), 2014, 644-653. doi:10.1109/TNSRE.2013.2279169
- [25] S.M. Lajevardi & Z.M. Hussain, "Feature selection for facial expression recognition based on optimization algorithm," Paper presented at the *Proceedings of 2009 2nd International Workshop on Nonlinear Dynamics and Synchronization*, INDS 2009, 2009, 182-185.
- [26] R.W. Levenson, "Basic emotion questions," in *Emotion Review*, 3(4), 2011, 379-386. doi:10.1177/1754073911410743
- [27] L. Likforman-Sulem, A. Esposito, M. Faundez-Zanuy, S. Clemençon & G. Cordasco, "EMOTHAW: A novel database for emotional state recognition from handwriting and drawing," Paper presented at the *IEEE Transactions on Human-Machine Systems*, 47(2), 2017, 273-284. doi:10.1109/THMS.2016.2635441
- [28] K. Matsuno, C. Lee, S. Kimura & S. Tsuji, "Automatic recognition of human facial expressions," Paper presented at the *IEEE International Conference on Computer Vision*, 1995, 352-359.
- [29] J.D. Mayer, P. Salovey, D.R. Caruso & G. Sitarenios, "Emotional intelligence as a standard intelligence," in *Emotion*, 1(3), 2001, 232-242. doi:10.1037/1528-3542.1.3.232
- [30] D. McDuff, R. El Kaliouby & R.W. Picard, "Crowdsourcing facial responses to online videos," Paper presented at the *IEEE Transactions on Affective Computing*, 3(4), 2012, 456-468. doi:10.1109/T-AFFC.2012.19
- [31] A. Mehrabian, "Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in temperament," in *Current Psychology*, 14(4), 1996, 261-292.
- [32] N. Mehta & S. Jadhav, "Facial emotion recognition using log gabor filter and PCA," Paper presented at the *Proceedings - 2nd International Conference on Computing, Communication, Control and Automation, ICCUBEA 2016*, 2017, doi:10.1109/ICCUBEA.2016.7860054
- [33] G. Muhammad & M.F. Alhamid, "User emotion recognition from a larger pool of social network data using active learning," in *Multimedia Tools and Applications*, 76(8), 2017, 10881-10892. doi:10.1007/s11042-016-3912-2
- [34] J.C. Mundt, A.P. Vogel, D.E. Feltner & W.R. Lenderking, "Vocal acoustic biomarkers of depression severity and treatment response," in *Biological Psychiatry*, 72(7), 2012, 580-587. doi:10.1016/j.biopsych.2012.03.015
- [35] R. Plutchik, (2001). *The Nature of Emotions*. *Am. Scientist*, vol. 89, 344
- [36] N. Rathee, A. Vaish & S. Gupta, "Adaptive system to learn and recognize emotional state of mind," Paper presented at the *Proceeding - IEEE International Conference on Computing, Communication and Automation, ICCCA 2016*, 2017, 32-36. doi:10.1109/CCAA.2016.7813726
- [37] L.D. Riek & R.N.M. Watson, "The age of avatar realism: When seeing shouldn't be believing," Paper presented at the *IEEE Robotics and Automation Magazine*, 17(4), 2010, 37-42. doi:10.1109/MRA.2010.938841
- [38] J.A. Russell, "Affective space is bipolar," in *Journal of Personality and Social Psychology*, 37(3), 1979, 345-356. doi:10.1037//0022-3514.37.3.345
- [39] C. Torres-Valencia, M. Álvarez-López & Á. Orozco-Gutiérrez, "SVM-based feature selection methods for emotion recognition from multimodal data," in *Journal on Multimodal User Interfaces*, 11(1), 2017, 9-23. doi:10.1007/s12193-016-0222-y
- [40] P. Visutsak, "Template-based affine registration of autistic brain images," Paper presented at the *IJCCI 2015 - Proceedings of the 7th International Joint Conference on Computational Intelligence*, 2, 2015, 188-192.
- [41] Y. Yang, C. Fairbairn & J.F. Cohn, "Detecting depression severity from vocal prosody," Paper presented at the *IEEE Transactions on Affective Computing*, 4(2), 2013, 142-150. doi:10.1109/T-AFFC.2012.38
- [42] A.W. Young, D. Rowland, A.J. Calder, N.L. Etcoff, A. Seth & D.I. Perrett, "Facial expression megamix: Tests of dimensional and category accounts of emotion recognition," in *Facial expression recognition: Selected works of Andy Young*, 2016, 68-110. doi:10.4324/9781315715933
- [43] M. Zhang, D. Lee, A. Desai, K.D. Lillywhite & B.J. Tippetts, "Automatic facial expression recognition using evolution-constructed feature," in. Paper presented at the *10th International Symposium on Visual Computing*, ISVC 2014, 2014, 282-291
- [44] M. Magdin, F. Prikler. Real Time Facial Expression Recognition Using Webcam and SDK Affectiva, *International Journal of Interactive Multimedia and Artificial Intelligence*, (2017), <http://dx.doi.org/10.9781/ijimai.2017.11.002>



M. Magdin

He works as a professor assistant at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly implementation interactivity elements in e-learning courses, face detection and emotion recognition using a webcam. He participates in the projects aimed at the usage of new competencies in teaching and also in the projects dealing with learning in virtual environment using e-learning courses.



F. Prikler

was student at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly face detection and emotion recognition using a webcam. At present he works as programmer in the company Muhlbauer Technologies.