

Evaluating the Emotional State of a User Using a Webcam

M. Magdin¹, M. Turčáni¹, and Lukáš Hudec²

¹Department of Computer Science, Faculty of Natural Science, Constantine The Philosopher University in Nitra

²Java developer in EmbedIT

Abstract — In online learning is more difficult for teachers identify to see how individual students behave. Student's emotions like self-esteem, motivation, commitment, and others that are believed to be determinant in student's performance can not be ignored, as they are known (affective states and also learning styles) to greatly influence student's learning. The ability of the computer to evaluate the emotional state of the user is getting bigger attention. By evaluating the emotional state, there is an attempt to overcome the barrier between man and non-emotional machine. Recognition of a real time emotion in e-learning by using webcams is research area in the last decade. Improving learning through webcams and microphones offers relevant feedback based upon learner's facial expressions and verbalizations. The majority of current software does not work in real time – scans face and progressively evaluates its features. The designed software works by the use neural networks in real time which enable to apply the software into various fields of our lives and thus actively influence its quality. Validation of face emotion recognition software was annotated by using various experts. These expert findings were contrasted with the software results. An overall accuracy of our software based on the requested emotions and the recognized emotions is 78%. Online evaluation of emotions is an appropriate technology for enhancing the quality and efficacy of e-learning by including the learner's emotional states.

Keywords — Emotion Recognition, Face Recognition, Facial Expressions, Emotion Classification.

I. INTRODUCTION

TODAY, ICT are fundamental for our society. Their task is to make information accessible from every place in the world, quickly and without effort [47]. ICT are the means, which from the global point of view can contribute to the development of knowledge and skills [37],[22]. The ICT with its basic character enables us to increase the quality of educational processes [26]. Nowadays, we can't even imagine education in our information society without its electronic form [25]. E-learning is multimedia support of a learning process, combined with modern information and communication technologies [24]. During the last decade, several new technologies have been adopted by e-learning specialists for enhancing the effectiveness, efficiency and attractiveness of e-learning. Development of new adaptive techniques, as well as modernization of old-fashioned technologies, causes also fundamental changes in the development of our society. Development of adaptive techniques for the sphere of e-learning systems, which allow for personalization of the student, has been known for long [18]. However, during the last 5 years it has come to a real progress in opportunities for use of these technologies [8], [45], [27]. Wide spectrum of opportunities and open space for experimenting allows us applying our creativity on the topmost level [12], [31]. Such system of learning by means of ICT, for example in cooperation with the

opportunities provided by e-learning systems and technologies Web 2.0 has been actively used mainly in the last two decades [20],[33].

However, recent developments of ICT, specifically input devices (such as webcams) for interacting with such environments are still underexploited. Such devices firstly offer opportunities for more natural interactions with the e-learning applications [1]. Secondly, they offer better ways for gathering affective user data, as they do not interfere with the learning like questionnaires often do. This is because of their unobtrusive and continuously nature of data gathering.

Education belongs to areas where extensive data exploration is needed [7]. There are various methods of gathering data by the use of which it is possible to adapt the learning process to the learner. These methods can be divided into direct and indirect. Direct methods are those which are used by the user directly during the learning process, they are known and can be partly influenced (e.g. a questionnaire). Indirect methods represent a way how to individualize the learning style of the learner without participating in entering input data into the adaptive process. The learner does not fill a questionnaire but his activity is evaluated during the learning process and the learning style of the learner is determined according to association rules. As an example of data mining of indirect methods is the use of interactive animations through which an overview can be gained about cognitive and intellectual skills of the learner (module IES – Interactive Element Stat, [30]).

Existing methods for gathering affective user data, like psychological sensors and questionnaires, are either obtrusive or discontinuous. They can hamper learning as well as issues in its suitability for elearning [16], [41].

Previous software primarily dealt with offline emotion recognition that cause post-processing of the learner's data [1]. They have a couple of limitations that mainly restrict their application context and might impede their accuracy. The application context is restricted by the fact that such software can only manage a small set of expressions from frontal view faces without facial hair, glasses provided that there is constant illumination. Furthermore, the software requires postprocessing steps for analysing videos and images and cannot analyse extracted facial information on different time scales [39]. In addition, their accuracy might also be impeded as this software used no databases for authentic emotions. In our research we will investigate the opportunities of a webcam for continuous online and unobtrusive gathering of affective user data in an e-learning context.

Emotions are a critical component of effective learning and problem solving, especially when it comes to interacting with computer-based learning environments (CBLEs; multi-agent systems, intelligent tutoring systems, serious games [17].

II. RELATED WORK

Major component of human communication are facial expressions which constitute around 55 percent of total communicated message

[32]. We use facial expressions not only to express our emotions but also to provide important communicative cues during social interaction, such as our level of interest, our desire to take a speaking turn and continuous feedback signaling understanding of the information conveyed.

The facial expression research is an actual topic. The basic essays about expressions which are forming the current ones can be found in the 17th century. A detailed description of various expressions and facial muscle movements was provided by John Bulwer in 1649 in his book “Pathomyotomia”.

Another important work dealing with facial expression analysis was by Charles Darwin. In his work he described and assorted groups of expressions into categories according to similarities. He described deformations as well by which each expression is formed [6].

Though facial expressions obviously does not necessarily convey emotions, in the computer vision community, the term —facial expression recognition often refers to the classification of facial features into one of the six so called basic emotions: happiness, sadness, fear, disgust, surprise and anger, as introduced by Ekman. The advantage of this categorization is its universality among races and various cultures. An alternative of this category is a categorization designed by Baron-Cohen which includes more complex emotions. The group contains 412 various emotions divided into 24 groups. There are emotions such as boredom, interest, frustration [5]. The problem of the classification is a smattering knowledge about the universality towards various cultures and races [4]. Researches have shown that, in fact, various facial expressions are hard to record because while following, people are changing them subconsciously. Thus differences between real spontaneous expression and played expression are created [14].

There are two approaches by the use of which emotional state evaluation can occur. The first approach is to have native coders see the images or videotapes, and then make holistic judgments concerning the degree to which they see emotions on target faces in those images. While relatively simple and quick to perform, this technique is limited in that the coders may miss subtle facial movements, and in that the coding may be based by idiosyncratic morphological features of various faces. Furthermore, this technique does not allow for isolating exactly which features in the face are responsible for driving particular emotional expressions. The second approach is to use componential coding schemes in which trained coders use a highly regulated procedural technique to detect facial actions. For example, the Facial Action Coding System [15] is a comprehensive measurement system that uses frame-by-frame ratings of anatomically based facial features (“action units”) [2]. While lot of work on FACS has been done and also FACS is an efficient, objective method to describe facial expressions, but coding a subject’s video is a time- and labor-intensive process that must be performed frame by frame. A trained, certified FACS coder takes on average 2 hours to code 2 minutes of video. In situations where real-time feedback is desired and necessary, manual FACS coding is not a viable option [42].

Automatic facial expression recognition and emotion recognition have been researched extensively [2]. In the last decade, both mentioned approaches have been used in facial expression and emotion recognition. In their cases, we cannot speak about automatic recognition of facial expressions and emotions. Gained images have been manually compared and progressively evaluated. The work of [39] is considered to be the first automatic comparison system. They developed a system for automatic recognition of facial action units and analyzed those units using temporal models from profile-view face image sequences. For quicker way of comparison, neural networks were starting to be used. Progressively, they way of facial expression recognition has shifted to the other approach – recognition in real time. They have developed a general computational model for facial

affect inference and have implemented it as a real-time system. This approach used dynamic Bayesian networks for recognizing six classes of complex emotions. Their experimental results demonstrated that it is more efficient to assess a human’s emotion by looking at the person’s face historically over a two second window instead of just the current frame. Their system was designed to classify discrete emotional classes as opposed to the intensity of each emotion [2]. Others, such as Corzilus and Smids related to the work but their designed system had no response [10].

The problem of real time smile detection related is facial expression recognition. Sensing component company Omron [38] in 2009 has released smile measurement software. It can automatically detect and identify faces of one or more people and assign each smile a factor from 0% to 100%. Omron uses 3D face mapping technology and claim its detection rate is more than 90%.

III. AUTOMATIC SYSTEM OF EMOTION RECOGNITION (THEORETICAL ANALYSIS)

The system for automatic facial expression recognition has to deal with these problems: facial detection and location in a chaotic environment, extraction of facial features and correct classification of the expression on face [53].

Drawing upon various published works [20], [16], [48], [1], [43], main system elements from various similar system have been identified. The main elements of each system to evaluate facial expression are:

1. facial detection,
2. extraction of features,
3. expression evaluation.



Fig. 1. Database operation – testing application (own creation, used image from Jaffe database)

Other common features can be found in all works as well. The general proposition for systems with automatic emotional state evaluation of the user can be divided into six parts [21]:

1. image acquisition,
2. facial detection,

3. preprocessing,
4. feature acquisition,
5. classification,
6. final processing and correction.

The proposition can be extended by access to emotions. According to them it would be decided which kinds of features are appropriate to use. Each solution is different only in approaches they have to each part of the proposition. In the proposition, the model has been enriched by other elements in order to acquire fully generalized proposition.

The aim was to create an application with a simple GUI which on the bases of the webcam inputs would evaluate the actual emotional state of the user. From the results it would provide output in the form of the given situation result.

Software which was designed for uses webcams allows to interpret the emotional state of students during their interactions with an e-learning environment. This can trigger timely feedback based upon learner's facial expressions and verbalizations. The following emotions were observed: sadness, anger, disgust, fear, happiness, surprise, and neutral.

The primary objective when realizing the system was a progressive problem analysis, feature detection, state recognition and state evaluation. The original six-level model has been extended, edited and parts were added concerning the data preparation and basis. The resulting model has the form:

1. choice of monitored emotions,
2. choice of monitored features,
3. database acquisition,
4. image acquisition,
5. facial detection,
6. preprocessing,
7. feature approximation,
8. feature acquisition,
9. classification and correction.

IV. TEST SAMPLE DATABASE

One of the most important aspects of creating a new system to detect or recognize, is a database choice which would be used for new system testing. If only one database was used for each research, then new system testing, comparing with other top-class systems and efficiency testing would be rather trivial tasks [6].

The standard database for testing the systems is currently the FERET Face database [6], [46]. The next used database is Jaffe (Japanese Female Facial Expression). There are images of women in gray scales with checked background and with seven different expression based on Ekman's classification. The advantage of the database is that the expressions are already classified and described.

The main system testing is carried out on a custom sample of data. The advantage of the sample is that it is directly aimed to reveal and test weaknesses of the system. Images of various quality, size, with barriers on the face (hairstyle or glasses) were raised. The images were taken from webcams with various settings and in contrast to the database Jaffe, they do not have checked background. They are characterized by wider snapshot of face rotation and tilt. Thus the sample reflected the conditions the most realistically in which the system would work (bad light conditions, averted face, glasses, etc.). According to results gained by the system, other optimization would take place and more exact conditions would be evaluated for the image parameters.

V. DETAILED ANALYSIS OF THE SIX-LEVEL MODULE

This section lists six basic steps that make up the analysis of emotions:

1. Image acquisition,
2. Detection and facial rejection,
3. Preprocessing,
4. Feature Acquisition,
5. Emotion classification,
6. Final processing and correction.

A. Image Acquisition

When evaluating the emotional state:

1. static images,
 2. sequences
- can be used as input.

The static images are enough to acquire the expression. Although, the result contains less information according to which it would be possible to evaluate the expression. A common problem is for example the found out neural emotion when transforming from one expression to another. The sequences carry more information and provide better possibilities to optimize, to evaluate emotions and facial detection [21]. From the programming point of view, the use of sequences is demanding because a new problem is needed to be solved – facial observation. By combining the techniques, a compromise can be done – from sequence images faces can be acquired and then assigned to faces found in other images from the sequence [6].

B. Detection and facial rejection

For facial detection, it is possible to use various methods, and according to evaluation access they can be divided into four groups:

1. knowledge-based methods
2. feature invariant approaches
3. template matching methods
4. appearance-based methods.

The first approach, knowledge-based method, is based on a man's knowledge about a typical facial appearance. Usually, it represents the resolution and the relation among the facial features.

The second approach, feature invariant approach, is based on determining the signs and the rules which define the face when changing poses, or at bad light conditions. The face is being searched for according to these signs.

The third approach, template matching method, compares parts of image with the templates and facial patterns or each features.

The fourth approach, appearance-based method, the templates/models are created from groups of images which correctly represent the variability of the face. Thus learned models are used for detection [52].

Those methods are considered to be the best ones which have a high percentage of successful recognition even in unsuitable conditions, unchecked environment (uniform background, various levels of light conditions). They also fulfil a requirement of an image evaluation in real time (Fig. 2).

The majority of software's work with a successfulness of about 70%. It was possible to reach higher scores by removing the common imperfections of the software (see the implemented research).

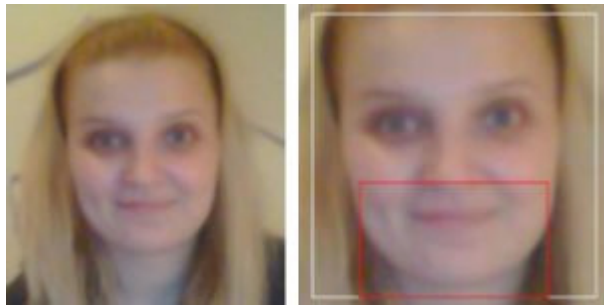


Fig. 2. Samples of custom database (on the left – an image taken by a webcam, on the right is facial and mouth detection – emotion evaluation)

The most used method is Viola-Jones, based on Haar cascade [50]. The method was used in the solution. Another method searched for connected groups of pixels which have the skin colour. According to Dadgostar and Sarrafzadeh [11] the skin colour is found in HSI model between values 0-33 for H and 15_250 for value I. The group of the pixels is processed and facial geometry is applied on them. The pixel group which fulfil the requirements would be evaluated as a human face. For facial detection, neural networks might be used, eigenfaces methods [49], support vector machine and others. For more complex view, we draw upon the work Yang, Kriegman and Ahuja which deals with description of various methods [52].

An important requirement is face rejection. All faces would be rejected where the wanted features or features needed for further evaluation are impossible to find. Also, all faces might be rejected which do not fulfil the requirement of size. Haar cascade is able to find features which have minimal size and it can be defined and calculated according to face size. That is why it is possible to exclude faces which do not fulfil the size requirement and thus there would be no wanted features. When finding more faces, the most dominant face would be left and that would mean, it has the biggest size. Next rejections would follow after feature approximation. In Fig.3 there is an image before and after rejection of incorrect faces.



Fig. 3. Before and after face rejection (own creation)

After finding the face, it is possible to use normalization methods and more accurate face designation. An example can be detection, cropped hair, removing background, brightness compensation and other.

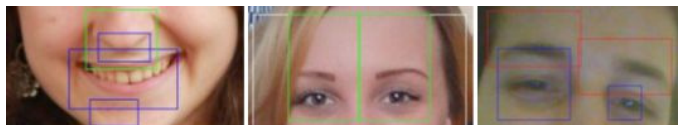


Fig. 4. Examples of mouth, eye and eyebrow detection (own creation)

C. Preprocessing

Image preprocessing often precedes face detection. The main way of preprocessing is removing noise and other processes done in order to improve quality and image appropriateness [21]. Other ways of preprocessing can be size changes, colour converting of the image into grey colour spectrum, intensity increase or applying other colour transformations. Thus, information about blushing can be lost which is lowering the accuracy of calculation [21].

D. Feature acquisition

Feature acquisition on the face is similar to facial detection on an image. Procedures and methods of facial detection can be easily modified and used for feature detection. To each feature should be approached distinctively to get the most precise feature detection and choose a method to acquire it [36], [34], [40], [3]. Precision is also important by which the feature is acquired. Optimized haar cascade adds approximated feature position but is not enough for expression evaluation and is more appropriate as approximator and a support for other methods.

When searching for lips position, it is possible to use colour transformation designed in ??? [9]. To acquire eyebrow position it is better to use method based on another principle [40].

Features can be processed according to two approaches [51]:

1. holistic,
2. local.

Holistic approach studies the face as a whole. The local approach is oriented on each features or face characteristics which can be the subjects of change [51].

The paper deals with the local approach mainly because of the easy implementation by the use of haar cascades.

According to feature detection it can be divided into:

1. approach based on geometry,
2. method based on template,
3. methods based on appearance,
4. methods based on colour and texture [36], [34], [40], [3].

E. Emotion classification

In classification, we draw on Ekman's definition of six emotions. We dare to complete it as below:

1. joy
2. sadness
3. surprise
4. fear
5. anger
6. disgust
7. neutral expression.

An important factor when creating the system to recognize expressions is a set of expressions which are needed to be recognized. In analysis a few approaches to emotion categorization have been mentioned. The observed features depend on the categorization. The system serves mainly as support in e-learning applications so the final list of expressions can be edited. Extreme expressions can be omitted such as scare, wrath, big shock and other the final categories can be divided into three groups:

1. positive
2. negative
3. neutral.

Drawing upon the hypothesis it can be possible to reduce wanted features. The nose and eyes play an important role in escalated expression and thus are not that important, in comparison to mouth and eyebrows which would be the important features.



Fig. 5. Mouth extraction (own creation)

Although, both are needed to be approached differently due to anatomic differences.



Fig. 6. Eyebrow extraction (own creation)

F. Final processing and correction

The last part of the evaluation of the emotional state of the user's is facial expression evaluation. As Sumathi, Santhanam and Mahadevi indicate to evaluate the expression, two approaches are used [44]:

1. frame based expression recognition,
2. sequence based recognition.

Frame based expression recognition uses a single image as an input. In the case of evaluation of other images, it approaches distinctively to each one and does not save information about their connection and progress. The main methods of frame based recognition are neural networks, SVM (Support Vector Machine), rule based classifiers and linear discriminant analysis.

Sequence based recognition uses image sequence as an input. Besides of images, it saves information about their progress and uses face observation. The main methods of sequence approach are recurrent neural networks, hidden Markov models, rule based classifiers [44].

According to Chibelushi another step can be result correction based on for example knowledge about common mistakes and incorrect classifications [21].

VI. PERCENTAGE SOFTWARE SUCCESSFULNESS – COMPARISON OF RESULTS ACQUIRED FROM THE SOFTWARE WITH REAL OBSERVED EMOTIONS

When verifying the software reliability, we draw on a similarly orientated research [1]. The following hypothesis has been made.

Hypothesis: There is a reliable use of data acquired by webcam and the designed software for user's emotion recognition.

Images with facial expressions were not used, but observations have been made about what emotions are evoked by the images. Each image represented a certain emotion and was provided to assistants at the department of psychology. This way has been chosen not for the users to not realize their emotions but to reach the truest result. Samples of

same size have been chosen – 10 students. These numbers are taken from all 1000 emotions (10 test persons displaying 100 emotions each) including the cases that one or more of the rates judged that the test person was unable to mimic the requested emotion correctly. Each requested emotion is separated in two rows that intersect with the recognized emotions by the software.

In the following table (Table 1) there are students' emotions in contrast to software results determined for recognition. To be able to compare the students' emotions with the acquired results, students were stealthily shot by the camera (with their permission). The video underwent an analysis. It was done by an assistant from the department of psychology and he provided detailed explanation to each analysed emotion (how did it contrast with the result achieved by the designed software).

Software that uses Bahreini, Nadolski and Westera [1] has the highest recognition rate for the neutral expression (77,2%) and the lowest recognition rate for the fear expression (50%). Note that the obtained differences between software and requested emotions are not necessarily software faults but could also indicate that participants were sometimes unable to mimic the requested emotions. The software had in particular problems to distinguish surprise from neutral. Error rates are typically between 1% and 14%. The software confused 11,3% of the neutral emotions as surprise and confused 12,5% of surprise as neutral.

The achieved results considerably differentiate from the results of Bahreini, Nadolski and Westera [1]. It is caused because of different recognition technology of each facial parts which contributes to detailed analysis and to more relevant results. The reliability of our software is 78%.

According to Table 1, the software has the highest rate of recognition for neutral expression (98,7%) and the lowest rate for sad expression (54,7%) (Table 1). Similarly, in our case there were relatively large deviations. We are inclined to the view of Bahreini, Nadolski and Westera [1] that the acquired differences between the software and the required emotions are not necessarily software mistakes. Mistake rate is in interval of 0 to 15,3%, specifically in the case of differentiating anger from disgust. What is interesting, that the software shows very different results in this cases: confused 12,5% of the neutral emotions as surprise and confused 7,2% of surprise as neutral.

TABLE I
REQUESTED EMOTIONS AND RECOGNIZED EMOTIONS BY THE SOFTWARE – THESE NUMBERS ARE TAKEN FROM ALL 1000 EMOTIONS INCLUDING 'UNABLE TO MIMIC' BY THE PARTICIPANTS (10 PARTICIPANTS DISPLAYING 100 EMOTIONS EACH).

	Recognized Emotion by the Software							Total
	Happy	Sad	Surprise	Fear	Disgust	Angry	Neutral	
Happy	94 78,3%	8 5,4%	7 2,9%	12 8,0%	3 1,3%	1 0,7%	5 3,3%	130 100%
Sad	0 0,0%	52 54,7%	7 7,4%	5 5,3%	12 12,6%	7 7,4%	12 12,6%	95 100%
Surprise	5 4,8%	2 2,1%	73 76,8%	1 1,1%	2 2,1%	5 5,1%	8 8,0%	96 100%
Fear	0 0,0%	5 5,2%	9 10,2%	45 60,0%	13 10,8%	4 2,6%	15 11,1%	91 100%
Disgust	8 8,4%	2 1,7%	1 1,2%	3 3,2%	68 80,0%	4 3,4%	3 2,1%	89 100%
Anger	0 0,0%	4 4,7%	2 2,4%	2 2,4%	13 15,3%	63 74,1%	1 1,2%	85 100%
Neutral	8 0,0%	17 0,2%	42 0,2%	12 0,4%	18 0,4%	3 0,0%	314 98,7%	414 100%
Total	113	87	145	74	130	89	362	1000

Table 1 is designed in a way it would be possible to differentiate all the seven basic emotions and easily identify the software results. Disgust has the second biggest value (80%) which is contrast to Bahreini, Nadolski and Westera. Apart from neutral, the emotion that shows best discrimination from other emotions is disgust, as disgust has a high score of 80% and is not confused with happy, sad, and

angry. The most difficult emotion is sad (54,7%) and fear (60%) and is easily confused with anger .74,1%. The difference is not a software mistake. In our opinion, the difference might have happened because the students observe the image differently and simultaneously achieve more emotions (anger and disgust). The affirmation is in accordance with Murthy [35] and Zhang [53] - the most difficult emotion to mimic accurately is fear and this emotion is processed differently from other basic facial emotions. According to various researches [35] the three emotions sad, disgust, and angry are difficult to distinguish from each other and are therefore often wrongly classified. This is confirmed by our acquired and processed results.

In following table it is visible that the analysis results of Agreements and disagreements about 1000 as students perceive various images - evaluation emotions from professor assistant at the department of psychology.

TABLE II
AGREEMENTS AND DISAGREEMENTS ABOUT 1000 AS STUDENTS PERCEIVE VARIOUS IMAGES - EVALUATION EMOTIONS FROM PROFESSOR ASSISTANT AT THE DEPARTMENT OF PSYCHOLOGY.

	Happiness	Sadness	Surprise	Fear	Disgust	Anger	Neutral	Total
Raters agree:	125	19	65	13	46	54	398	720
Able to perceive	91%	21,3%	67,8	15,4%	51,9%	61,7	89,3%	72%
Raters disagree:	21	24	25	16	21	15	61	183
Able/unable to	14,5%	25,6%	28,7%	23,9%	29,3	24,6	18,2	18,3%
Raters agree:	5	29	8	33	8	4	10	97
Unable to perceive	2,1%	35,6%	12,4%	51,3%	12,4%	5,8%	3,6%	9,7%
								100%

According to the assistant of psychology, Table 2 specifies that the students were able to perceive various images (requested emotion in 72% of the occurrences). In 183 occurrences (18,3%) there was disagreement between rater and software evaluation In 9,7% of the cases the rater stated that students were unable (on the basis of each other) to perceive requested emotions (97 times). It is interesting that students are best at perceiving neutral (89,3%) and worst at fear (15,4%). It should be noted that students are not apart. Therefore, an imitation of emotions there is possible.

For correct percentage software successfulness the results of each emotion separately and in total were re-calculated. In Table 3 the requested emotions of participants are shown (these numbers are taken by the raters from 720 emotions of the participants that were able to perceive the requested emotions) contrasted with software recognition results. Difference between Table 1 and Table 3 is that we removed both the 'unable to perceive' various images and the records from assistant psychology disagreed from the dataset.

TABLE III
REQUESTED EMOTIONS AND RECOGNIZED EMOTIONS BY THE SOFTWARE

		Recognized Emotion by the Software							Total	
		Happiness	Sadness	Surprise	Fear	Disgust	Anger	Neutral	Total	
Requested Emotions	Happiness	93	5	10	5	3	2	7	125	
		84,5%	5,3%	6,3%	1,9%	0,6%	0,3%	1,2%	100%	
	Sadness	0	12	4	1	1	0	1	19	
		0,0%	75,0%	6,7%	5,2%	10,4%	0,0%	2,8%	100%	
	Surprise	1	2	52	1	2	3	4	65	
		2,9%	5,0%	63,4%	2,9%	5,7%	8,6%	11,4%	100%	
	Fear	0	0	3	8	0	1	1	13	
		0,0%	0,0%	6,1%	72,7%	0,0%	11,1%	10,1%	100%	
	Disgust	6	0	1	1	33	2	3	46	
		7,1%	0,0%	1,1%	1,1%	82,5%	5,1%	3,1%	100%	
	Anger	0	1	0	0	3	49	1	54	
		0,0%	2,0%	1,1%	1,4%	10,0%	84,7%	0,8%	100%	
	Neutral	7	10	42	9	15	1	314	398	
		3,0%	3,4%	8,1%	1,1%	1,5%	0,1%	82,6%	100%	
	Total		107	30	112	25	57	58	331	720

In Table 3, there are results that have been achieved by removing incorrect records and repeatedly re-calculate the results achieved from evaluation when recognizing each expression of the students. The software's success rate has been set as follows: all percentage result values have been added for each expression on the diagonal in Table 3

and divided by the number of evaluated expressions (7). Thus was the success rate of 78% achieved.

VII. DISCUSSION

This study contrasted the requested emotions of participants with our designed software for the face emotion recognition. Results from assistant of psychology were used for evaluation. The best recognized emotion is anger 84,7% followed by happiness 84,5%, neutral 82,6%, disgust 82,5%, sadness 75%, fear 72,7%, and surprise 63,4%. These results are in stark contrast with the results of Bahrein, Nadolski and Westera [1]. Also, it has not been confirmed that the most intensive emotions are ranked higher than the less intensive emotions except the neutral emotion.

Anger and disgust have relatively lot of common facial signs [15], [1] and that is why high scores have been reached in Table 3 (82,5% and 84,7%).

Software precision can be verified in various ways. In the previous studies [28], [19], [1] the software has been verified on the basis of miming emotions. The respondents (note: they were not only university students, there were respondents of various age levels as well) were exposed to images with relevant expressions and their task was to mime them. The software scanned the faces and evaluated them. Consequently, results from the software were compared to required expressions.

The problem of this kind of determining emotion is the fact that the respondents might be aware about the testing and it might influence the results disconcertingly. In our case, it was not the student's task to mime the expressions from the images, but to express the emotion that the image evokes in them. Thus evoked emotions were evaluated by the designed software and, consequently, software results were confronted with the statements of an expert assistant from the departments of psychology. The comparison results are in Table 2 while in Table 3 there are results achieved by removing incorrect records and repeated calculations of results achieved from evaluation while observing each students' expressions. The software's success rate has been set as follows: all percentage result values have been added for each expression on the diagonal in Table 3 and divided by the number of evaluated expressions. Thus the success rate is on the limit of 78%. Thus the stated hypothesis can be accepted.

For this method of determining the level of success we have chosen because of that the youngsters and older adults are not equally good in miming different basic emotions (e.g., older adults are less good in miming sadness and happiness than youngsters, but older adults mimic disgust better than youngsters), it is acknowledged that the sample of test persons might influence the findings of the software accuracy [19].

VIII. CONCLUSION

In this paper we have proposed a system that automatically detects human emotions on the basis of facial expressions. We have implemented this system in LMS Moodle and is used to test students. Our interest is to determine what emotions have students in testing and help them to overcome for example stress, anger or disgust. The system is still in the testing phase.

The system works well for faces with different shapes, complexions as well as skin tones and senses basic seven emotional expressions. The success rate is on the limit of 78%. Facial expression recognition is a challenging problem in the field of image analysis and computer vision. Inclusion of emotions in human computer interface is an emerging field mainly in fields of acquiring data to support education. The solution provides us with many new opportunities. It is an assumption that the development of computationally of effective and robust solutions will

lead to increased importance of user in the process and set stage for revolutionary interactivity.

REFERENCES

- [1] K. Bahreini, R. Nadolski, & W. Westera, Towards multimodal emotion recognition in e-learning environments. *Interactive Learning Environments*, no. Ahead-of-print, 2014, p. 1-16.
- [2] J. N. Bailenson, E. D. Pontikakis, I. B. Mauss, J. J. Gross, M. E. Jabon, C. A. C. Hutcherson, . . . O. John, Real-time classification of evoked emotions using facial feature tracking and physiological responses. *International Journal of Human Computer Studies*, 66(5), 2008. 303-317.
- [3] U. Bakshi, & R. Singhal, A Survey of face detection methods and feature extraction techniques of face recognition. *International Journal of Emerging Trends & Technology in Computer Science (IJETTCS)*, 3(3), 2014, 233-237.
- [4] S. Baron-Cohen, Reading the mind in the face: A cross-cultural and developmental study. *Visual Cognition*, 3(1), 1996, 39-60.
- [5] S. Baron-Cohen, O. Golan, S. Wheelwright, & J. Hill, Mind reading: the interactive guide to emotions. 2004.
- [6] V. Bettadapura, Face Expression Recognition and Analysis: The State of the Art. arXiv: Tech Report, (4), 2012, 1-27.
- [7] I. Birzniece, P. Rudzajs, D. Kalibatiene, O. Vasilecas, & E. Rencis, Application of interactive classification system in university study course comparison. *Informatics in Education*, 14(1), 2015, 15-36.
- [8] P. Brusilovsky, I.H. Hsiao, & Y. Folajimi, QuizMap: open social student modeling and adaptive navigation support with TreeMaps. *Proceedings of the 6th European conference on Technology enhanced learning: towards ubiquitous learning (EC-TEL'11)*, Carlos Delgado Kloos, Denis Gillet, Raquel M. Crespo Garcia, Fridolin Wild, and Martin Wolpers (Eds.). Springer-Verlag, Berlin, Heidelberg, 2011, 71-82.
- [9] U. Canzlerm & T. Dziurzyk, T. Extraction of Non Manual Features for Video based Sign Language Recognition. *Proceedings of IAPR Workshop*, 2002, 318-321.
- [10] M. Corzilus, & M. Smids, *A Real-time Emotion Mirror*, 2006. Available: http://www.inf.unideb.hu/~sajolevente/papers2/emotion/emotion_mirror.pdf
- [11] F. Dadgostar, & A. Sarrafzadeh, An adaptive real-time skin detector based on hue thresholding: A comparison on two motion tracking methods. *Pattern Recognition Letters*, 27(12), 2006, 1342-1352.
- [12] K.E. DeLeeuw, R.E. Mayer, & B. Giesbrecht, How does text affect the processing of diagrams in multimedia learning. *Proceedings of the 6th international conference on Diagrammatic representation and inference (Diagrams'10)*, Ashok K. Goel, Mateja Jamnik, and N. Hari Narayanan (Eds.). Springer-Verlag, Berlin, Heidelberg, 2010, 304-306.
- [13] M. Drlik, & J. Skalka, Virtual Faculty Development Using Top-down Implementation Strategy and Adapted EES Model. *Proceedings of the World Conference on Educational Technology Research*. East Univ, Nicosia, Cyprus, Paper presented at the Procedia - Social and Behavioral Sciences, 28, 2011, 616-621.
- [14] P. Ekman & E. L. Rosenberg, *What the face reveals: basic and applied studies of spontaneous expression using the facial action coding system (FACS)*, Illustrated Edition, Oxford University Press, 1996, 3-559.
- [15] P. Ekman, & W. V. Friesen, Constants across cultures in the face and emotion. *Journal of Personality and Social Psychology*, 17(2), 1971, 124-129.
- [16] M. Feidakis, T. Daradoumis, & S. Caballe, Emotion Measurement in Intelligent Tutoring Systems: What, When and How to Measure. *Third International Conference on Intelligent Networking and Collaborative Systems*, 2011, 807-812.
- [17] J. M. Harley, F. Bouchet, M. S. Hussain, R. Azevedo, & R. Calvo, A multi-componential analysis of emotions during complex learning with an intelligent multi-agent system. *Computers in Human Behavior*, 48, 2015, 615-625.
- [18] I. -. Hsiao, P. Brusilovsky, M. Yudelson & A.Ortigosa, The value of adaptive link annotation in e-learning: A study of a portal-based approach. Paper presented at the HT'10 - *Proceedings of the 21st ACM Conference on Hypertext and Hypermedia*, 2010, 223-227.
- [19] I. Huhnel, M. Fölster, K. Werheid & U. Hess, Empathic reactions of younger and older adults: No age related decline in affective responding. *Journal of Experimental Social Psychology*, 50, 2014, 136-143.
- [20] Y. Chen, R. Hwang, & C. Wang, Development and evaluation of a web 2.0 annotation system as a learning tool in an e-learning environment. *Computers and Education*, 58(4), 2011, p. 1094-1105.
- [21] CC. Chibelushi, & F. Bourel, Facial expression recognition: a brief tutorial overview, CVonline: On-Line Compendium of Computer Vision, *Proceedings of the 13th Annual ACM International Conference on Multimedia*, 5, 2002, 669-676.
- [22] K. Jakobs, (2009). ICT standardisation in china, the EU, and the US. Paper presented at the *International Telecommunication Union - Proceedings of the 2009 ITU-T Kaleidoscope Academic Conference: Innovations for Digital Inclusion, K-IDI 2009, 2009*, 177-182.
- [23] Š. Koprdá, P. Brečka, & M. Maroš, Project and realization of WIFI net. *Proceedings of the 2009 conference Trends in education: Information Technologies and technical education*, Olomouc, Votobia, 2009, 301-304.
- [24] Š. Koprdá, M. Munk, D. Klocoková, & J. Záhorec, Základy práce s PC a v prostředí operačního systému MS Windows. In *Elearning 2006*. Hradec Králové: Gaudeamus, 2006, 239-249.
- [25] K. Kostolányová, E-learning Form of Adaptive Instruction. *Proceedings of the 9th International Scientific Conference on Distance Learning in Applied Informatics (DIVAI 2012)*. Nitra: Faculty of Natural Sciences Constantine the Philosopher University in Nitra, 2012, 193-201.
- [26] K. Kostolányová, O. Takács, & J. Šarmanová, Automated Assessment of Students' Knowledge in Adaptive Instruction. *Information and Communication Technology in Education (ICTE 2012)*. Univ Ostrave, Pedagog. Fak., Roznov pod Radhostem, Czech Republic, 2012, 119-130.
- [27] K. Kostolányová, O. Takács, & J. Šarmanová, Adaptive Education Process Modeling. *Proceedings of the 10th International Conference on Efficiency and Responsibility in Education (ERIE 2013)*, Prague 6th - 7th June 2013. Prague : Czech University of Life sciences, 2013, 300-308.
- [28] E. Krahmer, & M., Swerts, Audiovisual Expression of Emotions in Communication. *Philips Research Book Series*. Springer Netherlands, 12, 2011, 85-106.
- [29] H.-T. Lin, Ch.-H. Wang,, CH.-F. Lin, & S.-M. Yuan, Annotating Learning Materials on Moodle LMS. *Proceedings of the 2009 International Conference on Computer Technology and Development, Volume 02 (ICCTD '09)*, IEEE Computer Society, Washington, DC, USA, 2009, 455-459.
- [30] M. Magdin, M. Cápaj, New Module for Statistical Evaluation in LMS Moodle. *Proceedings of the 11th European Conference on e-Learning, ECEL 2012 : 26-27 October, Groningen, The Netherlands. - Groningen : Academic Publishing International Limited*, 2012. 304-310.
- [31] R. E. Mayer, Applying the science of learning to medical education. *Medical Education*, 44(6), 2010, 543-549.
- [32] A. Mehrabian, Communication without words. *Psychology Today*, 2(4), 1968, 53-56.
- [33] P. Miller, Web 2.0: Building the New Library. *Ariadne*, Issue 45, 2005, Available: <http://www.ariadne.ac.uk/issue45/miller/>
- [34] J. L. Moreira, A. Braun, & S. R. Musse, Eyes and eyebrows detection for performance driven animation. Paper presented at the *Proceedings - 23rd SIBGRAP Conference on Graphics, Patterns and Images, SIBGRAP*, 2010, 17-24.
- [35] G. R. S. Murthy, & R. S. Jadon, Effectiveness of Eigenspaces for facial expression recognition. *International Journal of Computer Theory and Engineering*, 1 (5), 2009, 638-642.
- [36] M. Nixon, & A. S. Aguado, *Feature Extraction and Image Processing*, 2nd ed. Academic Press, 2007.
- [37] A. Ollo-López, & M. E. Aramendia-Muneta, ICT impact on competitiveness, innovation and environment. *Telematics and Informatics*, 29 (2), 2012, 204-210.
- [38] Omron. *OKAO Vision*, 2009, Available: http://www.omron.com/r_d/coretech/vision/okao.html.
- [39] M. Pantic, N. Sebe, J. F. Cohn, & T. Huang, .Affective Multimodal Human computer interaction, 2005
- [40] U. Saeed, & J. C. Dugelay, Combining edge detection and region segmentation for lip contour extraction. *Proceedings of the 6th international conference on Articulated motion and deformable objects (AMDO'10)*, Francisco J. Perales and Robert B. Fisher (Eds.). Springer-Verlag, Berlin, Heidelberg, 2010, 11-20.
- [41] A. Sarrafzadeh, S. Alexander, F. Dadgostar, C. Fan, & A. Bigdeli, How

- do you know that I don't understand? A look at the future of intelligent tutoring systems. *Computers in Human Behavior*, 24(4), 2008, 1342-1363.
- [42] S. Srivastava, Real Time Facial Expression Recognition Using A Novel Method. *International Journal of Multimedia & Its Applications (IJMA)*, 4(2), 2012, 49-57.
- [43] M. Suk, & B. Prabhakaran, Real-time facial expression recognition on smartphones. Paper presented at the Proceedings - 2015 IEEE Winter Conference on Applications of Computer Vision, WACV 2015, 2015, 1054-1059.
- [44] C.P. Sumathi, T. Santhanam, & M. Mahadevi, Automatic Facial Expression Analysis A survey. *International Journal of Computer Science & Engineering Survey (IJCSES)*, 3(6), 2012, 47-59.
- [45] J. Sun, Y. M. Chai, Guo, & H. F. Li, On-line adaptive principal component extraction algorithms using iteration approach. *Signal Processing*, 92(4), 2012, 1044-1068.
- [46] The FERET Database [online]. Available: <http://www.itl.nist.gov/iad/humanid/feret/>
- [47] V. T. Tinio, ICT in Education. Presented By UNDP for the Benefit of Participants To The World Summit On The Information Society. UNDP's Regional Project, the Asia-Pacific Development Information Program (APDIP), In Association with the Secretariat of ASEAN, 2009, Available: <http://www.apdip.net/publications/iespprimers/eprimer-edu.pdf>
- [48] M. Trojahn, F. Arndt, M. Weinmann, & F. Ortmeier, Emotion recognition through keystroke dynamics on touchscreen keyboards. Paper presented at the ICEIS 2013 - Proceedings of the 15th International Conference on Enterprise Information Systems, 3, 2013, 31-37.
- [49] M. Turk, & A. Pentland, Eigenfaces for recognition. *Journal of Cognitive Neuroscience*, 3(1), 1991, 71-86.
- [50] P. Viola, & M. J. Jones, Robust real-time object detection. *International Journal of Computer Vision*, 57(2), 2004, 137-154.
- [51] J. Whitehill, G. Littlewort, I. Fasel, M. Bartlett, & J. Movellan, Toward Practical Smile Detection. *Pattern Analysis and Machine Intelligence*, 31(11), 2009.
- [52] M. Yang, D. J. Kriegman, & N. Ahuja, Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(1), 2002, 34-58.
- [53] Z. Zhang, Feature-Based Facial Expression Recognition: Sensitivity Analysis and Experiment with a Multi-Layer Perceptron. *International Journal of Pattern Recognition Artificial Intelligence*, 13 (6), 1999, 893-911.



Martin Magdin 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.



Milan Turčáni is head of Department of Computer Science and works as a professor at the Department of Computer Science. He deals with the theory of teaching informatics subjects, mainly implementation e-learning in learning process. 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. He is supervisor of important projects, with a focus to the area of adaptive hypermedia systems.



Lukáš Hudec was a student at the Department of Computer Science. In present is Java developer in company EmbedIT. As a student deals with facial expression recognition. At present, it continues as a consultant for the development and improvement of software for recognition of emotions students through webcams.