Pers Ubiquit Comput (2013) 17:43–52
DOI 10.1007/s00779-011-0477-y

ORIGINAL ARTICLE

Measuring instant emotions based on facial expressions during computer-based assessment

Vasileios Terzis · Christos N. Moridis · Anastasios A. Economides

Received: 14 February 2011 / Accepted: 17 August 2011 / Published online: 13 October 2011

Abstract Emotions are very important during learning and assessment procedures. However, measuring emotions is a very demanding task. Several tools have been developed and used for this purpose. In this paper, the efficiency of the FaceReader during a computer-based assessment (CBA) was evaluated. Instant measurements of the FaceReader were compared with the researchers' estimations regarding students' emotions. The observations took place in a properly designed room in real time. Statistical analysis showed that there are some differences between FaceReader's and researchers' estimations regarding Disgusted and Angry emotions. Results showed that FaceReader is capable of measuring emotions with an efficacy of over 87% during a CBA and that it could be successfully integrated into a computer-aided learning system for the purpose of emotion recognition. Moreover, this study provides useful results for the emotional states of students during CBA and learning procedures. This is actually the first time that student's instant emotions were measured during a CBA, based on their facial expressions. Results showed that most of the time students were experiencing Neutral, Angry, and Sad emotions. Furthermore, gender analysis highlights differences between genders' instant emotions.

Keywords FaceReader · e-Learning · Computer-based assessment · Emotion recognition

1 Introduction

Measuring emotions could be crucial in fields as varied as psychology, sociology, marketing, information technology, and e-learning. Consequently, several researchers have developed their own instruments to assess emotions [1]. The core channels/methods for measuring emotions are the following [2]: (1) questionnaire, (2) personal preference information, (3) speech recognition, (4) physiological data, and (5) facial expressions. Although this paper evaluates and uses facial expressions method, the following paragraphs briefly highlight some main points of the aforementioned emotion recognition methods.

Many researchers have used static methods such as questionnaires and dialogue boxes, in order to infer a user's emotions. These methods are easy to administer but have been criticized for being static and thus not able to recognize changes in affective states. Moreover, Oatley recognized that self-reporting of emotions simplifies the recognition problem [3]. However, Dieterich, Malinowski, Kühme, and Schneider-Hufschmidt stated that this approach transfers one of the hardest problems in adaptive affective interfaces from the computer to the user [4]. Thus, another advantage of the questionnaire is that it provides feedback from the user's point of view and not an outsider's [1]. Questionnaires can be used to infer users' emotions, either stand-alone or assisting another affect recognition method. On the other hand, the way questions are framed and demonstrated [5], the order in which questions are asked and the terminology employed in questions are all known to affect the subject's responses

V. Terzis (✉) · C. N. Moridis · A. A. Economides
Information Systems Department, University of Macedonia, Egnatia Street 156, Thessaloniki 54006, Hellas, Greece
e-mail: bterzis@otenet.gr

C. N. Moridis
e-mail: papaphilips@gmail.com

A. A. Economides
e-mail: economid@uom.gr

[6, 7]. Similarly, there is evidence that judgments on rating scales are non-linear and that subjects hesitate to use the extreme ends of a rating scale [8]. Hence, when using verbal scales, one should make sure that the terminology employed and the context in which it is to be presented, really reflect the subjective significance of the subject population [9].

Emotional recognition frameworks using personal preference information are based on the assumption that people do not necessarily recognize emotions just by signals seen or heard; they also use a high level of knowledge and reason, to be able to process the goals, situations, and preferences of the user. A person's emotions could be predictable if their goals and perception of relevant events were known [10]. Implemented in a computational model, this can be achieved by using agents, artificial intelligence techniques, reasoning on goals, situations, and preferences [11]. For example, if the system can reason about the reactions of a user from the input that the system receives, (assumption made derived from the time of day, speed of reading, provided personal information, etc.) appropriate content could be displayed in a way adapted for the emotion or the mood of the user.

The modulation of voice intonation is one (of the) main channel(s) of human emotional expression [12]. Certain emotional states, such as anger, fear, or joy, may produce physiological reactions [13], such as an increase in cardiac vibrations and more rapid breathing. These in turn have quite mechanical and thus predictable effects on speech, particularly on pitch (fundamental frequency F0), timing, and voice quality [14]. Some researchers have investigated the existence of reliable acoustic correlates of emotion in the acoustic characteristics of the signal [12, 15]. Their results agree on the speech correlates that are derived from physiological constraints and correspond with broad classes of basic emotions, but disagree and are unclear concerning the differences between the acoustic correlates of fear and surprise or boredom and sadness. This is perhaps explained by the fact that fear produces similar physiological reactions to surprise, and boredom produces similar physiological reactions to sadness, and consequently, very similar physiological correlates result in very similar acoustic correlates [14]. The task of machine recognition of basic emotions in non-formal everyday speech is extremely challenging.

Another valuable channel for emotional detection derives from the measurement of physiological quantities, such as temperature or blood pressure. This is important not only for the study of physiological processes and the clinical diagnostics of various diseases, but also for the estimation of emotional states. William James was the first who proposed that patterns of physiological response could be used to recognize emotion [16]. Psychologists have been using physiological measures as identifiers of human emotions such as anger, grief, and sadness [17]. Usually, changes in emotional state are associated with physiological responses such as changes in heart rate, respiration, temperature, and perspiration [18]. The use of engineering techniques and computers in physiological instrumentation and data analysis is a new, challenging research practice, especially when referring to emotional recognition. For instance, researchers at the MIT Media laboratory have been using sensors that detect galvanic skin response (GSR), blood volume pulse, respiration rate, and electromyographical activity of muscles [19]. The emotion mouse, an example of recent advances in affective computing, measures the user's skin temperature, galvanic skin response (GSR), and heart rate and uses this data to categorize the user's emotional state [20]. It has also been suggested that facial electromyography (EMG) could be potentially useful input signals in HCI [21, 22]. Therefore, there is a need for adequate measures to associate physiological measurements with definite emotional states in order to assign them to conditions meaningful to a computer [23]. Since the physiological state is so closely associated with the affective state, an accurate model of a physiological response could enable computer interactive environments to effectively determine a user's affective state in order to guide appropriate customized interactions [24]. Nevertheless, subjective and physiological measures do not always agree, which indicate that physiological data may detect responses that users are either unconscious of or cannot recall at post-session subjective assessment [25]. Moreover, the sensors might often fail and result in missing or unfavorable data, a common problem in many multimodal scenarios, resulting in a considerable reduction in the performance of the pattern recognition system [26].

Research evidence supports the existence of a number of universally recognized facial expressions for emotion such as happiness, surprise, fear, sadness, anger, and disgust [27]. Therefore, estimating emotional experiences from objectively measured facial expressions has become an important research topic. Other facial recognition systems employ advanced video-based techniques [28] or measure the electrical activity of muscles with EMG (facial electromyography) [21].

An important issue is that many of the existing facial recognition systems rely on analyzing single facial images instead of tracking the changes in facial expressions continuously [29]. It would be more meaningful if the computerized learning environments could analyze the student's facial expressions continuously to be able to react to changes in the student's emotional state at the right time. Relative to this, Essa and Pentland made the point that the lack of temporal information is a significant limitation in many facial expression recognition systems. Consequently, methods for analyzing facial expressions in human–computer interaction, especially those concerning computer-aided learning systems, should incorporate a real-time analysis [28]. This can be achieved either by using advanced video-based techniques [28] or by measuring the electrical activity of muscles with EMG (facial electromyography) [21].

At present, different machine vision techniques using video cameras are the predominant methods in measuring facial expressions [30–32]. A notable application is the FaceReader, lately developed by Vicar Vision and Noldus Information Technology bv. The FaceReader recognizes facial expressions by distinguishing six basic emotions (happy, angry, sad, surprised, scared, disgusted, and neutral) with an accuracy of 89% [33]. The system is based on Ekman and Friesen's theory of the Facial Action Coding System (FACS) that states that basic emotions correspond with facial models [34]. Several studies have used FaceReader for different purposes [35, 36].

With regard to learning, there have been very few approaches for the purpose of affect recognition. A real-time analysis should be incorporated in human–computer interaction [2], especially concerning computer-aided learning systems. Previous studies in different fields showed that FaceReader is a reliable measuring tool [35, 36]. However, learning and self-assessment are procedures with particular characteristics.

This paper evaluated the effectiveness of FaceReader 2.0 during a computer-based assessment (CBA). Accordingly, FaceReader's efficiency was measured in comparison with 2 experts' observations. Moreover, the proportions of seven basic students' emotions were estimated during the CBA and were also compared between genders.

2 Methodology

The course was an introductory informatics course, in the Department of Economic Sciences of a Greek University. The course contains theory and practice. In the theoretical module, students have to learn general concepts of Information and Communication Technology (ICT). In the practical module, students have to learn how to use Word Processing and Internet. Computer-based assessment (CBA) includes questions from both modules.

208 students enrolled to participate in computer-based assessment. The next step was the arrangement of the appointments. Finally, 172 applicants out of the 208 attended their appointments. There were 60 males (35%) and 112 females (65%). The average age of students was 18.4 (SD = 1.01). The CBA was voluntary. CBA consists of 45 multiple choice questions, and its duration was 45 min. Each question had 4 possible answers. The sequence of questions was randomized.

The use of the CBA was very simple. Each student had to choose the right answer, and then, he/she had to push the "next" button. Each page included the question, the 4 possible answers, and the "next" button. The text was in Greek. Teachers did not offer any additional instruction in the beginning. Only a few students, who were not very comfortable with the use of the assessment and asked for help with its use, received further information and instructions. The CBA's appearance was simple, too, in order to avoid any effects of design and esthetics.

During the evaluation stage of a system, the effects of human–computer interaction (HCI) are often examined by what is called the "wizard of oz mode," where a researcher hidden behind a curtain controls the system and makes observations [37]. Accordingly, each student took the test alone in a properly designed room. The room had two spaces. There was a bulkhead between the two spaces. At the first space, there was the PC on which the CBA took place. Moreover, the camera of the FaceReader was hidden in a bookcase. Besides, it is well known that people express themselves more freely when they feel that they are on their own.

In the second space were the 2 researchers. FaceReader was connected with another PC in that space, so the researchers were able to watch the facial expressions and the emotions of the participants in real time. The two researchers were also able to observe student's actions during the test through VNC viewer software, which was presenting the student's screen on a separate window of the researchers' screen (Fig. 1). Each researcher recorded the student's emotions measured by the FaceReader and his/her estimation regarding the student's emotions at the same time, based on student's facial expressions and actions.

In a live analysis, FaceReader's output is a number of charts and files. Each emotion is expressed as a value between 0 and 1, indicating the intensity of the emotion.

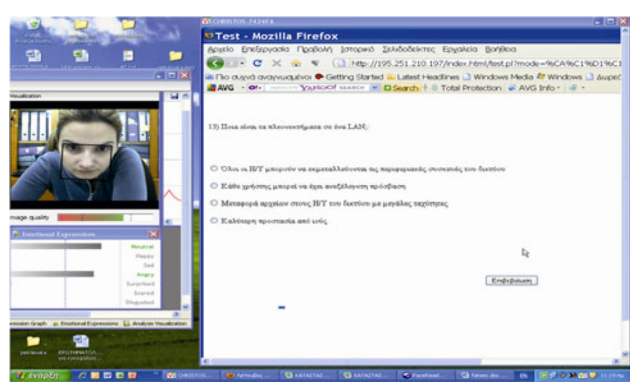

Fig. 1 Researchers' screen: FaceReader and VNC viewer (student's screen)

"0" means that the emotion is not visible in the facial expression, and "1" means that the emotion is fully present. Only emotions of value ≥0.5 were evaluated by the researchers. Changes at FaceReader's measurements in relation to student's facial expression or/and actions (observed by the researchers during the test) determined whether a FaceReader measurement was confirmed or not.

The purpose of this study has two dimensions in the context of CBA: The first is the examination of FaceReader's efficiency in measuring students' instant emotions, and the second is to provide empirical data concerning students' instant emotions.

3 Results

Firstly, it had to be examined whether the 2 researchers' estimations were statistically different. It was important to show that these estimations were free from researchers' opinions. This means that any researcher will have a good chance to show the same results if the experiment was repeated. Thus, a contingency table was created for each emotional state and overall. The 2 groups were the 2 researchers, and the outcomes were the agreement and the disagreement with the FaceReader (Table 1). The difference between the 2 researchers is not considered to be statistically significant in each emotional state and overall.

Secondly, for the 172 students, 7,416 different emotional states were recorded by the FaceReader. Table 2 shows the results for each emotional state. The second column shows confirmed records. Confirmed records are FaceReader's records that they are also confirmed by the researchers. In contrast, the third column shows all the records (Confirmed records + Not Confirmed records) of the FaceReader during CBA. Researchers and FaceReader had almost the same opinion regarding Neutral (99%) and Happy (90%) emotions. Moreover, researchers and FaceReader had high agreement for Scared (87%), Surprise (82%), and Sad (79%) emotions. However, the agreement results were lower regarding Disgusted (70%) and Angry (71%) emotions. Nevertheless, there was a high agreement overall between the emotions measured by the FaceReader and the researchers' opinions (87%).

Moreover, Table 3 shows the agreement between researchers and FaceReader on the emotional states observed in each gender. Thus, the fourth column of Table 3 presents the proportion of confirmed instances to total (confirmed and not confirmed by the researchers and FaceReader's records) FaceReader records for each emotion in each gender. Therefore, the null hypothesis was that

Table 1 Contingency table

Emotion	Researcher 1	Researcher 2	Total	Chi square	p value
Disgusted					
Agreement	170	125	295	1.03	0.31
Disagreement	80	46	126		
Surprised					
Agreement	130	85	215	0.29	0.59
Disagreement	31	16	47		
Neutral					
Agreement	1,985	1,576	3,561	1.29	0.26
Disagreement	30	16	46		
Happy					
Agreement	160	103	263	3.06	0.08
Disagreement	23	6	29		
Angry					
Agreement	694	631	1,325	2.73	0.1
Disagreement	309	236	545		
Scared					
Agreement	110	85	195	0.34	0.56
Disagreement	18	10	28		
Sad					
Agreement	281	305	586	0.28	0.59
Disagreement	70	85	155		
Total					
Agreement	3,530	2,910	6,440	2.33	0.12
Disagreement	561	415	976		

Table 2 FaceReader and researchers' agreement on various emotional states

Emotion	Confirmed records: FaceReader and researchers' agreement	Total records: confirmed and not confirmed FaceReader's records	Percentage of confirmed/total records (%)
Disgusted	295	421	70
Surprised	215	262	82
Neutral	3,561	3,607	99
Happy	263	292	90
Angry	1,325	1,870	71
Scared	195	223	87
Sad	586	741	79
Total	6,440	7,416	87

the proportions of confirmed instances to total FaceReader records for each emotion would not be statistically different in each gender. The results of the *Z* test are presented at columns 5 and 6 of Table 3. For Neutral, Happy, and Angry emotions, FaceReader showed almost the same results in both genders. Scared emotion was recognized better by FaceReader regarding males than females with statistically significant difference. Finally, Sad emotion was recognized better by FaceReader regarding females than males, also with statistically significant difference. Thus, gender differences, concerning FaceReader performance, were observed in 2 out of 7 emotional states.

Table 4 demonstrates the confirmed (column 2) and total (column 3) proportions of each instant emotion records out of overall records during CBA. *Z* test was also used to compare the proportions of the two groups determining whether they are significantly different from one another. It was expected that Neutral would be the instant emotion with the higher proportion. During the CBA, students' facial expressions stayed calm. Students changed their facial expressions instantly only if they read questions or answers that provoked them negative or positive emotions. However, the percentage of Neutral's appearances in the overall emotions, observed by the FaceReader alone, was less (48%) than the percentage of confirmed Neutral appearances (55%) in the overall confirmed emotions (observed by FaceRader and confirmed by the researchers). The co-appearance, in FaceReader's observations, of Neutral with other emotions such as Angry and Disgusted increased the total records and thus decreased the Neutral's percentage. For cases, such as this, the researchers agreed most of the times only on the Neutral observation.

On the other hand, the percentage of confirmed Disgusted and Angry emotions in the overall confirmed

Table 3 FaceReader and researchers' agreement on various emotional states observed regarding each gender

Emotion	Confirmed records: FaceReader and researchers' agreement	Total records: confirmed and not confirmed FaceReader's records	Percentage of confirmed/total records (%)	*Z* test	Significant difference
Disgusted male	131	198	66	1.544	No
Disgusted female	164	223	73		
Surprised male	82	93	88	1.743	No
Surprised female	133	169	78		
Neutral male	1,196	1,205	99	1.837	No
Neutral female	2,365	2,402	98		
Happy male	68	73	93	0.791	No
Happy female	195	219	89		
Angry male	563	779	72	1.088	No
Angry female	762	1,091	70		
Scared male	62	63	98	2.876	Yes
Scared female	133	160	83		
Sad male	200	272	74	2.736	Yes
Sad female	386	469	82		
Total male	2,302	2,683	86	1.959	No
Total female	4,138	4,733	87		

Table 4 Confirmed and total records percentages for each emotion records out of overall records during CBA

Emotion	Confirmed records: FaceReader and researchers' agreement (%)	Total records: confirmed and not confirmed FaceReader's records (%)	*Z* test	Significant difference
Disgusted	4.58	5.68	2.879	Yes
Surprised	3.34	3.53	0.565	No
Neutral	55.30	48.64	7.808	Yes
Happy	4.08	3.94	0.376	No
Angry	20.57	25.22	6.461	Yes
Scared	3.03	3.01	0.019	No
Sad	9.10	9.99	1.747	No

observations was lower than it was for the overall observations of FaceReader alone. However, Surprised, Happy, Scared, and Sad were not statistically different. This indicates that FaceReader's and researchers' observations agreed concerning these emotions during the CBA. The results also showed that "negative" emotions (Angry, Sad, and Disgusted) appeared more often than positive emotions such as Happy.

Table 5 demonstrates the confirmed (column 2) and total (column 3) percentages of instant emotions for each gender. Neutral and Angry were also statistically different for both genders. However, Disgusted was statistically different only for males. This indicates that there was an agreement between the FaceReader and researchers' observations concerning females' emotions of Disgusted. Thus, concerning Happy, Scared, Surprised, and Sad emotions, FaceReader's and researchers' observations were statistically indistinguishable in both genders.

Moreover, we compared the confirmed percentages of the two genders for each emotion records out of overall records during CBA. Table 6 shows whether the differences between the two genders are statistically significant. Results indicated that males were more Disgusted and Angry than females. On the other hand, females showed significantly more times Neutral and Happy facial expressions than males. Surprised, Scared, and Sad had no significant difference between the two genders regarding confirmed records.

4 Discussion

Measuring instant emotions by using facial expressions is a well-known method. However, this knowledge and technology have not been yet extensively used in learning environments. The aim of this study was firstly to examine the effectiveness of the FaceReader during a computer-based assessment. In parallel, we demonstrated the instant emotions' percentages that came up during the CBA. In other words, we presented how the students felt instantly while taking the CBA. Furthermore, we extended our analysis to genders in order to highlight differences between them.

Results showed that FaceReader is capable of measuring emotions with an efficacy of over 87% during CBA (Fig. 2)

Table 5 Confirmed and total records' percentages for each emotion records out of overall records during CBA in each gender

Emotion	Confirmed records: FaceReader and researchers' agreement (%)	Total records: confirmed and not confirmed FaceReader's records (%)	*Z* test	Significant difference
Disgusted male	5.69	7.38	2.339	Yes
Disgusted female	3.96	4.71	1.673	No
Surprised male	3.56	3.47	0.095	No
Surprised female	3.21	3.57	0.874	No
Neutral male	51.95	44.91	4.931	Yes
Neutral female	57.15	50.75	6.01	Yes
Happy male	2.95	2.72	0.403	No
Happy female	4.71	4.63	0.128	No
Angry male	24.46	29.03	3.595	Yes
Angry female	18.41	23.05	5.337	Yes
Scared male	2.69	2.35	0.675	No
Scared female	3.21	3.38	0.387	No
Sad male	8.69	10.14	1.695	No
Sad female	9.33	9.91	0.887	No

Table 6 Statistical significance of the differences between the confirmed percentages for each emotion records out of overall confirmed records during CBA in each gender

Emotion	Male (%)	Female (%)	*Z* test	Significant difference
Disgusted	5.69	3.96	3.12	Yes
Surprised	3.56	3.21	0.677	No
Neutral	51.95	57.15	3.996	Yes
Happy	2.95	4.71	3.354	Yes
Angry	24.46	18.41	5.724	Yes
Scared	2.69	3.21	1.091	No
Sad	8.69	9.33	0.811	No

and that it could be successfully integrated into a computer-aided learning system for the purpose of emotion recognition. Specifically, FaceReader successfully recognized Surprised, Happy, Scared, and Sad emotions (Fig. 2). FaceReader was also successful for Neutral (Fig. 2).

Moreover, results indicated that FaceReader did not have significant differences regarding emotion recognition between genders, except for Sad and Scared emotions (Fig. 3). For Sad, FaceReader was more successful for females. For males, FaceReader was more effective for Scared.

Our analysis showed limitations concerning the distinction between Neutral, Angry, and Disgusted for males during CBA. Practitioners and researchers could improve the effectiveness of emotion face recognition methods to be more effective in distinguishing between Neutral, Angry, and Disgusted in the context of CBA. Specifically, Figs. 4, 5, and 6 show examples of FaceReader's limitations during CBA. As we discussed earlier, most of the times FaceReader measured simultaneously Angry and Disgusted, the researchers agreed only with the presence of an Angry emotion (Fig. 4). Some movements of jaw, mouth, and nose may have interfered with the FaceReader's accuracy.

Additionally, many times FaceReader measured an Angry emotion simultaneously with a Neutral one, but Neutral was the only emotion confirmed by the researchers (Fig. 5). This particular disagreement was expected. When participants read the questions, many of them had clouded

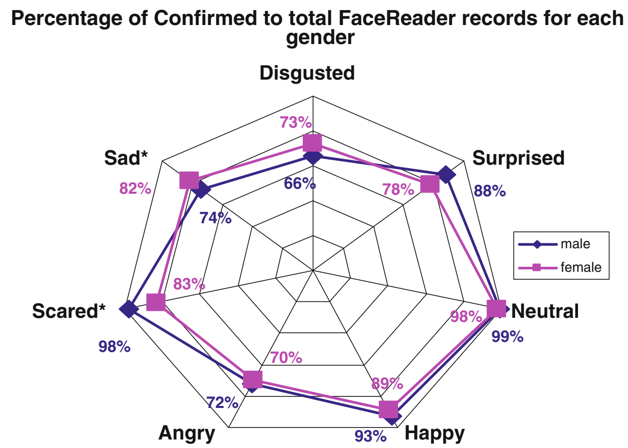

Fig. 3 FaceReader and researchers' agreement on various emotional states observed regarding each gender. *Emotions with significant differences regarding emotion recognition between genders

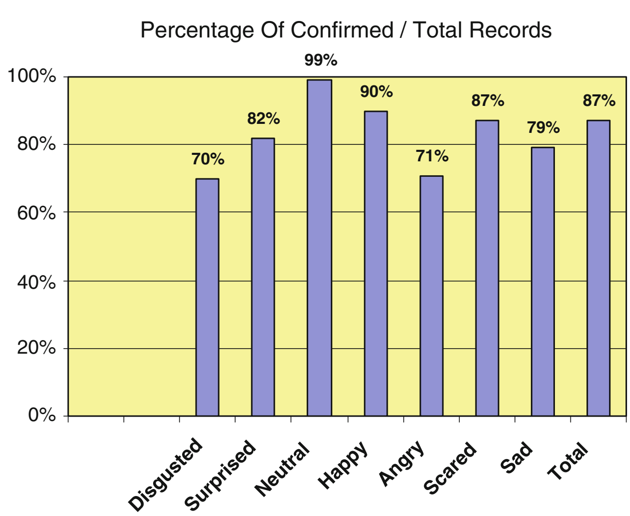

Fig. 2 FaceReader and researchers' agreement on various emotional states

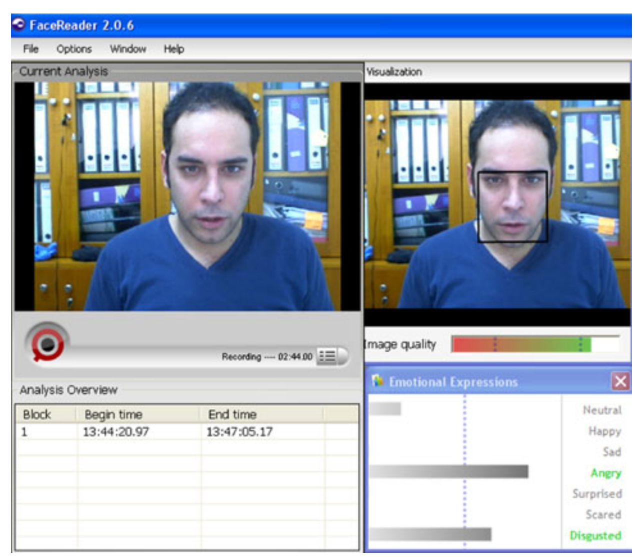

Fig. 4 Angry and Disgusted emotions co-appearance

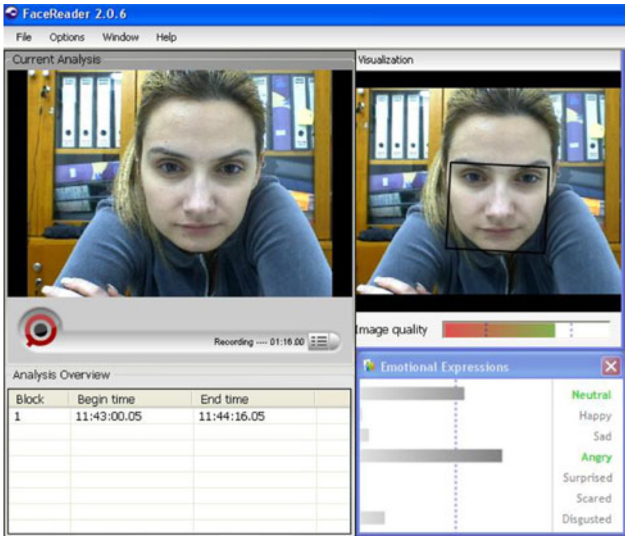

Fig. 5 Angry and Neutral emotions co-appearance

brow. People are taking this facial expression when reading something with great concentration. Zaman and Shrimpto-Smith came up to the same result [1]. This may be the reason for FaceReader measuring, so frequently, an Angry emotion at the same time with a Neutral one.

Moreover, FaceReader faced limitations with participants that wore glasses or had piercing. Other problems were caused by special characteristics of some persons like big noses, bushy brows, small eyes, or chins. Another difficulty was fringes reaching down to eyebrows (Fig. 6).

However, these limitations are being confronted. Researchers currently classify features that are located outside the modeled area of the face (e.g. hair) or features that are poorly modeled, such as wrinkles, tattoos, piercing, and birthmarks. Moreover, person identification will be added to the system [33].

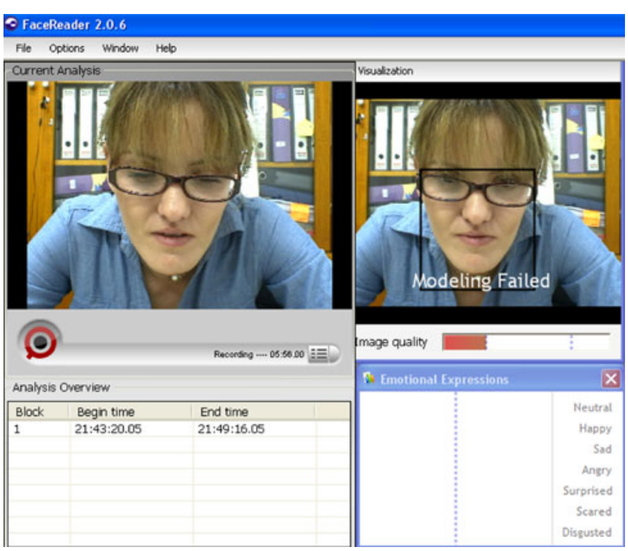

Fig. 6 Modeling failed

Our analysis also included the measurements of the different instant emotions that appeared during the CBA. Neutral was the most dominant of confirmed instant emotions with 55% (Fig. 7). As we said earlier, most of the time students' facial expressions stayed calm and they were changing their facial expressions only if they read something that changed their emotions, such as a very difficult or a very easy question. Besides Neutral, the appearance of confirmed Angry was also very large with 20% (Fig. 7). This is a very crucial result. Angry is a negative emotion that could disorganize student's effectiveness during a self-assessment or a learning procedure [38]. Another negative confirmed instant emotion with large percentage during the test was Sad (9.1%). Similarly, Sad could have negative effects on student's attention and motivation [39]. Disgusted (4.6%) and Scared (3%) are other two negative confirmed emotions that were not observed extensively (Fig. 7). However, their measurement is also important because if practitioners and researchers wish to manage student's instant emotions, they also have to take into account Disgusted and Scared [40]. During CBA, Disgusted and Scared are two negative emotions that can have an influence on student's emotional experience. Scared and Disgusted were observed most of the times after a big series of wrong answers. On the other hand, confirmed Happy (4%) had also a small percentage during the CBA (Fig. 7). This result may be justified, since a test is an anxiety provoking procedure. Happy was observed when students answered correctly a difficult question or during the last questions if they felt that they had already reached a good score.

Moreover, gender analysis revealed some useful results (Fig. 8). Surprised, Scared, and Sad had no significant difference between genders. Males presented significant larger percentages for Disgusted and Angry. This may indicate that males lose easier their temper and concentration. On

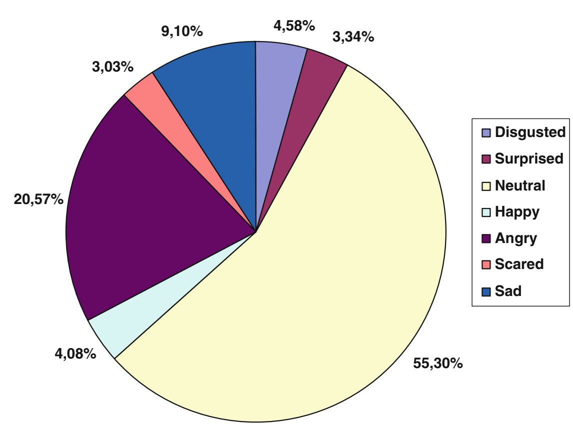

Fig. 7 Confirmed records percentages for each emotion records out of overall confirmed records during CBA

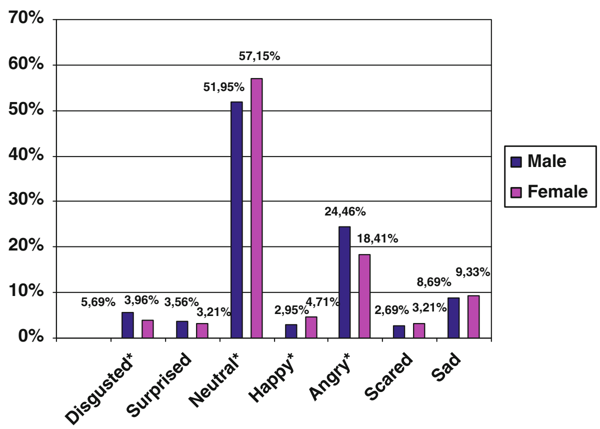

Fig. 8 Confirmed records for each emotion out of overall confirmed records during CBA for each gender. *Emotions with significant differences regarding confirmed records percentages between genders

the other side, females appeared to experience more Neutral and Happy emotions.

When the effect of negative emotions (such as Sad, Fear, or angry) is too intense, the student's performance can be seriously impaired. Frequent errors could create the expectation of more errors, thus increasing negative emotions, and leading to even more wrong answers until the student's performance collapses [41]. Positive emotions may also occasionally necessitate instruction. For instance, providing the correct answer to a hard question could induce positive emotions such as joy and enthusiasm, but also lead to loss of concentration if too much consideration is given to the elicited emotions.

Although fear was not often observed in this study, it is still an emotion that can have a detrimental effect on students' performance during a test [42, 43]. Neither was happy often observed, but positive emotions may also occasionally necessitate instruction. For instance, positive emotions can lead students to focus on the excitement and undervalue the effort required to achieve a successful result [44, 45]. On the other hand, Angry and Sad emotions were observed often enough in this study to be emotions "calling for feedback."

Regarding emotional feedback, Economides proposed an emotional feedback framework, taking as field of application the CAT (Computer Adaptive Testing) systems, in order to manage emotions [44, 46]. The emotional feedback can occur before and after the test, during the test, and before and after a student's answer to a question [46, 47]. In all these cases, emotional feedback can be provided either automatically according to the student's emotional state, either upon the student's or the teacher's request. Humor and jokes, amusing games, expressions of sympathy, reward, pleasant surprises, encouragement, acceptance, praises but also criticism are some of the possible actions that could be practiced by a testing system [44].

Finally, gender analysis revealed that females exhibited significantly higher percentages for Neutral and Happy emotions. On the other hand, males appeared to experience more Disgusted and Angry emotions. Therefore, the results of this study indicate that gender differences should be seriously taken into account when designing emotional feedback strategies for computerized tests.

5 Conclusions

An instrument like FaceReader is very crucial for the amelioration of computer-aided learning systems. Educators will have the opportunity to better recognize how their students are feeling during the learning procedures and they will also be able to give better and more effective emotional feedback in learning, self-assessment, or CAT (Computer Adaptive Testing) systems [41].

To our best knowledge, this is the first study that evaluated an emotional facial recognition instrument during CBA. Our analysis indicates some useful results. Firstly, FaceReader is efficient in measuring emotions with over 87% during CBA. Specifically, FaceReader successfully recognized Neutral, Surprised, Happy, Scared, and Sad emotions and it faces some limitations with Angry and Disgusted**.** Moreover, our research indicates that FaceReader did not have significant differences regarding emotion recognition between genders, except for Sad, in which it was more successful for females and for Scared, in which it was more effective for males.

Besides the evaluation of FaceReader, this study provides empirical data for the emotional states of students during computer-based assessments and learning procedures. Our analysis shows that Neutral (55%) was the dominant instant emotion, followed by Angry (20%) and Sad (9%). Students also experienced the other four instant emotions, that FaceReader is able to measure, at lower percentages such as Disgusted with 4.5%, Happy with 4%, Surprised with 3.3%, and Scared with 3%. Finally, gender analysis revealed that females presented significantly larger percentages for Neutral and Happy. On the other side, males appeared to experience more Disgusted and Angry emotions.

To conclude, our study provides useful and important results regarding the effectiveness of FaceReader and the students' instant emotions during CBA. These results could be useful for tutors, researchers, and practitioners.

References

1. Zaman B, Shrimpton-Smith T (2006) The FaceReader: measuring instant fun of use. In: Proceedings of the fourth Nordic

conference on human-computer interaction. ACM Press, Oslo, pp 457–460
2. Moridis C, Economides AA (2008) Towards computer-aided affective learning systems: a literature review. J Educ Comput Res 39(4):313–337
3. Oatley K (2004) The bug in the salad: the uses of emotions in computer interfaces. Interact Comput 16(4):693–696
4. Dieterich H, Malinowski U, Kühme T, Schneider-Hufschmidt M (1993) State of the art in adaptive user interfaces. In: Schneider-Hufschmidt M, Kühme T, Malinowski U (eds) Adaptive user interfaces: principles and practice. Elsevier, North Holland, pp 13–48
5. Lindgaard G, Triggs TJ (1990) Can artificial intelligence outperform real people? The potential of computerised decision aids in medical diagnosis. In: Karwowski W, Genaidy A, Asfour SS (eds) Computer aided design: applications in ergonomics and safety. Taylor & Francis, London, pp 416–422
6. Lindgaard G (1995) Human performance in fault diagnosis: can expert systems help. Interact Comput 7(3):254–272
7. Anderson NH (1982) Methods of information integration theory. Academic Press, London
8. Slovic P, Lichtenstein S (1971) Comparison of Bayesian and regression approaches to the study of information processing in judgment. Organ Behav Human Perform 6:649–744
9. Lindgaard G (2004) Adventurers versus nit-pickers on affective computing. Interact Comput 16(4):723–728
10. Ortony A, Clore GL, Collins A (1988) The cognitive structure of emotions. Cambridge University Press, Cambridge, UK
11. Conati C (2002) Probabilistic assessment of user's emotions in education games. J Appl Artif Intell 16(7–8):555–575 (special issue on managing cognition and Affect in HCI)
12. Banse R, Sherer KR (1996) Acoustic profiles in vocal emotion expression. J Pers Soc Psychol 70(3):614–636
13. Picard R (1997) Affective computing. MIT Press, Cambridge, MA
14. Oudeyer P-Y (2003) The production and recognition of emotions in speech: features and algorithms. Int J Human Comput Stud 59(1–2):157–183
15. Burkhardt F, Sendlmeier W (2000). Verification of acoustical correlates of emotional speech using formant-synthesis. In: Proceedings of the ISCA workshop on speech and emotion. Belfast, Northern Ireland
16. James W (1983) What is an emotion? In: James W (ed) Essays in psychology. Harvard University Press, Cambridge, pp 168–187 (Reprinted from Mind, 1884, 9:188–205)
17. Ekman P, Levenson RW, Friesen WV (1983) Autonomic nervous system activity distinguishes among emotions. Science 221:1208–1210
18. Frijda N (1986) The emotions. Cambridge University Press, Cambridge, UK
19. Picard R (1998) Toward agents that recognize emotion. In: Proceedings of IMAGINA. Monaco, pp 153–165
20. Ark W, Dryer D, Lu D (1999) The emotion mouse. In: Bullinger HJ, Ziegler J (eds) Human–computer interaction: ergonomics and user interfaces. Lawrence Erlbaum, London, pp 818–823
21. Partala T, Surakka V (2004) The effects of affective interventions in human-computer interaction. Interact Comput 16(2):295–309
22. Partala T, Surakka V (2003) Pupil size as an indication of affective processing. Int J Human Comput Stud 59(1–2):185–198
23. Bamidis PD, Papadelis C, Kourtidou-Papadeli C, Vivas A (2004) Affective computing in the era of contemporary neurophysiology and health informatics. Interact Comput 16(4):715–721
24. McQuiggan SW, Lee S, Lester JC (2006) Predicting user physiological response for interactive environments: an inductive approach. In: Proceedings of the second conference on artificial intelligence and interactive entertainment. Marina del Rey, pp 60–65
25. Wilson GM, Sasse MA (2004) From doing to being: getting closer to the user experience. Interact Comput 16(4):697–705
26. Kapoor A, Picard RW (2005) Multimodal affect recognition in learning environments. In: Proceedings of the 13th annual ACM international conference on multimedia. Hilton, Singapore, pp 677–682
27. Ekman P (1982) Emotion in the human face, 2nd edn. Cambridge University Press, Cambridge, MA
28. Essa IA, Pentland AP (1997) Coding, analysis, interpretation and recognition of facial expressions. IEEE Trans Pattern Anal Mach Intell 19(7):757–763
29. Partala T, Surakka V, Vanhala T (2006) Real-time estimation of emotional experiences from facial expressions. Interact Comput 18(2):208–226
30. Cohen I, Sebe N, Chen L, Garg A, Huang TS (2003) Facial expression recognition from video sequences: temporal and static modelling. Comput Vis Image Understand 91(1–2):160–187
31. Oliver N, Pentland A, Berard F (2000) LAFTER: a real-time face and lips tracker with facial expression recognition. Pattern Recogn 33(8):1369–1382
32. Smith E, Bartlett MS, Movellan J (2001) Computer recognition of facial actions: a study of co-articulation effects. In: Proceedings of the eighth annual joint symposium on neural computation
33. Den Uyl MJ, van Kuilenburg H (2005) The FaceReader: online facial expression recognition. In: Proceedings of measuring behaviour. Wageningen, The Netherlands, pp 589–590
34. Ekman P, Friesen WV (1977) Manual for the facial action coding system. Consulting Psychologists Press, Palo Alto, CA
35. Benţa K-I, Cremene M, Todica V (2009) Towards an affective aware home. In: Mokhtari M et al. (eds) ICOST 2009, LNCS 5597, pp 74–81
36. Truong KP, Neerincx MA, Van Leeuwen DA (2008) Measuring spontaneous vocal and facial emotion expressions in real world environments. In: Proceedings of MB 2008. Maastricht, The Netherlands, pp 170–171
37. Cassell J, Miller P (2007) Is it self-administration if the computer gives you encouraging looks? In: Conrad FG, Schober MF (eds) Envisioning the survey interview of the future. Wiley, New York, pp 161–178
38. Goleman D (1995) Emotional intelligence. Bantam Books, New York
39. Bower G (1992) How might emotions affect learning? In: Svenake C, Lawrence E (eds) Handbook of emotion and memory: research and theory. Erlbaum, Hillsdale, NJ
40. Economides AA, Moridis CN (2008) Adaptive self-assessment trying to reduce fear. In: Proceedings first international conference on advances in computer-human interaction. ACHI 2008, IEEE Press
41. Yusoff MZ, Du Boulay B (2009) The integration of domain independent strategies into an affective tutoring system: can students' learning gain be improved? Electron J Comput Sci Inf Technol 1(1):23–30
42. Achebe C (1982) Multi-modal counselling for examination failure in a Nigerian university: a case study. J Afr Stud 9:187–193
43. Thompson T (2004) Failure-avoidance: parenting, the achievement environment of the home and strategies for reduction". Learn Instruct 14(1):3–26
44. Economides AA (2005) Personalized feedback in CAT (Computer Adaptive Testing). WSEAS Trans Adv Eng Educ 2(3):174–181
45. Efklides A, Volet S (2005) Feelings and emotions in the learning process. Learn Instruct 15(5):1–10
46. Economides AA (2006) Emotional feedback in CAT (Computer Adaptive Testing). Int J Instruct Technol Dist Learn 3(2). Available online at: http://itdl.org/Journal/Feb_06/article02.htm
47. Economides AA (2006) Adaptive feedback characteristics in CAT (Computer Adaptive Testing). Int J Instruct Technol Dist Learn 3(8):15–26