Measuring Instant Emotions During a Self-Assessment Test: The Use of FaceReader

Vasileios Terzis
bterzis@otenet.gr

Christos N. Moridis
papaphilips@gmail.com

Anastasios A. Economides
economid@uom.gr

Information Systems Department, University of Macedonia,
156 Egnatia Avenue, Thessaloniki 54006, Greece

ABSTRACT

Emotions are very important during learning and self-assessment procedures. Measuring emotions is a very demanding task. Several tools have been developed and used for this purpose. In this paper we evaluate the efficiency of the FaceReader during a self - assessment test. We compared instant measurements of the FaceReader 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. Generally, results showed that FaceReader is capable of measuring emotions with an efficacy of over 87% during a self-assessment test, and that it could be successfully integrated into a computer-aided learning system for the purpose of affect recognition. Moreover, this study provides useful results for the emotional states of students during self-assessment tests and learning procedures.

Author Keywords

FaceReader, e-learning, self-assessment test, emotion recognition.

ACM Classification Keywords

H.5.1. Evaluation/methodology

INTRODUCTION

Measuring emotions could be crucial for many fields, such as psychology, sociology, marketing, information technology and e-learning. Consequently, several researchers have developed their own instruments to assess emotions [14]. Research evidence supports the existence of a number of universally recognized facial expressions for emotion such as happiness, surprise, fear, sadness, anger and disgust [5]. Therefore, estimating emotional experiences from objectively measured facial expressions has become an important research topic. Many facial recognition systems use single facial images instead of tracking the changes in facial expressions continuously [9]. Other facial recognition systems employ advanced video-based techniques [6] or measure the electrical activity of muscles with EMG (facial electromyography) [10].

Until now, machines using video cameras have been the predominant methods in measuring facial expressions [2, 8, and 11]. VicarVision and Noldus Information Technology launched FaceReader, a system for fully automatic facial expression analysis [13]. The FaceReader recognizes facial expressions by distinguishing six basic emotions (happy, angry, sad, surprised, scared, disgusted), plus neutral, with an accuracy of 89% [3]. Several studies have used FaceReader for different purposes [1, 12].

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 [7], especially concerning computer-aided learning systems. Previous studies in different fields showed that FaceReader is a reliable measuring tool. However, learning and self-assessment are procedures with particular characteristics. The aim of this paper was to evaluate the effectiveness of the FaceReader 2.0 during a self- assessment test. Accordingly, FaceReader's efficiency was measured in comparison to two experts' opinions.

METHODOLOGY

Participants were undergraduate students. The course was a basic IT (Information Technology) skills course and the syllabus included knowledge and techniques. The self-assessment test was optional. Students filled in an application form in order to participate in the self-assessment test. The test consisted of 45 multiple choice questions and the time limit was 45 minutes. 208 applications were collected. The next step was the arrangement of the appointments. Finally, 172 applicants out of the 208 came to their appointments.

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 self-assessments test took place. Moreover, the camera of the FaceReader was hidden in a bookcase. 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 two 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. Each researcher recorded the student's emotions measured by the FaceReader and his estimation regarding the student's emotion at the same time.

The purpose of this study was to examine the efficiency of the FaceReader during a self-assessment test. In order to accomplish this aim, the results of the FaceReader were compared to the researchers' estimations.

RESULTS

Firstly, it had to be examined whether the two researchers' estimations were statistically different. It was important to show that these estimations were free from the researchers' opinions. This means that any researcher would have a good chance to show the same results if the experiment was repeated. Thus, a contingency table was created. The 2 groups were the 2 researchers and the outcomes were the agreement and the disagreement with the FaceReader (Table 1). Pearson's Chi square was calculated in order to show the independence between the two groups. Chi squared equals 2.329 with 1 degree of freedom. The two-tailed P value equals 0.1270. Thus, the difference between the two researchers is not considered to be statistically significant.

Secondly, for the 172 students, we recorded 7416 different emotional states given by the FaceReader. Table 2 shows the results for each emotional state. 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 emotion measured by the FaceReader and the researchers' opinions.

Groups	Agreement	Disagreement	Total
Researcher 1	2910	415	3325
Researcher 2	3530	561	4091
Total	6440	976	7416

Table 1. Contingency Table.

Emotion	FaceReader and researcher agreement	Records for each emotion statement	Percentage
Disgusted	295	421	70%
Surprised	215	262	82%
Neutral	3561	3607	99%
Happy	263	292	90%
Angry	1325	1870	71%
Scared	195	223	87%
Sad	586	741	79%
Total	6440	7416	87%

Table 2. FaceReader and Researcher agreement for various emotional states.

Emotion	FaceReader and researcher agreement	Records for each emotion statement	Percentage
Disgusted male	131	198	66%
Disgusted female	164	223	73%
Surprised male	82	93	88%
Surprised female	133	169	78%
Neutral male	1196	1205	99%
Neutral female	2365	2402	98%
Happy male	68	73	93%
Happy female	195	219	89%
Angry male	563	779	72%
Angry female	762	1091	70%
Scared male	62	63	98%
Scared female	133	160	83%
Sad male	200	272	74%
Sad female	386	469	82%
Total male	2302	2683	86%
Total female	4138	4733	87%

Table 3. FaceReader and Researcher agreement for various emotional states observed regarding each gender.

Moreover, Table 3 shows the agreement between researchers and FaceReader for emotional states observed for each gender. From 172 students, 60 were male (35%) and 112 were female (65%). This sample is large enough for gender differences to be studied. For Neutral, Happy and Angry emotions, FaceReader showed almost the same results in both genders. For Surprised and Scared emotions FaceReader showed better results regarding males than females. Finally, for Disgusted and Sad emotions, FaceReader showed better results regarding females than males. Gender differences, concerning FaceReader performance, were observed in 4 out of 7 emotional states. Interpreting these differences is not part of this work. However, we plan to discuss these differences in another work in the near future.

In order to obtain the confidence interval for the agreement between researchers' opinions and FaceReader, a binomial proportion confidence interval was used (Table 4). The Adjusted Wald interval provides the best coverage for a specified interval.

DISCUSSION & CONCLUSIONS

Disgusted and Angry were the two emotions that FaceReader recognized less effectively. Examining the results revealed that Disgusted and Angry co-appeared frequently. Most of the times FaceReader measured simultaneously these two emotions, the researchers agreed only with the presence of an Angry emotion. Some movements of jaw, mouth and nose confused the FaceReader accuracy. Additionally, many times FaceReader measured an Angry emotion simultaneously with a Neutral one, but neutral was the only emotion confirmed by the researchers. This particular disagreement was expected. When participants read the questions, many of them had clouded brow. People are taking this facial expression when reading something with great concentration. Zaman and Shrimpto-Smith (2006) came up to the same result. This is the reason why FaceReader measured, so frequently, an Angry emotion at the same time with a Neutral one. Moreover, FaceReader faced problems 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 were fringes reaching down to eyebrows.

Emotion	95% confidence interval	95% conf. interv., males	95% conf. interv., females
Disgusted	78% - 85%	59% - 72%	67% - 79%
Surprised	77% - 86%	80% - 93%	72% - 84%
Neutral	98.3% - 99%	98.5% -99.6%	97.8% - 98.8%
Happy	86% - 93%	86% - 97%	84% - 93%
Angry	69% - 73%	69% - 75%	67% -72%
Scared	82% - 93%	91% - 99%	77% - 88%
Sad	76% - 82%	68% - 78%	79% - 86%
Total	86%-87,5%	84.4% - 87%	86.5% - 88%

Table 4. The overall and the genders' confidence Interval for the six emotions plus Neutral.

Hopefully, these problems may be confronted because FaceReader will be upgraded. VicarVision and Noldus Information Technology support that they classify features which are located outside the modelled area of the face (e.g. hair) or features which are poorly modelled wrinkles, tattoos, piercing and birthmarks. Moreover, they will add person identification to the system [3].

Generally, results showed that FaceReader is capable of measuring emotions with an efficacy of over 87% during a self-assessment test, and that it could be successfully integrated into a computer-aided learning system for the purpose of affect recognition.

An instrument like FaceReader is very crucial for the amelioration of computer-aided learning systems. Educators will have the opportunity to give better and more effective emotional feedbacks in learning, self-assessment or CAT (Computer Adaptive Testing) systems [4].

To conclude, to our best knowledge this is the first study that evaluated FaceReader during a self-assessment test. Besides the evaluation of FaceReader, this study provides useful results for the emotional states of students during self-assessment tests and learning procedures.

REFERENCES

1. Benţa K.-I., Cremene M., & Todica V. *Towards an Affective Aware Home***,** M. Mokhtari et al. (Eds.): ICOST 2009, LNCS 5597 (2009), 74–81.
2. Cohen, I., Sebe, N., Chen, L., Garg, A., & Huang, T. S. Facial expression recognition from video sequences: Temporal and static modelling. *Computer Vision and Image Understanding, 91,* 1-2 (2003), 160-187.
3. Den Uyl, M. J., & van Kuilenburg, H. The FaceReader: Online facial expression recognition. In *Proceedings of MB 2005*, 589-590.
4. Economides, A. A. Adaptive feedback characteristics in CAT (Computer Adaptive Testing). *International Journal of Instructional Technology & Distance Learning*, 3, 8 (2006), 15-26.
5. Ekman, P. *Emotion in the human face* (2nd ed.). Cambridge, MA: Cambridge University Press, 1982.
6. Essa, I. A., & Pentland, A. P. Coding, analysis, interpretation and recognition of facial expressions. In *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*, 7 (1997), 757-763.
7. Moridis, C. & Economides, A. A. Towards computer-aided affective learning systems: a literature review. *Journal of Educational Computing Research,* 39, 4 (2008), 313-337.
8. Oliver, N., Pentland, A., & Berard, F. LAFTER: A real-time face and lips tracker with facial expression recognition. *Pattern Recognition, 33,* 8 (2000), 1369-1382.

9. Partala, T., Surakka, V., & Vanhala, T. Real-time estimation of emotional experiences from facial expressions. *Interacting with Computers, 18*, 2 (2006), 208-226.

10. Partala, T.,& Surakka, V. The effects of affective interventions in human-computer interaction. *Interacting with Computers, 16*, 2 (2004), 295-309.

11. Smith, E., Bartlett, M. S., & Movellan, J. Computer recognition of facial actions: A study of co-articulation effects. In *Proceedings of the eighth annual joint symposium on neural computation*, 2001.

12. Truong K. P., Neerincx M. A., & Van Leeuwen D. A. Measuring spontaneous vocal and facial emotion expressions in real world environments. *Proceedings of MB 2008*, 170-171.

13. Van Kuilenburg H., Den Uyl M.J., Israel M.L. & Ivan P. *Advances in face and gesture analysis. Proceedings of MB 2008,* 371-372.

14. Zaman B. & Shrimpton-Smith T. *The FaceReader: Measuring instant fun of use*. NordiCHI 2006, 457-460.