Noldus
Information Technology

FaceReader methodology

WHAT IS FACEREADER?

FaceReader™ is a program for facial analysis. It can detect emotional expressions in the face. FaceReader has been trained to classify facial expressions in one of the following categories: happy, sad, angry, surprised, scared, disgusted and neutral. These emotional categories have been described by Ekman [1] as the basic or universal emotions. Obviously, facial expressions vary in intensity and are often a mixture of emotions. In addition, there is quite a lot of inter-personal variation.

FaceReader has been trained to classify the seven emotions mentioned above. It is not possible to add emotions to the software yourself. Please contact Noldus Information Technology if you are interested in the classification of other emotions.

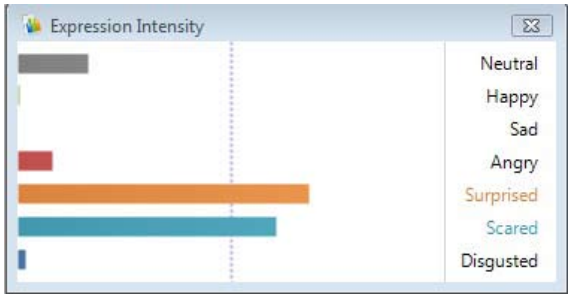

Figure 1. *Chart showing the intensity of the emotions that are visible in the face.*

In addition to facial expressions, FaceReader offers a number of extra classifications:

- Facial states: left and right eye open or closed, mouth open or closed and left and right eyebrow raised, neutral or lowered.
- FaceReader can track the orientation of the head. The head orientation is recorded in the X, Y and Z direction and is based on a 3D model of the face.
- Global gaze direction: left, forward or right. The gaze direction is derived from the position of the pupil in relation to the corners of the eyes and the orientation of the head.
- Subject characteristics: FaceReader can extract the following information from the face: the subject's gender, age and ethnicity and whether the person has a beard, moustache or glasses.

Gaze and head tracking may be used to get an indication of a test participant's interest/ engagement. If the participant is not looking in a particular direction, he/she is probably not (yet/anymore) interested in what is visible there.

FaceReader can classify facial expressions either live using a webcam, or offline, in video files or images. In a live analysis FaceReader can analyze up to 15 frames/second (depending on the computer you use). A pre-recorded video can be analyzed frame-by-frame.

HOW DOES FACEREADER WORK?

FaceReader works in three steps [2,3,4]:

1. The first step in facial expression recognition is detecting the face. FaceReader uses a unique combination of two face detection algorithms. The popular Viola-Jones algorithm [5] is used to roughly detect the presence of a face while a deformable template method [6] creates a more accurate framing containing information about the likely in-plane rotation of the face.
2. The next step is an accurate 3D-modeling of the face using an algorithmic approach based on the Active Appearance method described by Cootes and Taylor [7]. The model uses a database of annotated images. It describes 491 key points in the face and the facial texture of the face entangled by these points. The 491 key points include (A) the points that enclose the face (the part of the face that FaceReader analyzes); (B) points in the face that are easily recognizable (lips, eyebrows, nose and eyes). The texture is important because it gives extra information about the state of the face. The key points only describe the global position and the shape of the face, but do not give any information about the presence of wrinkles and the shape of the eye brows. This information is important to classify the facial expressions.
3. The actual classification of the facial expressions is done by training an artificial neural network [8]. As training material nearly 2000 manually annotated images were used

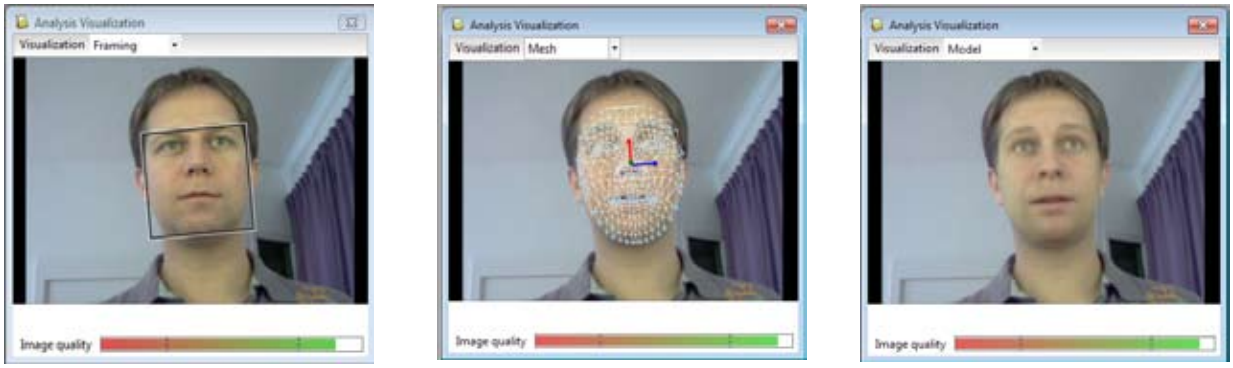

Figure 2. *Screenshots showing how FaceReader works. In the first step (left) the face is detected. A box is drawn around the face at the location where the face was found. The next step is an accurately modeling of the face (right). The model describes 491 key points in the face and the facial texture of the face entangled by these points (middle).*

There are multiple face models available in FaceReader. In addition to the general model which works well under most circumstances for most people, there are models for East Asian people, elderly and children. Before you start analyzing facial expressions, you must select the face model which best fits the faces you are going to analyze.

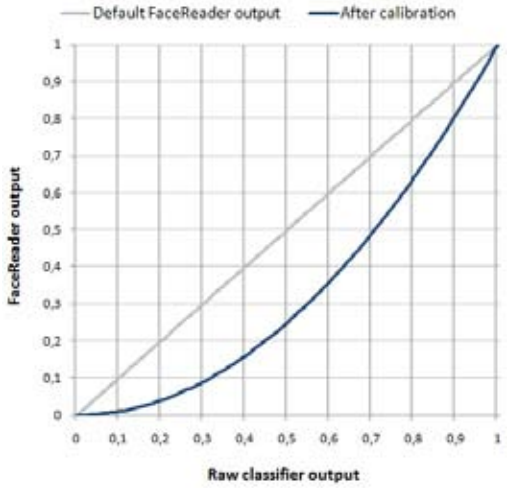

Figure 3. *An example of a possible classifier output correction for a specific facial expression using the calibration function.*

CALIBRATION

Some people look, for example, surprised or sad by nature. You can calibrate FaceReader to correct for these person-specific biases towards a certain emotion. Calibration is a fully automatic mechanism. The calibration procedure uses a set of images or videos (either calibration material you provide or the original video/images) to sample the intensities of the individual emotions of the test person. Using the distribution of these intensity samples, it applies a correction [9] resulting in an average distribution of the intensities. Consequently, the emotions are more balanced and personal biases towards a certain emotion are removed. The effect can best be illustrated by an example. For instance, a person is classified as 'angry' nearly all the time. This means that this test person should be classified as 'angry' when the classification value is, e.g., 0.7 rather than 0.5. Figure 3 shows how the classifier outputs are mapped to different values to negate the test person's bias towards 'angry'.

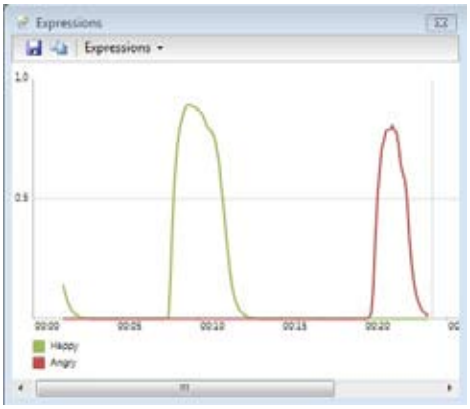

Figure 4. *An example of a chart displaying the facial expressions 'happy' and 'angry' over time.*

FACEREADER'S OUTPUT

FaceReader's output is a number of charts and log files. Each emotion is expressed as a value between 0 and 1, indicating the intensity of the emotion. '0' means that the emotion is not visible in the facial expression, '1'

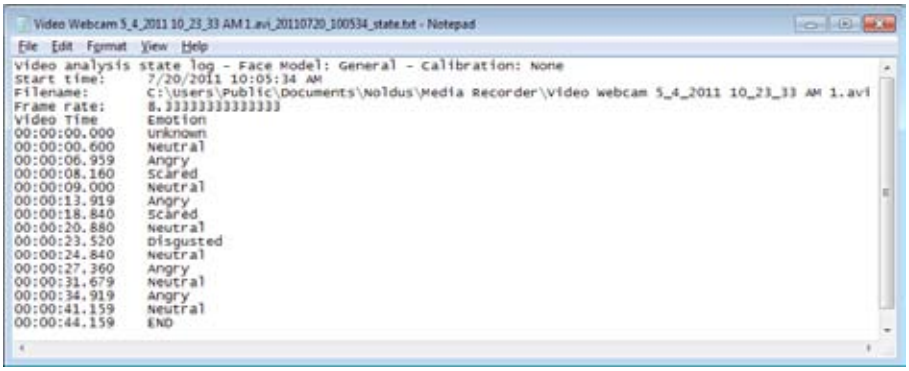

Figure 5. *Example of a State log.*

	A	B	C	D	E	F	G	H	I
1	Video analysis detailed log - Face Model: General - Calibration: None								
2	Start time:	7/20/2011 10:05							
3	Filename:	C:\Users\Public\Documents\Noldus\Media Recorder\Video Webcam 5_4_2011 10_23_33 AM 1.avi							
4	Frame rate:	8.333333333							
5	Video Time	Neutral	Happy	Sad	Angry	Surprised	Scared	Disgusted	Gaze
6	00:00.0	0.7056673	0.03295831	0.03829205	0.02418135	0.01173598	0.1281067	0.1131435	Left
7	00:00.1	0.6916715	0.03265892	0.03740995	0.02586098	0.01151441	0.1348285	0.1119588	Left
8	00:00.2	0.6623932	0.03129632	0.03744886	0.02763258	0.01161518	0.145324	0.1192957	Forward
9	00:00.4	0.6634784	0.03182124	0.03742323	0.02905909	0.01149597	0.1442428	0.1178577	Forward
10	00:00.5	0.6229243	0.03110085	0.03430537	0.03859083	0.01086978	0.1499203	0.1052296	Left
11	00:00.6	0.571656	0.0300923	0.03194452	0.04725857	0.01006697	0.1514511	0.1009451	Forward
12	00:00.7	0.5509137	0.02981828	0.02920196	0.06069282	0.009203221	0.1389731	0.1018396	Forward
13	00:00.8	0.5433869	0.02915456	0.02691063	0.07258482	0.008930645	0.1362792	0.09759995	Forward
14	00:01.0	0.5313807	0.02830554	0.02654836	0.08000304	0.008375062	0.1310891	0.09456912	Left
15	00:01.1	0.529245	0.02802264	0.02668181	0.08595741	0.00780268	0.1188672	0.09463192	Forward
16	00:01.2	0.5366116	0.02878796	0.02615617	0.08905451	0.00736994	0.1087872	0.09722186	Left
17	00:01.3	0.5367056	0.03069485	0.02525658	0.08623552	0.007467974	0.1094071	0.0947906	Left
18	00:01.4	0.5185009	0.03070518	0.02552071	0.08444178	0.007524171	0.09594008	0.1028275	Left
19	00:01.6	0.4887034	0.02469057	0.03529865	0.08509497	0.008016407	0.07098769	0.1073197	Forward
20	00:01.7	0.4351391	0.02025501	0.04233485	0.09424553	0.007867088	0.05195721	0.1244866	Forward
21	00:01.8	0.4093874	0.01922601	0.03678032	0.09750643	0.008520308	0.04157516	0.1276465	Left
22	00:01.9	0.3722185	0.0175673	0.03135795	0.1115556	0.009826086	0.03700506	0.1197161	Forward
23	00:02.0	0.3893852	0.01958246	0.02915991	0.1105299	0.008881401	0.03253517	0.1048316	Left
24	00:02.2	0.3571156	0.01879821	0.03087858	0.1325622	0.007845491	0.02720545	0.09963278	Left
25	00:02.3	0.3407884	0.01859379	0.03303112	0.1321689	0.00709053	0.02524665	0.103784	Forward
26	00:02.4	0.3363307	0.01941606	0.03296375	0.1390598	0.005973056	0.02405126	0.1145007	Left
27	00:02.5	0.4161603	0.02518262	0.03122611	0.1370987	0.005058454	0.02254134	0.1057243	Left
28	00:02.6	0.4721568	0.03344388	0.02935201	0.1227646	0.00452566	0.02234213	0.1039939	Left
29	00:02.8	0.5086641	0.04385349	0.02651583	0.1086887	0.004340693	0.02385403	0.1040817	Left
30	00:02.9	0.5498528	0.05744013	0.02272859	0.09322121	0.004373714	0.02707614	0.1071319	Left
31	00:03.0	0.5705476	0.06501897	0.02031526	0.08160476	0.004721094	0.03649638	0.1053972	Left
32	00:03.1	0.5865998	0.07192988	0.01882219	0.07082979	0.005278296	0.04484982	0.1024653	Left

Figure 6. *Example of a Detailed log.*

means that the emotion is fully present. FaceReader has been trained using intensity values annotated by human experts.

Facial expressions are often a mixture of emotions and it is very well possible that two (or even more) emotions occur simultaneously with a high intensity. The sum of the intensity values for the seven emotions at a particular point in time is, therefore, normally not equal to 1. In addition, the emotional state of the subject is estimated. The state values are an estimation of the emotional state of the subject based on the amplitude, duration and continuity shown in the recent emotional responses. Each time the emotional state changes, a record is written to the State log file (see Figure 5 for an example). The Detailed log contains all the emotional classifier outputs per time point (see Figure 6 for an example).

Both the State log and the Detailed log can be accessed real-time by other applications via an Application Programming Interface (API). This makes that FaceReader can be used for research into affective computing and the design of adaptive interfaces. In other words, FaceReader allows other programs to respond instantaneously to the emotional state of the test participant.

Instead of determining the intensities of individual emotions, it is also possible to calculate the valence. The valence indicates whether the emotional state of the subject is positive or negative. 'Happy' is the only positive emotion, 'sad', 'angry', 'scared' and 'disgusted' are considered to be negative emotions. 'Surprised' can be either positive or negative. The valence is calculated as the intensity of 'happy' minus the intensity of the negative emotion with the highest intensity. For instance, if the intensity of 'happy' is 0.8 and the intensities of 'sad', 'angry', 'scared' and 'disgusted' are 0.2; 0.0; 0.3 and 0.2, respectively, then the valence is 0.5. Valence values can be logged in the Detailed log file.

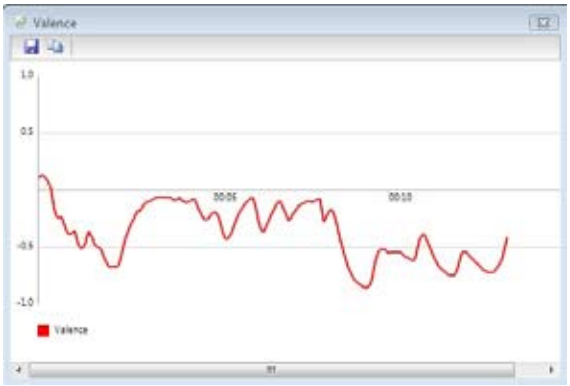

Figure 7. *Example of a Valence chart showing the valence over time.*

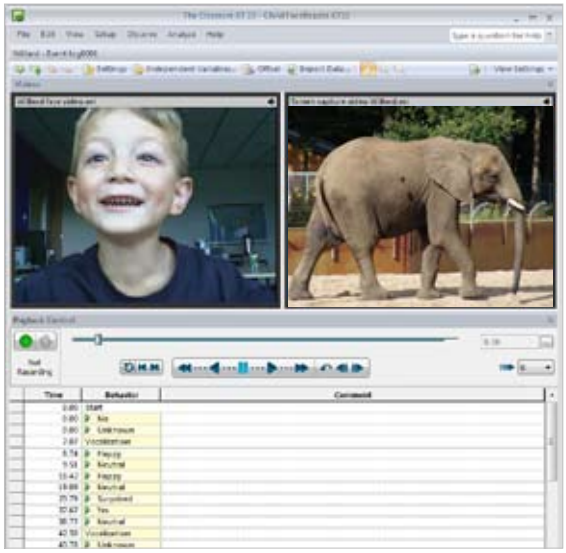

Figure 8. *FaceReader data in The Observer.*

You can save the log files as text files which you can open in most spreadsheet programs and text editors, for instance, in Excel or Notepad. You can also import FaceReader log files into The Observer® XT, our software package for collecting observational data. This makes it possible to combine the FaceReader data with your manually scored events and with data from other systems, like eye trackers or physiological data acquisition systems. The advanced data selection and analysis functionality in The Observer enables you to select, for instance, only that part of the observation when the test participant was looking at the homepage of your newly developed web site and calculate the mean classification values for each of the emotions during this period. The Observer also gives you the possibility to group data for analysis, for example compare results by gender or age, by using independent variables.

HOW WELL DOES FACEREADER WORK?

The current FaceReader version has a number of limitations:

- FaceReader is not trained to work with very young children, below the age of 3, nor with children from East Asia and South-East Asia.
- Pose, movement and rotation of the test person are limited. The test person should stand or sit and look frontally into the camera (angle < 40°). In addition, FaceReader requires strict light conditions.
- The face should not be partly hidden, for instance by very heavy facial hair or a person's hand during eating behavior. Glasses may hinder classification, especially thick and dark frames.
- FaceReader can analyze one face at a time. If there are more faces in an image these can be analyzed in different runs, provided positions of the faces do not change.
- FaceReader cannot classify facial expressions in test persons with a partial facial paralysis.

FaceReader contains an image quality bar, which gives you a good indication of how well the program is able to model the face depicted in the image.

Image quality bar.

Table 1 shows the performance results of FaceReader. Two sets of images have been analyzed in FaceReader (version 3): the Karolinksa data set [10] (980 images) and a subset of the Full Color Feret data set (1705 images) [11].

Table 1. *Performance results of FaceReader on the Color Feret data set and the Karolinska data set.*

Parameter	Color Feret (1705 images)	Karolinska Directed Emotional Faces (980 images)
Find failed (number of images)	4	0
Fit failed (number of images)	253	3
Percentages of faces found	99.8%	100%
Percentage fit succeeded of correctly framed	85.1%	100%

To validate FaceReader, its results have been compared with those of intended expressions [12]. Figure 9 shows the results of a comparison between the analysis in FaceReader and the intended expressions in images of the Radboud Faces Database [13]. The RaFD is a highly standardized set of pictures containing images of eight emotional expressions. The test persons in the images have been trained to pose a particular emotion and the images have been labeled accordingly by the researchers. Subsequently, the images have been analyzed in FaceReader. As you can see, FaceReader classifies 164 'happy' images as 'happy', five 'happy' images as 'Unrecognized', one as "Disgusted' and another 'happy' image as 'neutral', giving an accuracy of 95,9% for this emotion. For the other emotions the accuracy is lower. The overall accuracy is 90%.

In another study the results of FaceReader were compared with those of three human observers [14]. The observers appeared to vary in their ability to classify facial expressions. FaceReader performed as well as two of the three observers, the third observer scored significantly better. Terzis et al. [15] compared FaceReader with two human observers and found that the two observers

Noldus
Information Technology

FaceReader methodology

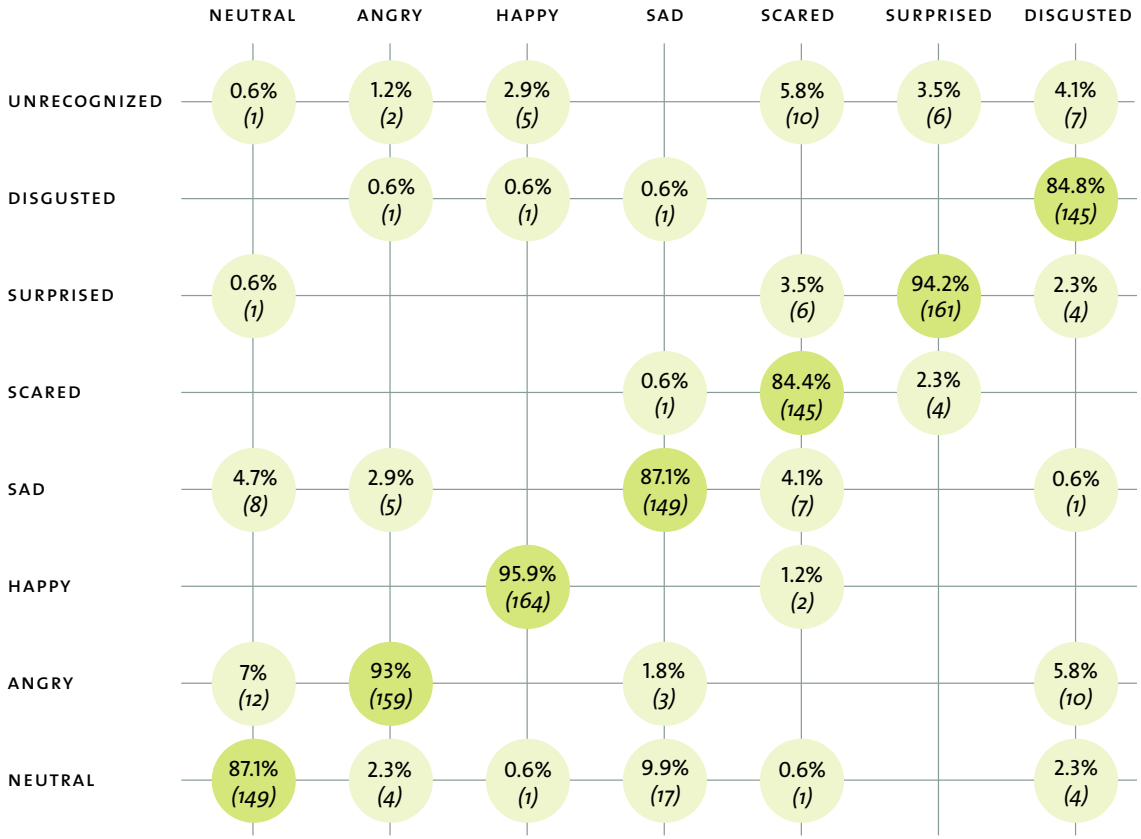

Figure 9. *Proportion of agreement between the facial expressions scored manually by the annotators of the Radboud Faces Database [13]. (horizontally) and the expressions scored by FaceReader (vertically).*

performed equally well. The results of FaceReader showed a high degree of agreement with those of the human observers, ranging from 99% for 'neutral' to 70% for 'disgusted'.

Up till now the most widely used method to code facial expressions is the Facial Action Coding System (FACS) developed by Ekman and Friesen [16]. This has become a standard for coding facial expressions. It codes the various possible facial movements ('action units'). Interpretation of the reported activated units is not included in FACS analysis, but is done in separate systems such as EMFACS [17] or FACSAID [18]. The FACS can detect very small differences in facial expressions, for example, the difference between a real (Duchenne) smile and a fake or social smile. The current version of FaceReader has not been trained to detect this difference.

FACS analysis is, however, extremely time consuming and it requires very intensive training. It is, therefore, not suitable for large data sets. FaceReader can analyze facial expressions real-time and is, thus, an interesting alternative.

Feel free to contact us or one of our local representatives for more references, clients lists, or more detailed information about FaceReader and The Observer.

WWW.NOLDUS.COM

REFERENCES

1. Ekman, P. (1970). Universal facial expressions of emotion. *California Mental Health Research Digest*, **8**, 151-158.
2. Van Kuilenburg, H.; Wiering, M; Den Uyl, M.J. (2005). A Model Based Method for Automatic Facial Expression Recognition. *Proceedings of the 16th European Conference on Machine Learning, Porto, Portugal, 2005,* pp. 194-205, Springer-Verlag GmbH.
3. Den Uyl, M.J.; Van Kuilenburg, H. (2008). The FaceReader: Online Facial Expression Recognition. *Proceedings of Measuring Behavior 2005, Wageningen, The Netherlands, August 30 - September 2, 2008,* pp. 589-590.
4. Van Kuilenburg, H.; Den Uyl, M.J.; Israël, M.L.; Ivan, P. (2008). Advances in face and gesture analysis. *Proceedings of Measuring Behavior 2008, Maastricht, The Netherlands, August 26-29, 2008*, pp. 371-372.
5. Viola, P.; Jones, M. (2001). Rapid object detection using a boosted cascade of simple features. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Kauai, HI, U.S.A., December 8-14, 2001.*
6. Sung, K.K.; Poggio, T. (1998). Example-based learning for view-based human face detection. *IEEE Transactions on Pattern Analysis and Machine Intelligence,* **20(1)**, 39-51.
7. Cootes, T.; Taylor, C. (2000). Statistical models of appearance for computer vision. *Technical report, University of Manchester, Wolfson Image Analysis Unit, Imaging Science and Biomedical Engineering*.
8. Bishop, C.M. (1995). *Neural Networks for Pattern Recognition*. Clarendon Press, Oxford.
9. Http://en.wikipedia.org/wiki/Histogram_equalization
10. Lundqvist, D.; Flykt, A.; Öhman, A. (1998). The Karolinska Directed Emotional Faces - KDEF. *CD ROM from Department of Clinical Neuroscience, Psychology section, Karolinska Institute.*
11. The Color FERET database. face.nist.gov/colorferet/, 2003.
12. Bijlstra, G., & Dotsch, R. (2011). *FaceReader 4 emotion classification performance on images from the Radboud Faces Database.* Unpublished manuscript retrieved from http://www.gijsbijlstra.nl/ and http://ron.dotsch.org/.
13. Langner, O., Dotsch, R., Bijlstra, G., Wigboldus, D. H. J., Hawk, S. T., & van Knippenberg, A. (2010). Presentation and validation of the Radboud Faces Database. *Cognition and Emotion*, **24(8)**, 1377-1388.
14. Benta, K.-I. ; Van Kuilenburg, H.; Eligio, U.X.; Den Uyl, M.; Cremene, M.; Hoszu, A.; Cret, O. (2009). Evaluation of a system for real-time valence assessment of spontaneous facial expressions. *Distributed Environments Adaptability, Semantics and Security Issues, International Romanian – French Workshop, Cluj-Napoca, Romania, September, 17-18, 2009*.
15. Terzis, V.; Moridis, Ch. N.; Economides, A.A. (2010). Measuring instant emotions during a self-assessment test: The use of FaceReader. *Proceedings of Measuring Behavior 2010, Eindhoven, The Netherlands, August 24 - 27, 2010, pp. 192-193*.
16. Ekman, P.; Friesen, W. (1978). *Facial action coding system: A technique for the measurement of facial movement.* Palo Alto, Ca.: Consulting Psychologists Press.
17. Friesen, W.; Ekman, P. (1983). EMFACS-7: *Emotional Facial Action Coding System.* Unpublished manual, University of California, California.
18. Ekman, P.; Hager, J. (2006). *Facial Action Coding System Affect Interpretation Database (FACSAID).* Retrieved from http://face-and-emotion.com/database/facsaid/description.jsp, September 2006.