

FaceReader™ is a tool that is capable of automatically analyzing facial expressions, providing the user with an objective assessment of a person's emotion.

FUNCTIONALITY

FaceReader can recognize a number of specific properties in facial images. First of all FaceReader can recognize the following six basic expressions:

- happy
- sad
- angry
- surprised
- scared
- disgusted

These are the six universal or basic emotions as described by Ekman. These emotions are cross cultural. Additionally, FaceReader can recognize a 'neutral' state. The primary output is a continuous value between 0 and 1 per expression. This value corresponds with the intensity and clarity of an emotion.

Identification or calibration is not required in order to start analysis. All emotions are represented as bar graphs. Emotions can also be displayed as a line graph (selectable). An additional graph summarizes the valence (negativity or positivity) of the emotional status of the subject. 'Happy' is regarded as a positive emotion, 'sad', 'angry', 'scared', and 'disgusted' as negative emotions. The maximum value of these negative emotions is the value of the negative valence. A negative valence can also be used as an indication of frustration.

Data are also visualized in a pie chart, showing the percentage per emotion. The line graphs and pie chart can be copied or saved as an image, for example for use in a Word report.

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. This information can also be stored in the logs created by FaceReader. When using multiple cameras, this information can be used to select the best analysis result.



Main screen of FaceReader. Main areas: 1. bar graphs; 2. line graphs expressions; 3. pie chart; 4. original image; 5. analysis view; 6. image quality bar; 7. settings.

In addition to emotions, FaceReader can also ‘read’ other information from the face:

- gender
- age
- ethnicity (Caucasian, Eastern Asian, African, South Asian, or Other)
- facial hair (beard and/or moustache)
- glasses

A range is given for age. For gender, ethnicity, and wearing of glasses a certainty level (between 0 and 100%) is given. The amount of facial hair is indicated by a value between 0 and 1. This output can be also be regarded as an indication of analysis quality.

Finally FaceReader offers a number of extra classifications:

1. Facial states (classification of certain parts of the face):
 - left and right eye open or closed
 - mouth open or closed
 - left and right eyebrow raised, neutral or lowered
2. Global gaze direction: left, forward or right.
3. Track head orientation, enabling tracking of the full 3D orientation of the head of the analyzed person. Head orientation is recorded in three angles, which denote the head orientation in the X, Y and Z direction.

The extra classifications can also be visualized in line graphs.

INPUT

FaceReader can be used with different input sources:

1. *Off-line using video.* FaceReader supports most common video codecs, including: MPEG1, MPEG2, XviD, DivX4, DivX5, DivX6, H.264 DivX, DV-AVI and uncompressed AVI. The advantage of off-line analysis is that you can achieve a higher resolution. Video can be analyzed frame-by-frame, or at a higher speed, analyzing every 2nd or every 3rd frame. When using frame-by-frame, the sample rate depends on the frame rate of the camera, i.e. 30 frames/sec. The off-line mode makes it possible to detect even short emotion changes (100 msec). For demonstration purposes the video can run in loop-(repeat-) mode. When logging from video, the video time is used in the log (not the PC time). It is possible to select which part of the video has to be analyzed. It is possible to analyze multiple videos in one batch, without having to open every file separately. Analysis results of different videos are stored in separate files.
2. *On-line, for live analysis using a webcam (via USB).* Using a DV camera (camcorder) via FireWire as well as recording and analyzing simultaneously is possible by using a software tool for converting and splitting

the camera signal (though keep in mind this requires a lot of processor power). In on-line mode the sample rate is dependent on processing power and on image quality. Typical sampling rates will be 5-10 frames/sec. On-line analysis is also possible using an IP-camera.

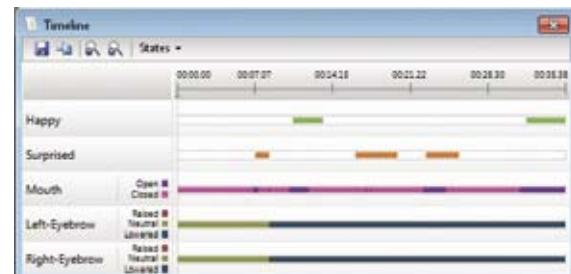
3. *Still Images.* Use *.jpg, *.bmp, *.gif, *.png, or *.tga. Animated gifs are not supported; in this case the first frame of the animated gif will be used. Advised minimum resolution is 640x480 pixels (200x200 pixels for the face area). Advised maximum resolution is 2000x2000 pixels.

The image to be analyzed can be rotated 90°, 180°, or 270°.

VISUALIZATION

The subject’s face can be visualized in different ways:

- *Show framing:* draws a box around the face at the location where the face was found.
- *Show model:* shows the model that has been constructed by the Active Appearance Model (refer to the section background information for more details about this model).
- *Show mesh:* shows the positions of 491 key points used by the Active Appearance Model and head orientation.
- *Show classifications and gaze direction.*



Timeline with a dynamic visualization of facial expressions and classifiers.

OUTPUT

FaceReader can generate two types of text files. The first is the detailed log. This contains all the emotional classifier outputs. If no information was available at a certain time, the log will show the text ‘Missing’, or in case no correct model could be built, ‘FIT-FAILED’ for each column in the record. Additionally, the identifier results, optional classifications and the emotional valence can be logged (selectable).

The second file is the state log. This file contains the ‘emotional state’ a person is in, which is an emotional category shown clearly and with a significant duration. An update to the log file is only made when the state changes.

The data is tab separated, making it easily importable in most spreadsheet programs or text editors. When using a live stream, one can also specify in the settings menu whether to sample at a constant frame rate of 5 frames per second (with possible missing values) or to just log the results whenever they're available. It is important to choose for the first option if results in the detailed log have to be available real-time for use in another program or if they need to be synchronized with another constant frame rate source. This option has to be selected when detailed log data have to be imported in The Observer® XT (as a continuous signal using the External Data Module). FaceReader can also store data in an .ODX format, for direct import into The Observer XT, the software package for collection, analysis, and presentation of observational data. The data can then be synchronized with event logs, keystrokes, mouse clicks, video, screen capture, etc. The state logs are stored as an event file showing the dominant emotional state as behaviors. The detailed log is entered as numerical modifiers in a second event log, showing the exact value of every emotion per sample. Facial states and gaze direction are stored in a third event log. FaceReader data can be visualized and analyzed in The Observer XT, enabling integration with other observational and physiological data.

The advanced video editing functions of The Observer XT can be used to create highlight clips in no time. Finally, it is possible to control FaceReader from within The Observer XT. This enables you to start classification when you start an observation and stop classification when you stop an observation.

COMMUNICATION WITH OTHER APPLICATIONS

Facial expressions detected by FaceReader can be accessed real-time by other applications, making the program an ideal tool for research into affective computing and the design of adaptive interfaces. In other words, FaceReader allows other software programs to respond instantaneously to the emotional state of the user. There are two ways in which other applications can access FaceReader data:

1. FaceReader can write data real time to a .txt file. This file can be accessed by other applications.
2. An additional FaceReader API can send classification results to remote programs over TCP/IP. This FaceReader API can be used in a .Net application to easily make a connection with FaceReader.

MONITOR IDENTITY

FaceReader also includes a basic person identification routine. Once processing starts, the identifier waits until at least 5 frames have been analyzed. When 5 frames have been sampled, a prediction of the current person's identity will be made. These frames will continuously be refreshed and guesses at the current person's identity will keep being made. The name of the test person can be included in the log files.

INDIVIDUAL CALIBRATION

This function enables the correction of person specific biases towards a certain emotional expression. A calibration model can be created using images or video. A minimum of 100 images/samples is required in order to create a new model. This set should contain all expressions relevant for the experiment at least a few times. If there is no calibration model selected, there is a possibility of using continuous calibration. In this mode, FaceReader continuously adapts to the bias of a certain user. In video batch analysis mode, it is possible to select a different calibration model per video.

SET-UP

FaceReader achieves the best performance if it gets a good (video) image. Both the placement of the camera and the lighting of the subject's face are of crucial importance in obtaining reliable classification results.

CAMERA POSITION

The ideal position for the camera is directly in front of the subject's face. If the subject faces a computer screen, the camera can be placed either directly above or directly below the screen. Classification output might have a small bias towards the angry emotion when the camera is placed on top of the monitor and a small bias towards surprised when the camera is placed below the monitor. This is due to the fact that people tend to tilt their head when showing these emotions.

ILLUMINATION

The best results are achieved with diffuse frontal lighting. The light intensity or the color is less relevant. Strong reflections or shadows, for example caused by lights from the ceiling, should be avoided. Sideward lighting from a window will generally degrade performance drastically. If the subject faces a computer screen, two columns of LED

lights to either side of the monitor will give a good result under most circumstances. Two TL-tubes to either side of the monitor will also give good performance under most circumstances.

In situations where interior lighting cannot be controlled, stronger lights (e.g. 4 25W TL-tubes) can be used to negate the effect of other undesirable light sources. Noldus provides illumination for the optimization of your set-up.

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.

BACKGROUND INFORMATION

HOW DOES IT WORK?

FaceReader classifies the face in three consecutive steps:

1. *Face Finding.* First an accurate position of the face is found, using a method called the 'Active Template Method' (ATM). The ATM algorithm displaces and transforms a face template over an image, calculating the most likely face position. If the ATM does not find a face, a 2nd algorithm is used, the Viola-Jones cascaded classifier algorithm, to find a face.
2. *Modeling.* The next step is an accurate 3D modeling of the face using an algorithmic approach based on the Active Appearance Method (AAM). The created model describes 491 key points in the face, including points on the eyebrows, eyes, and mouth which are strong indicators for certain emotions. The texture of the face is described by entanglement of these points, which can contain information on skin wrinkling and lighting changes that are equally important cues for the expression of the face. The AAM uses a set of annotated images and calculates the main sources of variation found in the images. The model (or AAM fit) closely resembles the original face. New faces can be described as deviations from the mean face using a compact vector called the 'appearance vector'.
3. *Classification.* An artificial neural network is trained to classify the six basic emotional expressions and a number of properties. Over 2000 pictures, labeled by experts, were used to train the emotional expression classifiers.

The actual accuracy depends on the quality of the input and the environmental circumstances. As an indication, you may expect an expression recognition accuracy of around 90% when using high-quality image material. For some emotions (happy, scared) the accuracy is higher, for others (angry, sad, surprised, disgusted) lower. It is essential to compare the performance with the performance of a human observer as facial coder, rather than for example a question list completed by the test participant afterwards.

For more details about the methods and algorithms used in FaceReader, please refer to our White Paper 'FaceReader methodology'.

TECHNICAL GUIDELINES

There are a number of limitations to keep in mind before starting to use FaceReader:

- Currently FaceReader is not well-trained to work with children below the age of 3.
- Glasses may hinder classification. Especially thick and dark frames can reduce performance significantly. A polarization filter can help to reduce reflection in the glasses.
- Ethnic groups: in order to analyze expressions from people with different skin colors, it is advised to adjust the illumination or camera settings. FaceReader is trained to work with people from different ethnicities, but may perform less if a face is far from average. The calibration function can be used to improve the performance in these cases. At this moment FaceReader is not well-trained for analysis of children from East Asia and South-East Asia.
- Acceptable pose, movement, and rotation of the test participant are limited. The test participant should stand or sit, and look into the camera frontally (angle <40°).
- The face should not be partially hidden, for instance by a hat, hands, or very heavy facial hair.
- FaceReader can analyze one face at once. If there are more faces in an image, these can be analyzed in different runs, provided the positions do not change significantly.

MISCELLANEOUS

ON-LINE HELP

A comprehensive set of help information is available throughout the program.

DOCUMENTATION

The documentation consists of a comprehensive reference manual, which is also included in *.PDF format on the FaceReader distribution CD.

SYSTEM REQUIREMENTS

Computer

If you order a complete solution from Noldus Information Technology, you will obtain a Dell Precision™ T3500 Workstation (or its successor) with the FaceReader software installed and ready to use.

Technical specifications Dell Precision T3500:

- Processor: Intel® Xeon Quad Core CPU, 2.8 GHz
- Internal memory: 6 GB

If you choose to order a PC from another supplier, you can use the above specifications as a guideline. We

recommend that you use a professional workstation. It is possible to buy consumer-range computers with a high processor speed and plenty of memory, but in order to remain competitive regarding price, the manufacturers often economize on the underlying system architecture. That means those computers are suitable for home use, but not for running professional scientific software. You should select a computer which is intended for professional use or labeled by the manufacturer as a workstation.

Operating system

FaceReader has been thoroughly tested using a US English version of Windows XP (with Service Pack 3) and Windows 7 (32 bit and 64 bit Professional edition). Like any software package, it is possible that minor differences in the operating systems of certain local language versions may affect how well FaceReader runs.

Camera

CCD webcam or FireWire camera with a resolution of at least 640x480 pixels. We strongly recommend that you use a high-quality webcam. Simple webcams are not suitable. You can also use an IP camera or convert a webcam into an IP camera.

A complete FaceReader system includes the following components:

- FaceReader software CD
- printed manual
- PC
- webcam
- illumination

PUBLICATIONS

For an overview of recent publications about FaceReader or research in which FaceReader has been used, please visit <http://www.noldus.com/facereader/recent-publications>

CUSTOMER SERVICE AND SUPPORT

TECHNICAL SUPPORT

Noldus prides itself in superior customer support provided by technically skilled support engineers in our helpdesks, located in The Netherlands, Germany, USA west coast, USA east coast, and China. As a registered user, you are entitled to 3 years of technical support as part of the Basic service contract which is included in the license fee. Extended support is available in the form of higher level service contracts. Technical support questions can be submitted via our web site, by e-mail, telephone, or fax. You can expect an answer within 24 hours (the next working day). An archive of Frequently Asked Questions and Tips and Tricks can be found in a Technical Support Knowledgebase on our web site, offering answers to your questions on a 24/7 basis. You will also receive free subscription to the Noldus Newsline newsletter, which informs you about new product developments, applications, and tips and tricks.

CONSULTING SERVICES

Our customer support goes far beyond solving technical problems. Our staff includes several consultants with an academic degree in the behavioral sciences, as well as qualified engineers. They are happy to help you with research problems related to the use of our solutions, the design of your experimental set-up, the choice of the appropriate hardware for your application, etc.

CONTACT INFORMATION

Please check our website for contact details.

www.noldus.com

FaceReader is a trademark of VicarVision. The Observer is a registered trademark of Noldus Information Technology bv. Due to our policy of continuous product improvement, information in this document is subject to change without notice.