FaceReader™ is a tool that automatically analyzes 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, including the following six basic expressions:

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

These are the six universal or basic emotions as described by Ekman (Universal facial expressions of emotion, California Mental Health Research Digest, 8, 151-158, 1970), which are cross cultural. Additionally, FaceReader can recognize a 'neutral' state and analyze 'contempt' as an emotional state. The primary output is a continuous value between 0 and 1 per expression, which corresponds with the intensity and clarity of an emotion. Furthermore,

identification or calibration is not required in order to start analysis. All emotions are represented as bar graphs and can also be displayed as a line graph. 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.

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.

Valence can also be visualized in a circumplex model, as described by Russell, J. A. (A circumplex model of affect. Journal of personality and social psychology, 39(6), 1161. bce c, 1980). In this model, valence is displayed on the horizontal axis. On the vertical axis the level of arousal is shown. The calculation of arousal in FaceReader is based on activation of Action Units and volatility. Arousal can also be displayed as a line graph.

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.
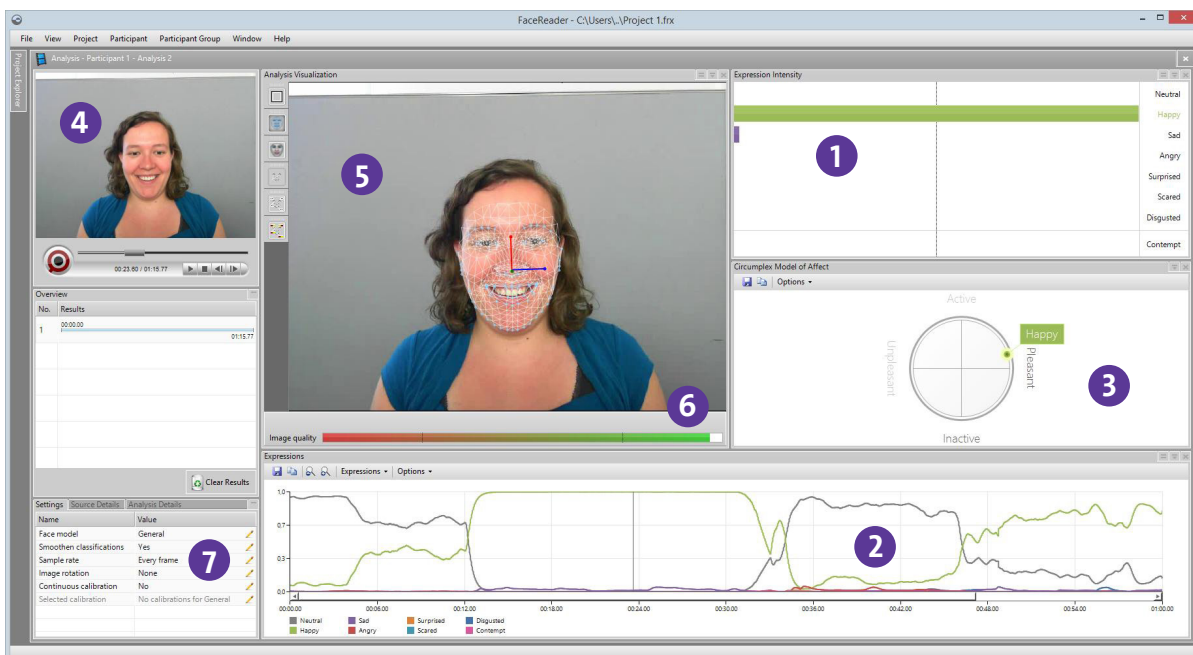
Main screen of FaceReader. Main areas: 1. bar graphs; 2. line graphs expressions; 3. circumplex model of affect; 4. original image; 5. analysis view; 6. image quality bar; 7. settings.
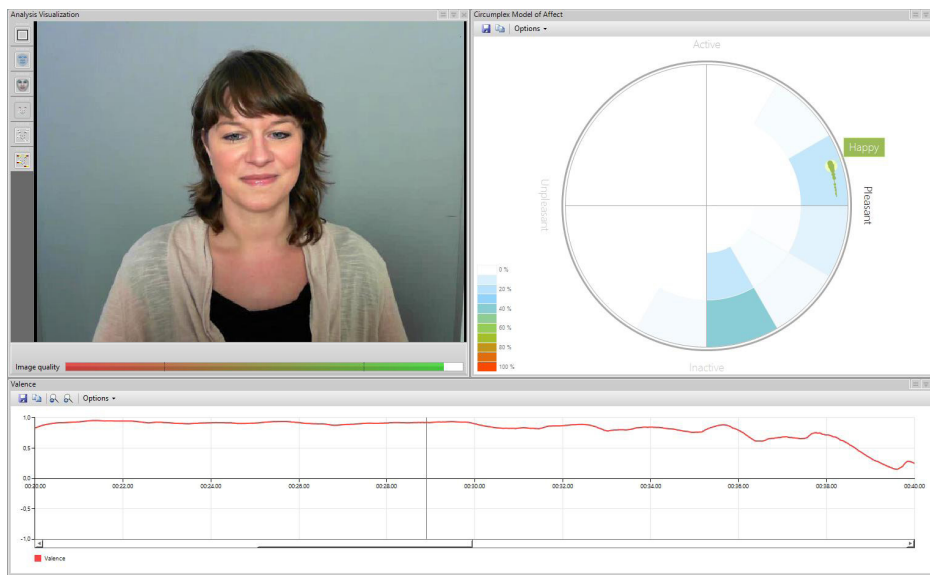
In the circumplex model, valence is displayed on the horizontal axis. On the vertical axis the level of arousal is shown.

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.

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.

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. *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 video 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 video, i.e. 30 frames/sec. This mode makes it possible to detect even short emotion changes (100 msec) and micro-expressions. 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 also possible to play back audio together with the video, FaceReader supports the following audio formats: **xxx**. Multiple videos can be analyzed in one batch, without having to open every file separately. Analysis results of different videos are stored in separate files. The maximum advised duration for video analysis is 2 hours.

2. *Live analysis using a webcam (via USB)* — In this case the sample rate is dependent on processing power and on image quality. Typical sampling rates will be 5-10 frames/sec. Live analysis is also possible using an IP-camera. The live image can be recorded in a DivX/MP4V format, allowing for more detailed frame-by-frame analysis afterwards in FaceReader or integration with other video and data modalities in The Observer® XT. It is possible to record audio together with the video signal. The maximum duration for camera analysis is 2 hours.

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). If the resolution is larger than 1024x1024 pixels, the image is downscaled to maximum 1024x1024 pixels, keeping the original aspect ratio. 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 key points used by the Active Appearance Model and head orientation.
- *Show classifications and gaze direction.*

These visualizations can also be shown simultaneously.
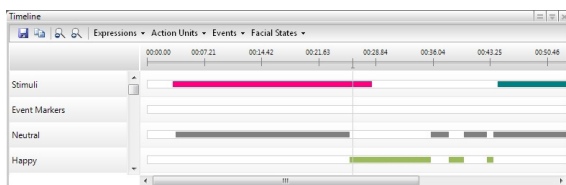
Timeline with a dynamic visualization of facial expressions and classifiers.

OUTPUT

FaceReader can generate two types of text files, the first being a detailed log that 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, image quality, and the emotional valence can be logged (selectable).

It is also possible to log the X- and Y-coordinates of all key points in the mesh. Or just the main ones. The main key points are the points displayed in blue around the mouth, nose, eyes, and eyebrows. The values of the coordinates are relative to the upper-left-corner of the image.

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).

Combining FaceReader with The Observer XT

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, physiological data, eye-tracking data, etc. The state logs are stored as an event file showing the dominant emotional state as behaviors. Facial states, gaze direction, valence, arousal, and head orientation are also included. The detailed log is entered as numerical modifiers in a second event log, showing the exact value of every emotion per sample.

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. 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.

If you have The Observer XT 13 (or newer) with the External Data Module and you observe live, you can use the Noldus network communication protocol N-Linx to control FaceReader with The Observer XT. When you start or stop the observation in The Observer XT, the FaceReader analysis will also start or stop. The FaceReader analysis and FaceReader video can automatically be imported into the observation. In addition it is possible to create a participant with camera analysis when creating an observation in The Observer XT.

Communication with other applications using the FaceReader API

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. The FaceReader API can send classification results to remote programs over TCP/IP. This API can be used in a .Net application to easily make a connection with FaceReader. For more information, refer to the Technical Note *'FaceReader 8 Application Programming Interface'*.

Monitor identity

FaceReader 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 live camera input, or images or video of the test participant showing a neutral expression. 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. During video batch analysis, it is possible to select a different calibration model and face 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. If the subject faces a computer screen, two columns of LED lights or two TL-tubes to either side of the monitor will give a good result under most circumstances.

In situations where interior lighting cannot be controlled, stronger lights (e.g. professional photo lamps) 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.

ADD-ON MODULES

Project analysis module

This add-on module allows detailed analysis and visualization of expression analysis results. The project explorer in FaceReader gives an overview of analyses per participant, and of participants per project. Projects which contain image analysis results cannot be analyzed in the Project Analysis Module.

In order to indicate and select relevant episodes and events for analysis, the Project Analysis Module offers two types of markers. First, it is possible to specify Stimuli. A stimulus has a fixed duration specified by the user, and can be accompanied by a video or image. In case of a video, the start time of the stimulus within the video file can be specified by the user. Next to Stimuli, it is possible to specify Event Markers, for example to indicate that the test participant is being distracted, eating, etc. Both Stimuli and Event Markers are mutually exclusive. They can be regarded as two different behavioral classes

For more advanced coding, it is possible to use The Observer XT.
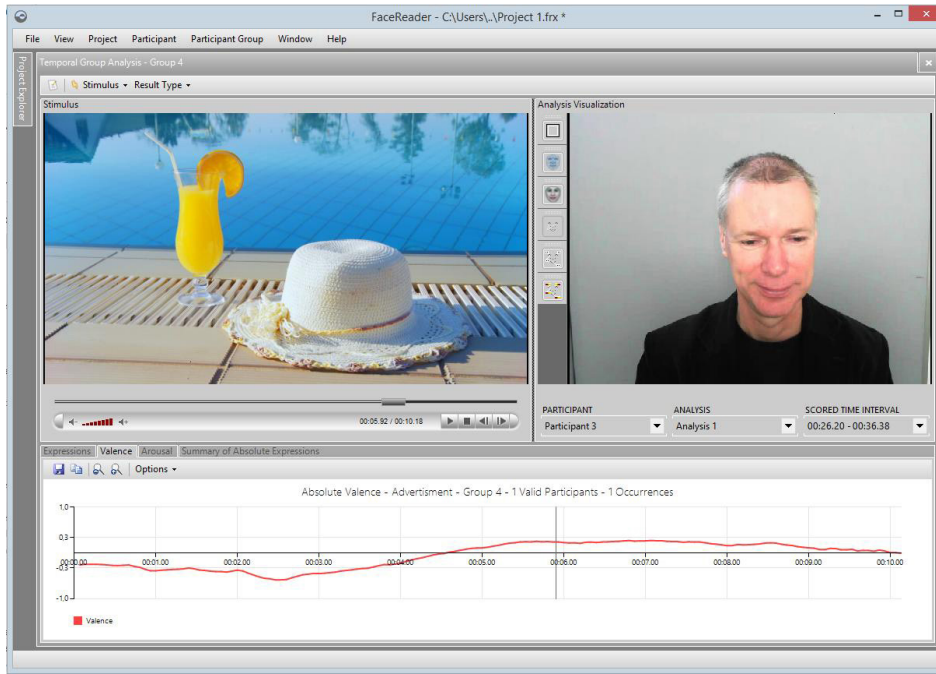
View the participant and the stimulus video.

It is possible to specify 36 Event Markers and Stimuli in total. The start of a Stimulus or the Start/Stop of an Event is triggered by a key (lower case letter or number).

Markers are visualized in the Timeline, the color can be specified by the user. Event Markers or Stimuli can also be triggered externally, using the FaceReader API. Furthermore, this API can be used to generate triggers in case a Stimulus or Event Marker is scored in FaceReader.

Markers can be placed both during live and during video analysis. Time information (start and stop of Stimuli and Event Markers) can be added in the detailed logs and .odx for export to The Observer XT.

It is possible to synchronize stimulus presentation with the trigger of a stimulus marker in the FaceReader pro-

Stimulus Presentation Tool   *FaceReader*

Local network

ject. The Stimulus Presentation Tool can be used on the same computer where FaceReader is running, or on a separate computer. In this case the computers are connected via a local network. Stimulus movies or images can be presented in a fixed or random order. Test participants can enter their name and independent variable values, such as age.

The Project Analysis Module offers a number of different analysis and reports:

A. *Temporal Analysis* — This type of analysis is based on a group response toward a single Stimulus. The average response of the group can be viewed side-by-side with the stimulus video, the participant videos and line-charts of the absolute and relative average expression, arousal and valence. The results can also be presented as a pie chart. When the stimulus includes audio, this is audible when reviewing your results.

B. *Numerical Analysis* — Averages and standard deviations of the responses of groups towards stimuli and event markers can be viewed in tables, bar charts and box-plots. The numerical report of statistics of expressions and valence includes mean and standard deviation.
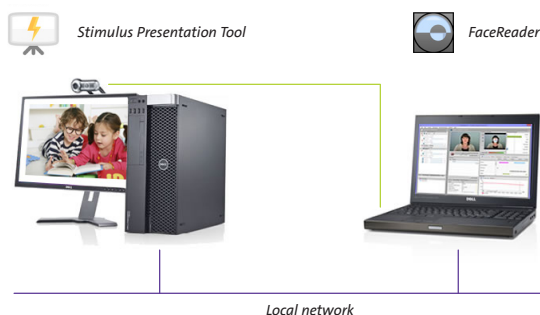
Statistics can be absolute or relative. For relative analysis there are four options:

a. Relative versus all other measurements

b. Relative versus all other Stimuli

c. Relative versus a specific Stimulus which can be selected by the user

d. Relative versus a specific interval before a Stimulus or Event Marker. This option can be used to compare the results with a baseline of a participant.

Line graphs, pie charts and box plots can be copied to clipboard or saved as an image (format .png, .bmp, .jpg, .tiff, .gif, .emf). Analysis data can also be exported in a .txt format.

The participant groups enable the analysis of the average response of separate groups of participants towards stimuli and event markers. Groups can be selected manually or automatically, based on independent variable values. Selections based on combinations of independent variables are also possible.

Independent variables are participant specific variables, defined by the user. There are 2 types of independent variables:

▪ *Numerical* — can take any numerical value.

▪ *Nominal* — predefined values. Examples of nominal independent variables are native language or experience level.

There are 2 default independent variables, which are always present: Age & Gender. These are the independent variables that can also be determined automatically by FaceReader.

Action Unit Module

Action Units analysis is possible with the Action Unit Module. FaceReader can analyze the following 20 Action Units:

Action Unit / Description

1 / Inner Brow Raiser*	15 / Lip Corner Depressor
2 / Outer Brow Raiser*	17 / Chin Raiser
4 / Brow Lowerer*	18 / Lip Pucker
5 / Upper Lid Raiser*	20 / Lip Stretcher*
6 / Cheek Raiser*	23 / Lip Tightener
7 / Lid Tightener*	24 / Lip Pressor
9 / Nose Wrinkler	25 / Lips Part
10 / Upper Lip Raiser	26 / Jaw Drop
12 / Lip Corner Puller*	27 / Mouth Stretch
14 / Dimpler	43 / Eyes Closed*

*For Action Units marked with a *, unilateral analysis is possible: you can choose whether the value of the Left and Right Action Units should be analyzed independently or not.*

Intensities are annotated by appending letters A (Trace), B (Slight), C (Pronounced), D (Severe), or E (Max). Action Units and their intensities can be visualized in the Time-line, and in the Analysis Visualization, and exported in the detailed log. Export in the detailed log as numerical values is also possible. Action Unit intensities can be exported real-time, using the FaceReader API.

It is also possible to define and analyze your own custom expressions. Users of FaceReader can build their own algorithms, using the following measurements as inputs:

▪ Facial expressions

▪ Action Units

▪ Derived expressions (valence & arousal)

▪ Custom expressions

▪ Head orientation

▪ Heart rate, heart rate variability

▪ Constant values

These inputs can be combined in a self-defined algorithm, using the following processors:

▪ Mathematical operations (maximum, minimum, sum, average, scale, offset, subtract, divide, multiply, scale to range)

▪ Logical operations (condition, and, or, not, if..else..)

▪ Temporal operations (average, weighted average, sum, maximum, minimum)

A number of custom expressions is already default available. These are the commonly occurring affective attitudes: interest, boredom, and confusion. With Baby FaceReader, Baby Cry and Baby Smile are available as default custom expressions.

The intensities of custom expressions can be displayed as a line chart, and can also be exported for analysis in other software packages as well.
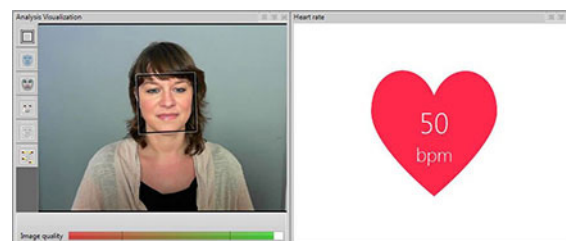
Heart rate monitoring supplies an additional indicator of arousal.

Remote PPG Module

With this module, you can analyze heart rate and heart rate variability (HRV) of the test participant without additional hardware, using the FaceReader camera. Photoplethysmography (PPG) is a simple and low-cost optical technique that can be used to detect blood volume changes in the tissue under the skin. It is based

on the principle that changes in the blood volume result in changes in the light reflectance of the skin. With each cardiac cycle the heart pumps blood to the periphery. Even though this pressure pulse is somewhat damped by the time it reaches the skin, it is enough to distend the arteries and arterioles in the subcutaneous tissue. PPG is often used non-invasively to make measurements at the skin surface. In remote PPG (RPPG), FaceReader can detect the change in blood volume caused by the pressure pulse when the face is properly illuminated. The amount of light reflected is then measured. When reflectance is plotted against time, each cardiac cycle appears as a peak. This information can be converted to heart rate (expressed in beats per minute). The heart rate and HRV can be visualized as a line chart, and exported in the detailed log.

For an overview of the accuracy of the estimated heart rate, see: Tasli, H.E. Gudi, A., and den Uyl, M. (2014b). Remote PPG based vital sign measurement using adaptive facial regions. Proceedings of 2014 IEEE International Conference on Image Processing (ICIP), October 27-30, 2014, pp. 1410-1414. doi: 10.1109/ICIP.2014.7025282.

Consumption Behavior Module

With this add-on module it is possible to analyze behavior of the test participant that is related to eating and talking. In FaceReader 8, this functionality is available on experimental basis for evaluation and validation purposes. The following behavior can be analyzed:

- Chewing
- Taking a bite
- Talking

Also occlusion is detected, which can be useful to select afterwards during which episodes the face was visible or not. Consumption behavior can be visualized in the Timeline, and exported in the detailed log.

BABY FACEREADER

The regular FaceReader enables expression analysis of participants starting from 3 years old. For analysis of younger test subjects you can use Baby FaceReader. This special version supports the analysis of infants between

6 – 24 months. It is possible to either only have Baby FaceReader or the regular FaceReader, or combine both versions in one installation.

Analysis

Baby FaceReader works in a comparable way as the regular FaceReader, but there are a number of differences between the analyses of both versions. The table gives an overview of the analysis offered by both versions.

	Regular FaceReader	Baby FaceReader
Basic expressions	✓	✗
Valence	✓	✓
Arousal	✓	✓
Circumplex model	✓	✓
Gaze direction	✓	✓
Subject Characteristics	✓	✗
Head orientation	✓	✓
Facial states	✓	✓
Monitor identity	✓	✗
Action units	✓	✓
Custom expressions	✓	✓
Remote PPG	ADD-ON MODULE	ADD-ON MODULE
Project Analysis	ADD-ON MODULE	✗
Consumption behavior	ADD-ON MODULE	✗
Calibration	ADD-ON MODULE	ADD-ON MODULE

BACKGROUND INFORMATION

How does it work?

FaceReader classifies the face in the following consecutive steps:

1. *Face Finding* — First an accurate position of the face is found. FaceReader uses the popular Viola-Jones algorithm to detect the presence of a face while eye-detection is used for containing information about the likely in-plane rotation of the 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 over 500 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 basic emotional expressions and a number of properties. Over 10.000 pictures, labeled by experts, were used to train the emotional expression classifiers.

4. *Deep face classification* – with this method, FaceReader directly classifies the face from image pixels, using an artificial neural network to recognize patterns (so no face modeling is done). This has the advantage that FaceReader can analyze the face under more difficult circumstances, for example when part of it is hidden. Deep face classification is done side by side with the Active Appearance Model and enhances the accuracy of facial expression analysis. In addition to this, deep face classification is used stand-alone if modeling with the Active Appearance Model fails, but FaceReader is still able to determine the position of the eyes. In this case, the following analyses can be carried out:

   - facial expression classification
   - valence calculation
   - arousal calculation
   - Action Unit classification
   - Subject characteristics analysis

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 performance and accuracy, please read: *FaceReader AU Module Validation* and the *White Paper – FaceReader Methodology*.

TECHNICAL GUIDELINES

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

- *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.* FaceReader is trained to work with people from different ethnicities. You can calibrate FaceReader to correct for person-specific biases towards a certain facial expression. At this moment FaceReader is less 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°).

- *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. It is possible to specify an area within the image where FaceReader should search for the face.

MISCELLANEOUS

Documentation

The documentation consists of a comprehensive reference manual, which is included in HTML5 format on the FaceReader USB-stick.

Languages

When installing the application, you can choose for an English or Chinese (simplified) version of the user interface. Documentation is available in both languages.

System requirements

Computer

If you order a complete solution from Noldus Information Technology, you will obtain a Dell Precision™ T3620 Workstation (or its successor) which is optimized for this

application and with the FaceReader software installed and ready to use.

Main technical specifications Dell Precision T3620:

- Processor — Intel Xeon E3-1240 v6 (Quad Core 3.7GHz, 4.1Ghz Turbo, 8MB)
- Internal memory — 8GB (2X4GB) 2400MHz DDR4 UDIMM Non-ECC

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 10 Pro (64-bit).

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 with a resolution of at least 640x480 pixels or Imaging Source USB. We strongly recommend that you use a high-quality webcam. Simple webcams are not suitable. You can also use an IP camera. For use of FaceReader in areas with low light conditions, it is possible to use IR cameras. In this case it is not possible to analyze heart rate or heart rate variability with the RPPG module, because this requires color video or camera images.

A complete FaceReader system includes the following components:

- FaceReader software on USB-stick
- manual
- PC
- webcam
- illumination

PUBLICATIONS

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

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 technical support as part of NoldusCare. Technical support questions can be submitted via our website, by e-mail, or phone. 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 Noldus Newsline, 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 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