General set up of a Stress Assessment System based on Multimedial Input data

L.J.M. Rothkrantz, R.J. van Vark, E.J.H. Kerckhoffs
Faculty of Technical Mathematics and Informatics
Julianalaan 132
2628 BL Delft, The Netherlands

At Delft University of Technology, since medio 1992 a research project on stress assessment is running. One of the aims of this project is the development of an instrument for automated stress assessment. The underlying system is based on the analysis of facial expressions, voice analysis and the analysis of physiological signals such as heart rate and blood pressure. Analyses of these multimedial data take place in parallel and are based on AI technology, such as knowledge-based systems and artificial neural networks. In each of the parallel subsystems, corresponding to the sensor, image and sound data, the functionality is split into a number of layers: filtering and reduction layer, preprocessor layer, medium specific layer, application layer and output layer. The results of the analyses are combined by a central interpreter, resulting in an overall stress measure.
The developed system produces feedback about the observed amount of stress and how this stress level has developed in time. In the paper a short overview will be given of the multimedia project, followed by a detailed description of the model and the underlying theory. Finally, some test results and possible fields of applications will be mentioned. A more detailed description of the results will be presented in a separate paper.
Stress involves the internal and external stimulation of an organism which upsets its internal balance. It involves a series of physiological and behavioural changes that prepare the organism for the appropriate coping response. These changes are characterized by some degree of activation (arousal or emotional tension in the organism) which insures the body to be ready for action. The physiological stress response of an organism is dominated by sympathetic arousal characterized by changes in heart rate, blood pressure, respiration patterns, and muscle tension. Of these, respiration and muscle tension are likely to have effect on speech production. From both real life studies of stress and laboratory induced stress experiments, it is found that a number of speech parameters might be viewed as stress indicators: the fundamental frequency, the first and second formant, intensity, speech rate, zero crossings. In nonverbal voice analysis for stress assessment purposes it is necessary to have good software tools and hardware equipment, for the extraction and analysis of the relevant sound and speech parameters. In the paper more details will be given which parameters are chosen and which computer system is chosen to run the developed prototypes.
From the theory of P. Watzlawick it is known that human communication has not only a verbal aspect but also a nonverbal aspect. The full understanding of a message in human interaction is only partly based on the contents of the message. The way of speaking, the corresponding facial expressions and the movements of head and hands greatly influence the understanding of the message. The movement of the mentioned body parts can be registrated by object tracking techniques. The registration and analysis of facial expressions is based on the theory of P. Ekman. He developed a classification model for emotions, based on special "action units" in the face. The interesting parts in the face are the regions of the mouth, the eyes and the eyebrows. Examples of action units are the openess of the eylid, the slanting of the eye, the distant between the brows, the width of the mouth and the openess of the mouth. The challenge is to extract these parameters in an automatic way. Koboyashi and Hara from the University of Tokyo are involved in a project to develop a human robot with human facial expressions. They also

try to generate and recognize facial expressions in an automated way with different techniques.
Recent developments in multimedia offer great opportunities for face to face communications for people on distant processing locations. Communicating by video phone contains much more nonverbal information than a fax message. The data from the different media contains of course some redundancy. The main reason that data from three different media is used, is to reduce the ambiguity in the different data streams. It is however possible to fool the system as it is possible to fool a human observer. The behaviour feedback system has not the pretention to surpass the human observer, but is mainly designed to be used in absence of a human observer, during which situation it can generate the best guess of the amount of observed stress.
The system can be used in several applications:

Training in stress management of pilots and train drivers.
Pilots and train drivers have a stressfull job. Analysis of accidents and near accidents proves that human factors and stress play an important role during the onset of these accidents. Some recently developed training programs in flight simulators include stress modules to train people to manage and control their stress. For autotraining or training without supervision an automated assesment and behaviour feedback system is indispensable.

Teleconferencing
In traditional meetings all participants watch the speaker and listeners simultaneously. Sudden movements of heads, hands and feet draws their attention, so that they inmediately zoom in. In teleconferencing they miss these nonverbal features and need an automated (stress) assessment and alarming system.

Virtual reality
To make the reality less virtual, it is desirable to have feedback from all the senses. The stress assessment system integrates sensor, sound and image data and can be a part of the whole feedback system. It can als be used to guide the future.

Monitoring system
Intensive Care patients in a hospital or babies can get in a case of severe shock, which is another type of stress. In such a life threatening situation a stress monitoring system can generate an alarm if the amount of stress surpass a certain level.

Human user interface
To make computers and robots more human friendly, it is desirable to use emotional expressions as feedback of a computer or robot in a computer or robot interaction.

Psychotherapy/ Assessment of persons
Psychotherapy biofeedback has been used for a long time in stress reduction or stress managing therapy. The use of voice analysis and facial expression analysis can enlarge the reliability of the system.

Lie detector
Traditionally the lie detector is based on sensor data. Using the data from other media such than sound and image can improve the system.

Identification verification

Face recognition and voice recognition is used in person verification. Up to now emotions are considered as noise. But emotion can be a characteristic personal trait . So extracting features of emotions from sensor, sound and image data, can improve the results of person recognition and verification. Especially when the traditional recognition or verification techniques fail due to special masking techniques.
In the paper the above is dealt with in more detail. Special emphasis is put on the multi processing aspects