AUTOMATION OF NON-VERBAL COMMUNICATION OF FACIAL EXPRESSIONS

M. Pantic, L.J.M. Rothkrantz, H. Koppelaar

Delft University of Technology
Faculty of Information Technology and Systems (ITS),
Zuidplantsoen 4, 2628 BZ Delft, the Netherlands
{M.Pantic,L.J.M.Rothkrantz}@cs.tudelft.nl.

KEYWORDS

Intelligent Analysis, Interpretation of MM data, Motion Capture, Computer based training, Desktop video conferencing

ABSTRACT

Human Expression Recognition Clips Utilised Expert System [0] was designed to recognise facial expressions of the observed person in an automatic way. HERCULES forms a part of an Automated System for Non-verbal Communication [1] and represent the implementation of the facial expression analysis that the system performs. The theoretical formulation of facial expression recognition, whereupon inference engine of HERCULES is built, has been acquired from FACS [2]. HERCULES contains also a second inference engine - emotion recognition engine that offers an emotional interpretation to the recognised facial expression. Currently HERCULES' emotion recognition is based on the recognition of so-called six basic emotions defined by Ekman [3a][3b]: happiness, sadness, fear, surprise, disgust and anger. Making the system independent of any of the propagated psychological theories about emotions will be the next step in the development of HERCULES. The system now accepts manually made measurements on a full-face photograph and returns a description of the shown facial expression, quantitative measurement of it, its interpretation in terms of the six basic emotions and quantitative measurement of that emotional interpretation. Although the current version of HERCULES operates with a manual input, a fully automated version of the system is in its latest phase [4].

INTRODUCTION

Human communication has two main aspects: verbal and non-verbal. As well as the words are the atomic information units of verbal communication, the phenomena such as facial expressions, total body posture, skeletal muscle movements and physiological reactions are the atomic units of non-verbal communication. The non-verbal communication forms guidance, amplification or even a replacement of the verbal communication.

At the moment there are several systems available for automatic speech recognition. On the other hand only a handful of attempts have been made for automating the non-verbal communication.

Automated Systems for Non-verbal Communication is an ongoing project carried out by the Knowledge Based Systems Department of the TU Delft. The goal of the project is development of an automated system for non-verbal communication that has to provide qualitative and quantitative information for assessment of mood and intentions of the observed communicator.

The system for the automatic recognition of the non-verbal communication has to cope with the registration, processing and interpretation of the non-verbal communication signals. Based on the input to the system consisting of sound-, image-, and sensor data, parallel performing of voice-, facial expression-, body movements-, and physiological appearances analysis have to take place [1]. All those analyses have to be implemented as modules that can operate as separate units as well as integral parts of the system. The final output of the system would be a combination (made as a so-called data-fusion) of the results of each separate module, representing a hypothesis about the mood and the intentions of the observed person.

Human Expression Recognition Clips Utilised Expert System (HERCULES) has to perform automatic facial expression analysis of the above described system. In other words, it has to perform:

1. automatic facial features extraction,
2. reasoning about shown facial expression (with eventual interpretation of it in terms of emotions), and
3. displaying of the acquired results.

APPLICATION AREA

Automated System for Non-verbal Communication and HERCULES as a part of it has and will have quite a broad area of application. Firstly, the system can be used in different training courses.

- **Training in enlarging the communicating skills** – In the companies that provide services (banks, house-agencies, insurance-agencies) an employee does not have to be only customer friendly but also clear and consistent so that there are no ambiguities between his/her words and actions in order to be clearly understood by the client. An automated feedback system based on the verbal and non-verbal information transfer will reduce the costs and enhance employee's training in achieving an appropriate level of communicating skills. In the scope of HERCULES, by monitoring the employee and compering the automatically obtained interpreted reactions with some to-be-achieved reactions, the employee can be advised automatically what he/she has to change in his/her reaction in order to achieve better communication performance. Costs of the training will be reduced for the salary of human trainer. Furthermore, monitoring and advising one employee at the time (*Fig. 1*) will enhance the training process.

Fig. 1: Human versus automatic observer

This initiated currently debating about applying HERCULES as an automatic training system of the employees of ING bank (the Netherlands).

- **Training in emotion and stress control** – For policeman, pilots, tram and train chauffeurs as well as for all of the jobs were stressful situations are daily present (e.g. nuclear plant operator) such a training is very important. An accurate reliable feedback system for the automatic facial expression recognition and interpretation can form a supplement to, or a replacement of the human observer so that the training becomes cheaper and enhanced. Currently preliminary talks take place with Circon (Amsterdam, the Netherlands), a specialised training centre for training in stress control, aimed at applying HERCULES as an automatic trainer in order to establish the performance of the system, the reduce of the costs and the enhancement of the performed training.

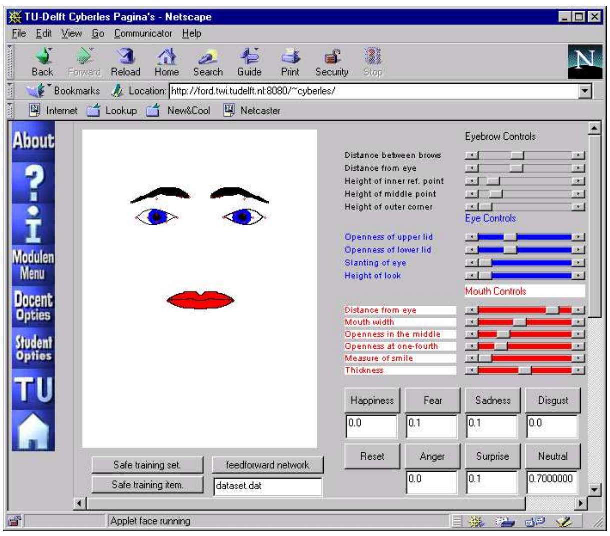


Fig. 2: Cybercourse of Neural Networks

- **Practical training in AI subjects** – HERCULES is an easy understandable expert system made in accordance to all of the basic rules that has to be applied in an expert system development. This makes it very suitable for practical training of the students on the topic of expert systems. Currently, exploring HERCULES forms a part of the practical work required by the Expert Systems course of the ITS faculty of TU Delft. HERCULES has been also used in the practical part of the Neural Networks course of the ITS faculty. The system forms a control mechanism for establishing how good a neural network has been trained on an emotion generator illustrated in *Fig. 2*.

From the rest of the possible applications of the system three have to be mentioned.

- **Monitoring hazardous work places** – All of the airline companies and all their clients fear the moment when the pilot panics. Similarly, each nuclear plant and all of the inhabitants of the area surrounding the plant pray for a cool-headed plant operator. In such hazardous working environments a system like HERCULES which can "on-line" assess the stress of the employee, inform head centre and even turn on/off an "automatic employee" will be of an enormous advantage of keeping the safety on an appropriate level. Although, HERCULES isn't a real-time system yet this sort of applications can become its major goal in the future.
- **Automated information systems** – Many companies have information desks available via telephone. In order to reduce the costs of such a service, the human operators are being replaced by the automated speech recognition systems. The related problem is that in a human dialogue the meaning of words together with the intonation is the key of understanding the intentions of the speaker. Unfortunately, no speech recognition system can understand the intonation. Still the intonation of the speaker is almost always attended by his/her facial expression. In order to come closer to the performance of the human operator the automated information system for providing services via telephone should be expanded with the module containing HERCULES.
- **Multi-modal communication systems** – In nowadays multi-modal communication systems such as videoconferencing, the real-time transmission of the images, the fast image-exchange between different participants and a perfect synchronisation between the sound and the images are not generally available. As a consequence, a lot of non-verbal communication information can be lost and in turn that can lead to

miscommunication. The participants are interested in gaining maximum information from the verbal and non-verbal communication of the partners. HERCULES can provide them with the wanted feedback and can focus the attention of the user on the interesting phenomena's of the non-verbal communication of the participants (e.g. on the bad-mood-face of some partner).

THEORETICAL BASIS

Facial expression recognition engine of HERCULES is based on FACS [2]. Facial Action Coding System is a system that can measure all possible visually distinguishable facial movements. The measurements units of FACS, so called **Action Units (AUs)**, are described in terms of facial muscle actions responsible for each and every facial movement.

Recognition of AUs' activation can be performed directly by connecting the observed person with wires and needles to a measuring instrument that reveals facial muscles contraction. However, this is not only uncomfortable, but also impractical if the person must be observed remotely (e.g. videoconferencing). Another (indirect) way of performing automatic recognition of AUs' activation is automatic facial features extraction.

The facial features movement can be determined by comparing some default position of the facial features with the position of those features found in the currently examined facial expression. Knowing that the movement of facial features is caused by the contraction of a particular facial muscle(s) that, in turn, underlies the activation of a particular AU(s) - by tracing the movement of the facial features, the activation of AUs can be traced.

Once the activation of AUs has been revealed, the shown facial expression is also revealed (in terms of activated AUs). This forms the basic principle of facial expression recognition engine of HERCULES.

Emotion recognition currently performed by HERCULES is based on Ekman's definition of six basic emotions [3]. Considering the fact that each of the six basic emotions is described in terms of the facial expression characteristic for it, when the shown facial expression is automatically recognised by the first part of HERCULES' inference engine, shown emotion will also be recognised.

However there is a lot of argues in the psychological circles about correctness of Ekman's theory. To make HERCULES independent of psychological debates and still to give the possibility of interpreting shown facial expression, a next step in the development of the system must be performed. Learning engine, which will facilitate the user to define his/her own emotional labels coupled with different facial expressions, must be developed. The user can then decide whether he/she will used predefined Ekman's emotional labels or his/her self-defined labels.

MODEL

There are several models describing facial features that can be used. Examples are FCP model [5] and Facial Landmarks model [6]. However, those models are not suitable for an automatic extraction of facial features position. Therefore, a **new model** has been built based on the following two constraints:

1. Defined facial points from which the facial features will be derived must be traceable automatically.

2. It must be possible to establish a simple and unique relation between facial features movement and the activation of separate AUs.

The facial points are described in *Table 1* and illustrated in *Fig. 3*.

Point	Facial position	Stability & Methods of Automatic Tracing
B	right eye inner corner	Stable point; tracing using the micro-features of the eye traced with a NN [4]
B1	left eye inner corner	Stable point; tracing using the micro-features of the eye traced with a NN [4]
A	right eye outer corner	Stable point; tracing using the micro-features of the eye traced with a NN [4]
A1	left eye outer corner	Stable point; tracing using the micro-features of the eye traced with a NN [4]
H	right nostril inner corner	Stable point; tracing by using nostrils template [4]
H1	left nostril inner corner	Stable point; tracing by using nostrils template [4]
C	middle point between nostrils	calculated as ½d(HH1) from the traced nostrils template [4]
D	on the right brow above B	Non-stable point; tracing by using eyebrow template [4]
D1	on the left brow above B1	Non-stable point; tracing by using eyebrow template [4]
E	on the right brow above A	Non-stable point; tracing by using eyebrow template [4]
E1	on the left brow above A1	Non-stable point; tracing by using eyebrow template [4]
F	top of the right eye ellipse	Non-stable point; tracing using the micro-features of the eye traced with a NN [4]
F1	top of the left eye ellipse	Non-stable point; tracing using the micro-features of the eye traced with a NN [4]
G	bottom of the right eye ellipse	Non-stable point; tracing using the micro-features of the eye traced with a NN [4]
G1	bottom of the left eye ellipse	Non-stable point; tracing using the micro-features of the eye traced with a NN [4]
K	top of the upper lip	Non-stable point; tracing by applying colour analysis or active-contour method [4]
L	bottom of the lower lip	Non-stable point; tracing by applying colour analysis or active-contour method [4]
I	right corner of the mouth	Non-stable point; tracing by applying colour analysis or active-contour method [4]
J	left corner of the mouth	Non-stable point; tracing by applying colour analysis or active-contour method [4]
M	bottom of the chin parabola	Non-stable point; tracing by using chin parabola template [4]

Table 1: Facial points

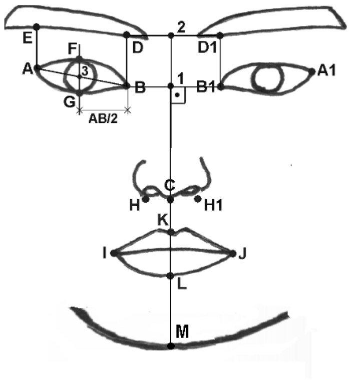
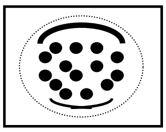


Fig. 3: Facial Features Model

Used facial features are:

- Two angles ∠BAD and ∠B1A1D1, and sixteen distances (*Fig. 3*): d(AE), d(A1E1), d(3F) where point 3 is the midpoint of the line AB, d(4F1) where point 4 is the midpoint of the line A1B1, d(3G), d(4G1), d(FG), d(F1G1), d(CK), d(IB), d(JB1), d(CI), d(CJ), d(IJ), d(KL), and d(CM) where M is the point of chin parabola.
- The intensity of black pixels in seven facial areas (see *Fig. 3*): in the upper half (above the line DD1) of the circle with r(½d(BB1)), C(2); in the lower half (below the line DD1) of the circle with r(½d(BB1)), C(2); in the outer half (right from the line AE) of the circle with r(½d(AB)), C(A); in the outer half (left from the line A1E1) of the circle with r(½d(A1B1)), C(A1); in the outer half (right from the line through point I, normal to the line BB1) of the circle with r(¼d(BB1)), C(I); in the outer half (left from the line through point J, normal to the line BB1) of the circle with r(¼d(BB1)), C(J); and the number (or existence) of white pixels on the line KL.
- Four facial shapes: the pattern of black pixels on the chin that becomes more visible if the chin is raised (*Fig. 4*) and three mouth shapes illustrated in *Fig. 5*.



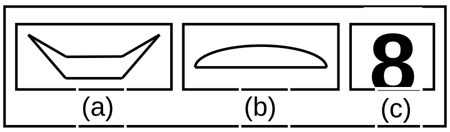


Fig. 4: Chin pattern *Fig 5*: Shapes of the mouth

Throughout performed tests wherein different human observers made manual measurements of the facial features defined by the model it has been confirmed that the facial features defined by model are reliable and quite accurately traceable even manually. Tests of the system confirm that the activation of the recognisable AUs is uniquely defined using this model.

KNOWLEDGE IMPLEMENTATION

Although with the described set of facial features not all of 44 AUs defined in FACS can be described, 32 Action Units can. Furthermore, by observing more than 500 photos of the people photographed when participating in some discussion we came to the conclusion that 12 AUs, which cannot be recognised by the system, seldom occur. Such a conclusion is also logical when man knows that those unrecognisable AUs are “Tongue Show”, “Jaw Sideways”, “Bulge on the cheeks produced by the pushed tongue”, etc.

As an example of describing AUs with HERCULES’ model of facial features we can consider AU6, AU12 and AU25. *Fig. 6* illustrates facial expression caused by the activation of AU6 + AU12 + AU25, and *Table 2* describes the rules of HERCULES’ facial expression recognition engine. The intensity of shown facial expression will be defined as the intensity of activated (recognised) AUs. Currently if a particular AU is recognised an intensity level of 100% will be assigned to it. Otherwise an intensity level of 0% will be assigned to it.

It will be better if the intensity of the AU activation is related to the intensity of the facial appearance change. In that case more than one level of intensity could be associated to the AU activation. However, the way of achieving this is quite ambiguous. To determine the intensity of facial appearance change extremities of that change must be known. The problem is that the extremities of facial appearance changes can, and probably will differ from person to person.

Fig. 6: Default expression & AU6+AU12+AU25

HERCULES’ emotion recognition engine will come to the conclusion that the shown emotion is happiness by compering the AU-description of the facial

AU	Ekman’s description	Mapped on our Model	Rule Name
AU6	cheeks raised, wrinkled outer corners of the eyes	the number of black pixels in the outer half (right from the line AE) of the circle with r(½d(AB)), C(A) becomes greater; the number of black pixels in the outer half (left from the line A1E1) of the circle with r(½d(A1B1)), C(A1) becomes greater.	rAU6
AU12	mouth corners pulled up	d(IB) becomes smaller; d(JB1) becomes smaller; d(CI) becomes greater; d(CJ) becomes greater.	rAU12
AU25	lips parted	d(KL) becomes greater; d(CM) does not become greater.	rAU25

Table 2: Implementation of FACS - HERCULES’ facial expression recognition

expression illustrated in *Fig. 6* and the “trees” of AUs that represent six basic emotions. Those trees are partially illustrated in *Fig. 7*. Intensity of the emotional representation will be defined as the ratio of matched AUs representing recognised facial expression to the number of AUs in each tree representing one of six basic emotions. Emotional interpretation of facial expression illustrated in *Fig. 6* will be 75% of happiness because three from four AUs that form happiness-tree represent facial expression description.

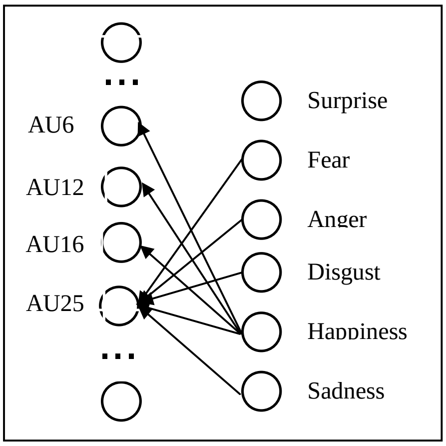


Fig. 7: Mapping of emotions on AUs [3]

SYSTEM STRUCTURE

The input of the system is the set of deviations:

1. between facial distances measured in examined and in some default facial expression,
2. between number of black pixels traced in examined and in default facial expression, and
3. between existence of facial shapes in examined and in default facial expression.

Based upon the registered deviations each facial expression is recognised and classified by the first part of the systems’ inference engine in terms of activated AUs. An intensity level (currently either 0% or 100%) is associated with each particular activated AU. The output of the first part of the system is a description of the shown facial expression (that is, a description of each of the activated AUs) and quantitative measurement of that facial expression (that is, the intensity of each of the activated AUs).

The second part of HERCULES matches the set of activated AUs with the set of AUs characteristic for the six basic emotions. The output of this part of the system is an interpretation of the examined facial expression in terms of the six basic emotions.

The first two parts of the system, forming HERCULES’ kernel, consider processing of one intensity image, that is, they deal with time dependent reasoning about the shown facial expression. The last part of the system (which isn’t implemented yet) has to deal with relating the separate intensity images to each other as a sequence of intensity images in time, that is, it has to deal with time independent reasoning about the shown facial expressions. The output of this part of the system has to be a graphical representation of the registered sequence of facial expressions as a function of time.

The structural representation of HERCULES as a whole can be represented as illustrated in *Fig. 8*.

In **pre-processing** the facial features have to be extracted, the result of the extraction has to be controlled in order to avoid incompleteness and/or redundancy, normalisation of extracted facial distances has to be performed, and the set of deviations has to be calculated.

The set of deviations will form the input to HERCULES’ **kernel** which, based on it, will give the description and the interpretation of the currently examined facial expression.

Post-processing of facial expressions involves time-independent reasoning aspects - graphical representation of the registered sequence of facial expressions as a function of time.

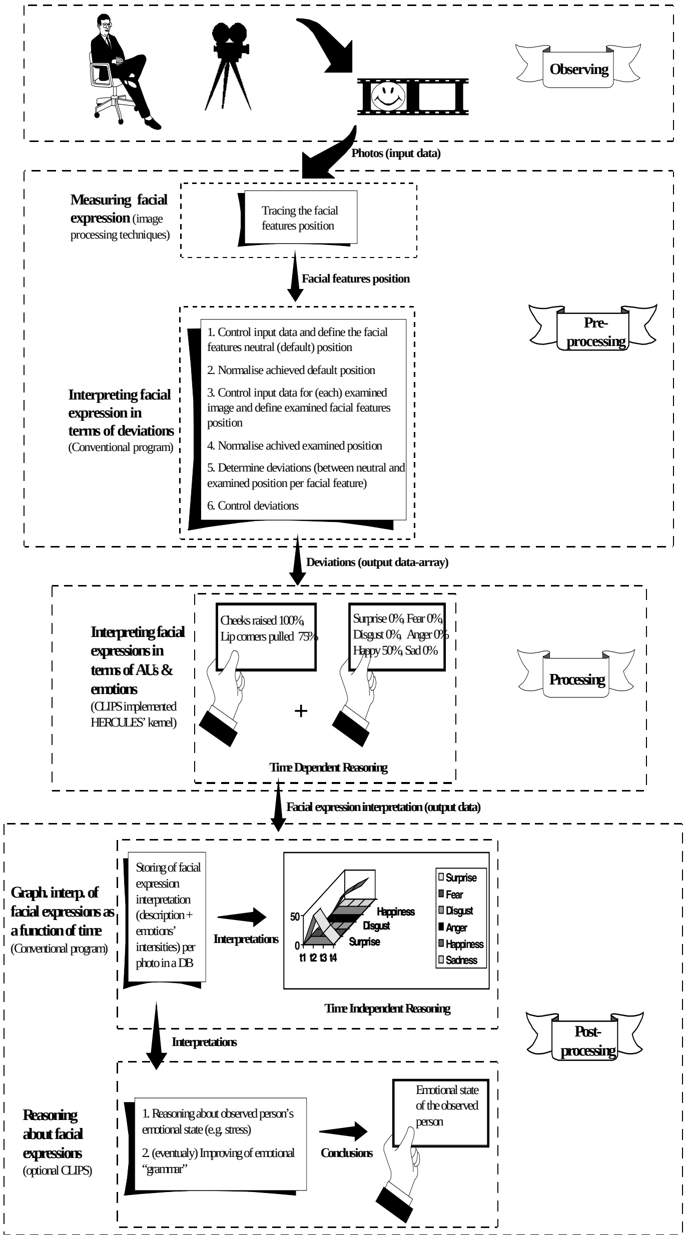


Fig 8: HERCULES

SYSTEM IMPLEMENTATION

HERCULES is an Expert System. The main reason for implementing facial expression recognition and interpretation as an Expert System lies in the characteristics of that task.

1. Knowledge about facial expression recognition and its emotional interpretation has to be implemented so that the reasoning about it can be performed.

2. There is no algorithmic solution to it that can be well structured and, at the same time, which can cover all of the possibilities of the mapping between facial appearance changes, facial expressions and emotions.

3. Knowledge, which has to be built in the system, involves psychological area of facial expressions. It does not involve common sense or general-world knowledge. Therefore it can be said that the domain of knowledge is narrow.

4. Facial expression recognition and interpretation involves reasoning rather than numerical computation.

Global characteristics of a task which can be best implemented using an Expert System are complexity of the task, requiring of reasoning about well defined problem domain rather than numerical computation and impossibility of solving it using a conventional computing method [7]. Knowing this it isn't so strange that the facial expression recognition task has been implemented as an Expert System.

HERCULES is built as a Rule-Based Expert System. There are two major reasons for this. The recognition of shown facial expression is done through recognition of the set of activated AUs. The activation of each particular AU is described by a set of condition clauses (see *Table 2*) - if the condition clauses are true, that AU is activated. The best way of implementing condition clauses is to implement them as rules. The LHS of a rule will contain condition clause(s), describing the activation conditions of a particular AU, while the RHS of the rule will state that the activation of that particular AU is present. Secondly, in a rule-based expert system, the knowledge base can grow little by little as rules are added so that the performance and correctness of the system can be continually checked. Together with the fact that the interactions between rules can be minimised or eliminated, it is possible to build different parts of the system separately, creating independent units that can be checked individually and still be integrated in the final system in the later stages.

HERCULES' kernel is implemented in CLIPS. CLIPS is an acronym for C Language Integrated Production System. CLIPS is a forward reasoning rule-based language, designed at NASA with the specific purpose of providing high portability, low cost and easy integration with external systems [7].

Still, CLIPS isn't a final solution. **Java will be used** in fully automated version of the system. There are several reasons for this. First, the framework containing the modules for automatic tracing of facial features is implemented in Java [4]. When HERCULES is Java implemented, an easy integration to fully automated version of HERCULES will be facilitated. Second, graphical, interactive and user friendly interface will be build into the system. Third, Java is an object-oriented language [8] and, as discussed below, OO paradigm is probably the best paradigm for development of learning facility [6]. Finally, an easy expansion of the whole system is possible by integration of new modules with the existing ones.

An example of HERCULES' processing is given on *Fig. 9*.

(a) default (b) examined

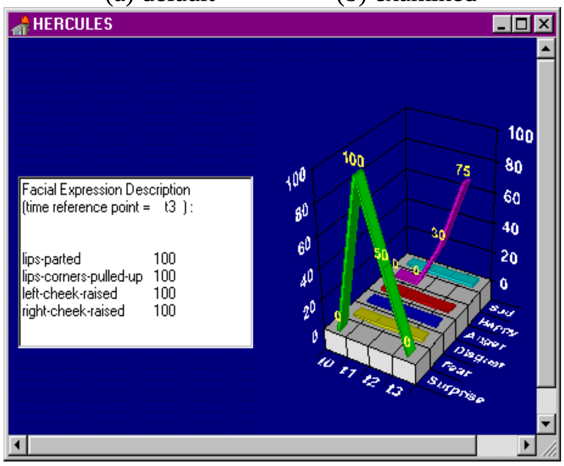


Fig. 9: Processing of HERCULES for manually supplied input measured on photo (b) in comparison to photo (a)

TESTING

The system has been tested using the set of manually made measurements on:

- 125 photographs representing facial expressions caused by the activation of separate AUs,
- 30 photographs representing facial expressions of *pure* emotions (i.e. emotions expressed at maximal intensity [2] [3]), and
- 125 photographs representing *blended* emotions (i.e. some combinations of the six basic emotions or pure emotions expressed at lower than maximal intensity [3]).

Considering the system's main functional requirement – recognition of shown facial expression – the system has been verified and validated.

The purpose of verification was establishing if the product is a product of quality in terms of correctness and reliability of the generated result [9]. Various tests on the above summarised set of input images have been performed. No fault has been encountered and for the same input (same set of facial features' deviations) the same result has been always generated (same description of facial expression).

The purpose of validation was to establish if the product is right product [9]. In the case of HERCULES it means that the system has to perform facial expression recognition. The reasoning mechanism of the system is completely based on, and acquired from FACS [2] which is proved as objective many times in more than fifteen years of its existence. When applied on same set of pictures, the system results always mach the results acquired by applying FACS rules. Therefore the validity of the facial expression recognition performed by HERCULES is ensured.

FUZZY EXPERT SYSTEM

During the process of verifying and validating HERCULES' model several human observers have been asked to define manually the position of the points defined by model in the full-face images. The aim of this manual measurement of the points was determining of the error that a human observer will made when measuring. The reasoning that has been followed concerns the extent of the found error. If the error made by manual measuring of a particular point extends extreme deviation in the position of that point then the model isn't reliable neither for human observers nor for automatic extraction.

Against all expectations the average error, calculated from 225 results measured on 75 different photos, lies in some close neighbourhood of half a pixel error in the case of all points except for the points representing corners of the eyebrows and tip of the chin. In other words, the points defined by HERCULES' model are rather accurately traceable. The exact errors per point are given in *Table 3*.

Point	$\nabla_{average}$	**Point**	$\nabla_{average}$
B	(0.3377, 0.2666)	E	(0.3422, 0.3466)
B1	(0.4044, 0.2622)	E1	(0.3733, 0.3066)
A	(0.3377, 0.3288)	F	(0.2355, 0.3688)
A1	(0.3777, 0.3288)	F1	(0.3022, 0.3333)
3	(0.2133, 0.2533)	G	(0.2400, 0.3422)
4	(0.2711, 0.2311)	G1	(0.3022, 0.2666)
H	(0.4044, 0.3955)	K	(0.3200, 0.3777)
H1	(0.5333, 0.5066)	L	(0.4044, 0.4088)
C	(0.2355, 0.3866)	I	(0.4577, 0.2844)
D	(0.9511, 0.7288)	J	(0.3688, 0.2711)
D1	(0.8177, 0.7422)	M	(0.5777, 0.6222)

Table 3: Average error by manual measuring

For two different photos of the same person the stable points remain stable and will be measured by different human observers with a divergence of the average error given for the particular point in *Table 3*. Further, the average errors of stable points' measurements are far smaller than the distances between stable points used for generalisation of the measurements. Similarly, the average errors of non-stable points' measurements are much smaller than the extreme deviations in the position of non-stable points. Considering these facts it can be concluded that the model is reliable. Testing of the system confirms that the change in the position of the points defined by model uniquely reveals the activation of Action Units, that is shown facial expression. Knowing this it can be concluded that the model is valid. Measuring the position of the points of the model ensures that precisely what has to be measured is indeed measured – the changes in facial appearance that reveal shown facial expression.

No matter how small are the errors encountered when measuring the position of the points defined by model, the measuring isn't completely accurate. Inaccuracy of made measurements must be taken into account as well in the manual input version of the system as in the fully automated version of the system. This is of crucial importance for automatic recognition of facial expressions where image-processing techniques used for automatic tracing of facial features are far from accurate [4].

Assessment of errors made when extracting automatically the facial points defined by the model must be obtained in the same way the average errors made when measuring manually have been assessed. Obtained manual measurements of the points on 75 photos can be adjusted for the value of defined average errors (*Table 3*). Adjusted measurements can be used further as a referent measurement for determining the errors made when extracting automatically the position of the points from the same set of 75 photos. This error assessment can be used further to make HERCULES a fuzzy expert system that will reason about shown facial expression according to the possibility/probability principle [10] so that the imprecision of facial features extraction is taken into account when system's knowledge is processed.

CURRENT STATE-OF-THE-ART

Goal placed on the development of HERCULES is threefold:

1. automatic facial features extraction,
2. reasoning about shown facial expression and eventual emotional interpretation of it,
3. displaying of the acquired results.

Although the system must still be improved in terms of final connection between modules for automatic facial features extraction [4] and the system's kernel, posed goal has been achieved. For the same set of deviations in facial features position, shape or intensity HERCULES will always give the same result – same description of facial expression and same emotional interpretation of it according to Ekman's theory.

In spite of the fact that during the last twenty years Paul Ekman's theory is the most prominent and most commonly used theory of emotions, correctness and universality of it has been set under psychological lens. While Tomkins [11], Izard [12], Ekman [3] and other psychologists of that school state that facial appearances of six basic emotions are indeed universal another psychological circle doubts even the correctness of the definition of the term "emotion". Psychologists like Frijda [13], Smith [14] and Fridlund [15] argue that Ekman's term "facial expressions of emotions" is merely stressing of the verbal communication and has nothing to do with an actual emotional state.

Considering the goal of the research, which isn't solving above described psychological debate but interpretation of recognised facial expression, making the system independent of all psychological polemics about emotions is the best thing to do. The way of achieving this and still retaining the facility of interpreting recognised facial expression is development of a learning facility that will allow the user to define his/her own emotional labels. The user can then decide whether he/she will used predefined Ekman's emotional labels or his/her self-defined labels.

Except for this and making HERCULES a fuzzy expert system several other improvements are to be made.

As we have seen a real-time processing of the system is very important for some of above named areas of application. In order to achieve this goal parallel execution of different modules for automatic facial features extraction has to be performed. Also parallel processing of sequence of images, in contrast to the processing of the single image currently implemented, must still be developed.

Although the system must be improved in all of the above described terms, the mechanics of the system's reasoning have been set.

Considering the fact that the replacement of the manual input handling module with an automatic input handling module is only a question of day, the fully automated version of HERCULES will be released very soon.

REFERENCES

[0] Pantic, M. (1996). *HERCULES*. Master thesis, Delft University of Technology, Department of Knowledge Based Systems, Delft, NL.

[1] van Vark, R.J., Rothkrantz, L.J.M. & Kerckhoffs, E.J.H. (1995). *Prototypes of Multimedial Stress Assessment subject to Laboratory Experiments*. Proc. of MediaComm95, Society for Computer Simulation International, Ghent, Belgium, ISBN 156555-046-3, pp. 108 - 112.

[2] Ekman, P. & Friesen, W.V. (1978). *Facial Action Coding System (FACS) : A Technique for the Measurement of Facial Action. Manual*. Consulting Psychologists Press, Palo Alto, CA, USA.

[3a] Ekman, P. (1982). *Emotion in the Human Face*. Cambridge University Press, Cambridge, UK.

[3b] Ekman, P. & Friesen, W.V. (1975). *Unmasking the Face. A Guideline to Recognising Emotions from Clues.* Prentice Hall, Englewood Cliffs, New Jersey, USA.

[4] Rothkrantz, L.J.M, Vollering, J., Ververs F. & van Schouwen, M. (1997). *A multimedial workbench to analyze facial expressions*. This conference proceedings.

[5] Kobayashi, H. & Hara, F. (1992). Study on Face Robot for Active Human Interface. *IEEE International Workshop on Robot and Human Communication*, pp. 276-281.

[6] Kearney, G.D. & McKenzie, S. (1993). Machine Interpretation of Emotion : Design of a Memory-Based Expert System for Interpreting Facial Expressions in Terms of Signaled Emotions (JANUS). *Cognitive Science* 17, pp. 589-622. University of Greenwich, London, UK.

[7] Giarratano, J. & Riley, G. (1990). *Expert Systems - Principles and Programming*. PWS-KENT Publishing Company, Boston, USA.

[8] Flanagan, D. (1996). *Java in a Nutshell*. O'Reilly & Associates Inc., Sebastopol, CA, USA.

[9] Kan, S.H. (1995). *Metrics and Models in Software Quality Engineering*. Addison-Wesley, Reading, Massachusetts, USA.

[10] Schneider, M., Kandel, A., Langholz, G. & Chew, G. (1996). *Fuzzy Expert System Tools*. John Wiley & Sons Ltd., Chichester, West Sussex, UK.

[11] Tomkins, S.S. (1962). *Affect, imagery, consciousness: Vol. 1*. New York: Springer.

[12a] Izard, C.E. (1971). *The face of emotion*. New York: Appleton-Century-Crofts.

[12b] Izard, C.E. (1994). Innate and universal facial expression. *Psychological Bulletin, 115*, 288-299.

[13a] Frijda, N.H. (1982). The meanings of facial expression. M. Ritchie-Kye (Ed.), *Nonverbal communication today*, pp. 103-120. The Hague: Mouton.

[13b] Frijda, N.H., Kuipers, P. & Terschure, E. (1989). Relations between emotions, appraisal, and emotional action readiness. *Journal of Personality and Social Psychology, 57*, 212-228.

[14] Smith, C.A. (1987). The informational structure of the facial expression of emotion. *Dissertation Abstracts International, 47*, 4002B.

[15a] Friduland, A.J. (1991). Evolution and facial action in reflex, social motive, and paralanguage. *Biological Psychology, 32*, 3-100.

[15b] Friduland, A.J., Kenworthy, K.G. & Jaffey, A.K. (1992). Audience effects in affective imagery: Replication and extension to affective imagery. *Journal of Nonverbal Behavior, 16*, 191-212.