Generation of Facial Expressions from Emotion Using a Fuzzy Rule Based System

The Duy Bui, Dirk Heylen, Mannes Poel, and Anton Nijholt

University of Twente
Department of Computer Science
The Netherlands
{theduy,heylen,mpoel,anijholt}@cs.utwente.nl

Abstract. We propose a fuzzy rule-based system to map representations of the emotional state of an animated agent onto muscle contraction values for the appropriate facial expressions. Our implementation pays special attention to the way in which continuous changes in the intensity of emotions can be displayed smoothly on the graphical face. The rule system we have defined implements the patterns described by psychologists and researchers dealing with facial expressions of humans, including rules for displaying blends of expressions.

1 Introduction

In this paper we introduce a fuzzy rule-based system that generates lifelike facial expressions on a 3D face of an agent based on a representation of its emotional state.

Within the Parlevink research group at the University of Twente, previous work has dealt with natural language interactions between humans and embodied conversational agents in virtual environments ([12], [13]). Our aim is to build believable agents for several application areas: information, transaction, education, tutoring and e-commerce. For an embodied agent to be believable it is necessary to pay attention not only to its capacities for natural language interaction but also to non-verbal aspects of expression. Furthermore, the mind of believable agents should not be restricted to model reasoning, intelligence and knowledge but also emotions and personality. We have therefore started exploring computational models of emotional behaviour, [10]. The representations used in this work form the basis for the results reported here on the facial expressions of emotions.

Based on the descriptive work by Ekman and Friesen in [2] we define rules to map emotion representations onto the contraction level of facial muscles. In the research reported on in this paper, we focus on two aspects of facial expression modeling. First, we want to take into account the continuous changes in expressions of an emotion depending on the intensity by which it is felt. Our fuzzy-rule based approach is chosen to assure smooth results. Secondly, we want to find a way to specify combinations of expressions, i.e. blends, in accordance with the literature mentioned.

Earlier work on computational models of emotion and facial expression includes the directed improvisation system of Hayes-Roth and van Gent [5], which makes an emotion-based selection among animation and audio sequences. Perlin and Goldberg's

Improv animation system [16], [17] layers small animations under the control of scripts and state variables including mood and personality. Stern, Frank and Resner [22] develop an animated pets system named Petz, with facial expression, posture, and vocalizations corresponding to each personality profile and internal emotion state of the character. In most of this work, our concerns with modelling intensity as well as blends figure less prominently. Blends of emotions are often defined in terms of graphics algorithms combining single emotion expressions (using interpolation for instance, [6], [14], [18]) instead of relying on the empirical rules described in the literature. Hendrix et al. [6] use interpolation to display the intensity of emotions. For expressions of blends emotions the basic emotions are arranged on an Emotion Disc with the neutral face in the centre and maximal expressions of emotions on the perimeter. Each position in the Emotion Disc corresponds to an expression obtained by interpolation between the predefined expressions positioned on the disc. This method also does not rely on the empirical literature as the emotion intensity may be represented differently in different facial regions. Beside using interpolation, Pighin et al. [18] also use regional blending to create blends of expressions. However, this method creates uncorrelated facial regions which do not appear in the human face. Moreover, they need to use a very complex 3D face mesh and to collect photographs of expressions of each basic emotion using camera at different positions in order to generate blends of expressions. Therefore, this approach is not suitable for our project which aims at realtime animation of the agent.

The ideas from emotion theory and facial expression on which our work is based are summarised in section 2. In section 3 we give an overview of the complete system that we have implemented. We then discuss the fuzzy rule based system in section 4 in more detail. Some results and the evaluation of the system are presented in section 5.

2   Emotions and Facial Expressions

The rule-based system presented below is based on a collection of theories of emotion and facial expression proposed by [2], [7], [9] and others that has been labeled as "The Facial Expression Program" by Russell [21]. In this program, it is assumed that emotions can be distinguished discretely from one another. A limited number of these are called basic. Opinions differ on what it means for an emotion to be called basic. Russell (o.c.) summarises this discussion as follows: "Each basic emotion is genetically determined universal and discrete. Each is a highly coherent pattern consisting of characteristic facial behavior, distinctive conscious experience (a feeling), physiological underpinnings, and other characteristic expressive and instrumental actions."

In this paper we consider the following six emotions: Sadness, Happiness, Anger, Fear, Disgust and Surprise. These are said to be universal in the sense that they are associated consistently with the same facial expressions across different cultures ([2]). In this book, Ekman and Friesen also describe in details what the expressions for these emotions and certain blends look like.

Emotion feelings may differ in intensity. In [2] it is pointed out how for each of the basic emotions the expression can differ depending on the intensity of the emotion. It is therefore important for us to build our system on a representation that takes intensities into account.

The human face is also able to show a combination of emotions at the same time. These are called blends. Ekman and Friesen describe which blends of the basic emotions occur and what these blends look like universally. We have used their descriptions as the basis for our fuzzy rules.

3   Overview of the System

Our system maps a representation of the emotional state to a vector of facial muscle contraction intensities which is used to control the facial expressions of the 3D face. The system, as shown in figure 1, consists of six components:

1. The input is an Emotion State Vector (ESV). This is a vector of basic emotion intensities represented by a real number:
   $ESV = (e_1, e_2, ..., e_6)$ where $0 \leq e_i \leq 1$
2. The output is a Facial Muscle Contraction Vector (FMCV):
   $FMCV = (m_1, m_2, ..., m_{18})$ where $0 \leq m_i \leq 1$
   This is a vector of facial muscle contraction intensities.
3. The Expression Mode Selection determines whether a single emotion or blend of two emotions will be expressed in the 3D face model.
4. In the Single Expression Mode muscle contraction intensities from a single emotion intensity are produced.
5. In the Blend Expression Mode FRBS muscle contraction intensities from two emotion intensity values are produced.
6. The muscle based 3D face model expresses the emotions.

Fig. 1. The proposed system

FRBS. The core is formed by the fuzzy rule-based system (FRBS). Two collections of fuzzy if-then rules are used to capture the relationship between the ESV and the FMCV. During fuzzy inference, all the rules that are in one of the collections are fired and combined to obtain the output conclusion for each output variable. The fuzzy conclusion is then defuzzified with the Center of Area(COA) [11] method, resulting in a final crisp output.

The fuzzy if-then rules for both single and blend expressions are based on Ekman and Friesen's summary of the facial movements to express basic emotions [2]. We

used several other information sources to map the descriptions of faces from Ekman and Friesen onto the values for the muscle contraction intensities that generate the expressions. These sources were the Facial Action Coding System (FACS)[3], and the book and tutorial by, respectively, Waters and Parke [15], and Prevost and Pelachaud [20]. Also our own observations on emotion expression in human faces have played a role. We will discuss these rules in more detail in section 4.

As can be seen from the system, the FRBS is actually broken up into three components: the Expression Mode Selection, the Single Expression Mode FRBS and the Blend Expression Mode FRBS. The expression of an emotion in a blend may differ in important ways from the expression of the emotion occurring on its own. Typically, for a single emotion expression several regions of the face are involved whereas in blends one of these regions may be used for the other emotion. We therefore do not want the single expression rules to fire when blends occur. It might be possible to build a system with just a single collection of fuzzy rules. However this will complicate the statement of the rules.

The emotional state vector, ESV, represents the emotional state of the agent. The human face cannot display all the combinations of emotion intensities that can be felt at the same time universally and unambiguously. It seems that only two emotions can be displayed at the same time, because the face has only a limited number of regions to display emotions. The mapping between emotional state and facial display is not direct also for other reasons. Several factors may be involved in real persons to decide for an emotion that is felt whether or not it will be displayed. There may be cultural rules for instance that inhibit showing certain emotions. An Expression Mode Selection module can mediate between the emotion state as it is felt and the rules for representing the emotions to be displayed. In our current implementation we select either the single or blend expression mode based on the intensities of the emotions felt[1].

FMCV. The muscle contraction intensities which the rules give rise to are used to manipulate the 3D face. Currently, we use 17 muscles and an additional parameter, Jaw Angle. The latter determines how far the mouth will be opened.

We have used the 3D face from Waters [15] for this project. The reason for using this face is that it is detailed enough to generate almost every visually distinguishable facial movement. It also provides an easy way to define a suitable muscle system.

The muscle system was defined for the face on the basis of anatomical information ([15] and [20]). We first created the muscles in the 3D face model at the positions similar to those of real muscles. Next we adjusted our muscle system until it produced reasonably lifelike effects on the 3D face model. For the adjustments we also relied on the photographs in [2].

---

[1] The Single Expression Mode is selected when a single emotion has to be expressed. This the case when only one emotion has an intensity bigger than 0.1 while other emotions have intensities close to zero (smaller than 0.1). In this case, the Single Expression Mode Fuzzy Rule Based System (FRBS) will be used, and the input of the FRBS is the single emotion with highest intensity. When the Blend Expression Mode FRBS is used, the input of the FRBS is the pair of emotions with highest intensity (in the case that more than two emotions have the same highest intensity, two emotions will be randomly selected to express). We certainly do not claim psychological realism here.

Some muscle types, such as the sphincter muscles, have not been implemented in Waters' 3D model. We therefore replaced the sphincter Orbicularis Oculi and Orbicularis Oris, by a collection of other muscle types combined into a circle to produce similar effects as the real sphincter muscles. The muscles that are implemented in our system can be seen in figure 2 and table 1.
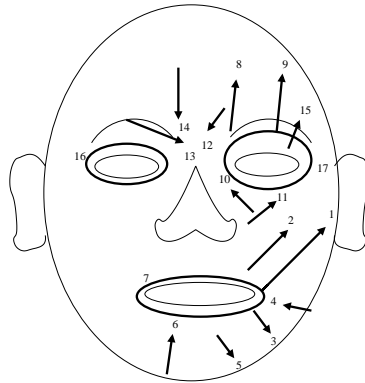
Fig. 2. Schematic of all implemented muscles

Table 1. Implemented muscles in the system

No.	Muscle name	No.	Muscle name
1	Zygomatic Major	10	Levator Labii Nasi
2	Zygomatic Minor	11	Levator Labii Superioris
3	Triangularis	12	Depressor Supercilli
4	Risorius	13	Corrugator Supercilli
5	Depressor Labii	14	Depressor Glabelle
6	Mentalis	15	Levator Palebrae Superios
7	Orbicularis Oris	16	Orbicularis Oculi Palebralis
8	Frontalis Medialis	17	Orbicularis Oculi Orbitalis
9	Frontalis Lateralis		

The system is designed to take into account future expansions. First, the introduction of the FMCV enables the combination of an agent's lip movements during speaking with facial emotion expression. Secondly, the use of the ESV and the Expression Mode Selection allows the integration of the agent's intention and personality into the model without changing the fuzzy rules for expressing emotions. This can be done by distinguishing the real ESV as felt from something like a "to-display" ESV. The "to-display" ESV which does not represent the agent's real emotion state but the emotion state the

agent want to express. For example, with a strong personality, the agent may display a fake smile to mask sadness by increasing the intensity of happiness in the "to-display" ESV.

4   The Fuzzy Rule Based System

The subsystems "single expression mode" and "blend expression mode" are both implemented using fuzzy logic. Both subsystems must convert an emotion state to a contraction level for the facial muscles taking into account the intensity of the emotions. In the literature on facial expressions of emotions qualitative descriptions like "surprise then lift eyebrows" can be found. In order to take intensities into account as well these (logical) rules where transformed into fuzzy rules. The fuzzy rule approach allows us to incorporate qualitative descriptions as above with quantitative information (emotion intensity and contraction level). Moreover we still have a comprehensible rule-based system in which the logical descriptions are visibly encoded. We would miss out on that when using other models like neural networks, for instance.

First we model the emotion intensity by five fuzzy sets (figure 3): **VeryLow**, **Low**, **Medium**, **High**, and **VeryHigh**. The contraction level of each muscle is described by again five fuzzy sets (cf. figure 4): **VerySmall**, **Small**, **Medium**, **Big**, and **VeryBig**. The exact form of the membership functions and the support of each membership function are experimentally determined by hand.
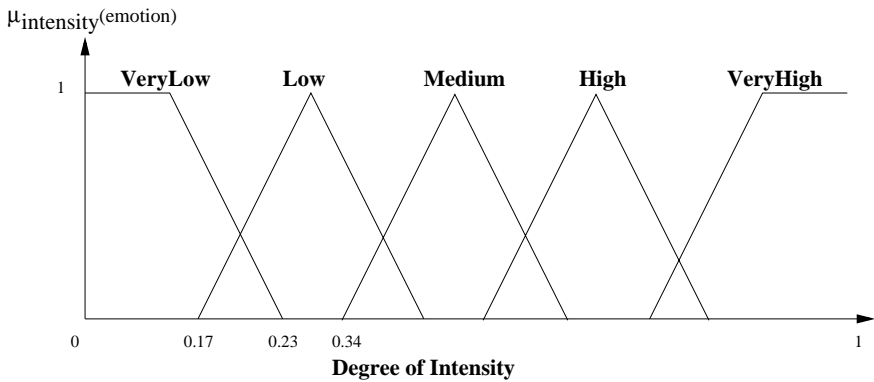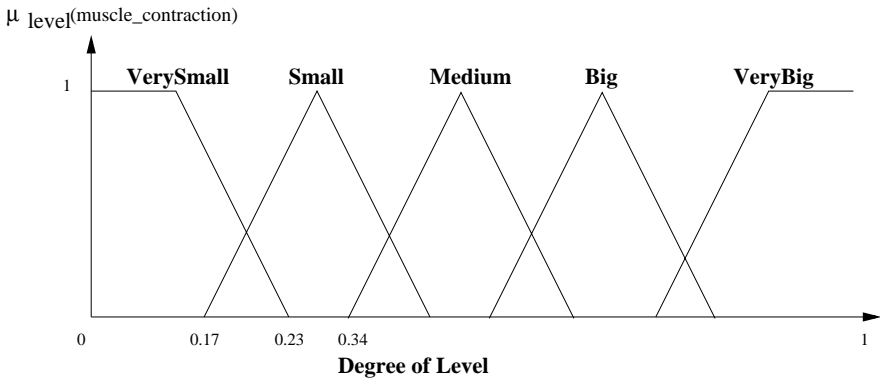

Fig. 3. Membership functions for emotion intensity

As we explained in the previous section, the Expression Mode Selector decides whether a single emotion has to be displayed or a blend. The rules in the single-expression mode FRBS take on the following form.

If Sadness is **VeryLow** then muscle 8's contraction level is **VerySmall**, muscle 12's contraction level is **VerySmall** ...

μ level(muscle_contraction)

Fig. 4. Membership functions for muscle contraction level

The sample rule above encodes the information presented in the first row of table 2. Note that the relation between the emotion intensities and the muscle contraction level is not so straightforward that we can use a simple mapping system. The name and position of the muscles can be seen in table 1 and figure 2. All the rules for other single emotions are represented in the table form and can be found in [1].

Table 2. Fuzzy rules for emotion **Sadness**, vh:**VeryHigh** h:**High** m:**Medium** l:**Low** vl:**VeryLow** vs:**VerySmall** h:**Small** m:**Medium** l:**Big** vl:**VeryBig** -:no contraction

E. Intensity	m8	m12	m13	m14	m17	m16	m3
vl	vs	vs	vs	vs	vs	-	-
l	s	s	s	s	s	vs	-
m	m	m	m	m	m	s	-
h	b	b	b	b	b	m	m
vh	vb	vb	vb	vb	vb	m	b

If the Single Expression Mode is not used, then the Blend Expression Mode is selected. In this mode two emotions are displayed on the face. Normally each of the two emotions is displayed in a separate region of the face. The fuzzy rules for the blend of expressions reflect this fact. The contraction level of a muscle is determined by the intensity of the emotion that will be displayed in the facial region to which this muscle belongs. As the contraction level of each muscle is determined by the intensity of only one of the emotions, there will not be conflict values placing on any muscle's intensity. We will illustrate this with a description of the blend of sadness and fear.

Ekman and Friesen [2] describe how in a such a blend sadness is expressed in the brows and eyelids while fear is expressed in the mouth. Combining this with the specification of muscle movements in the FACS, we can define the emotions in muscle terms.

Sadness is expressed by contracting the muscles Frontalis Medialis(8), Depressor Supercilli(12), Corrugator Supercilli(13), Depressor Glabelle(14) and Orbicularis Oculi(16 and 17). Fear is expressed by contracting the muscles Triangularis(3), Risorius(4), Depressor Labii(5) and by opening the jaw. The level of contraction of each of those muscles is then determined by the intensities of sadness and fear. The format of such a rule in our system looks as follows.

If Sadness is **Low** and Fear is **Medium** then muscle 8's contraction level is **Small**, muscle 3's contraction level is **Medium** ...

Some examples of the rules are presented in table 3. The full set of rules can be found in [1].

5   Result and Evaluation

The expressions of six basic emotions and a neutral face are displayed in figure 5. In figure 6, surprise is shown with increasing intensity. The increasing intensity of surprise can be seen in the increase in the raising of the eyebrows and the increase in the opening of the mouth. Figure 7 (left) shows the blend of anger and disgust. It can be seen that anger is represented in the eyebrows and eyelids while disgust is represented in the mouth. Blend of happiness and surprise are shown in figure 7 (right). This is a combination of surprised eyebrows and a happy smiling mouth.

Table 3. Fuzzy rules for blend of **Sadness** and **Fear**

| Sadness	Fear	m8	m12	m13	m14	m17	m16	m3	m4	m5
vl	vl	vs	vs	vs	vs	vs	-	vs	vs	vs
vl	l	vs	vs	vs	vs	vs	-	s	s	s
m	l	m	m	m	m	s	vs	s	s	s
...	...	...	...	...	...	...	...	...	...	...

The results also show that the emotions are not only displayed in the main parts of the face like mouth and eyebrows but also in very detailed parts like eyelids and lips. The blends of expression are displayed according to the rules as described by psychologists instead of being computed by some graphics algorithm that combines values for single emotion expressions (morphing, interpolation). And finally, the quality of the facial expressions is improved by the smooth relationship function between emotion intensities and muscles contractions level. This smooth relationship function is obtained with fairly simple fuzzy if-then rules rather than with complicated formulas or intensively trained Neural Networks.

For a first evaluation, questionaires were set up to assess the recognizability of the expressions generated by the system. The expression of six basic emotions and a neutral
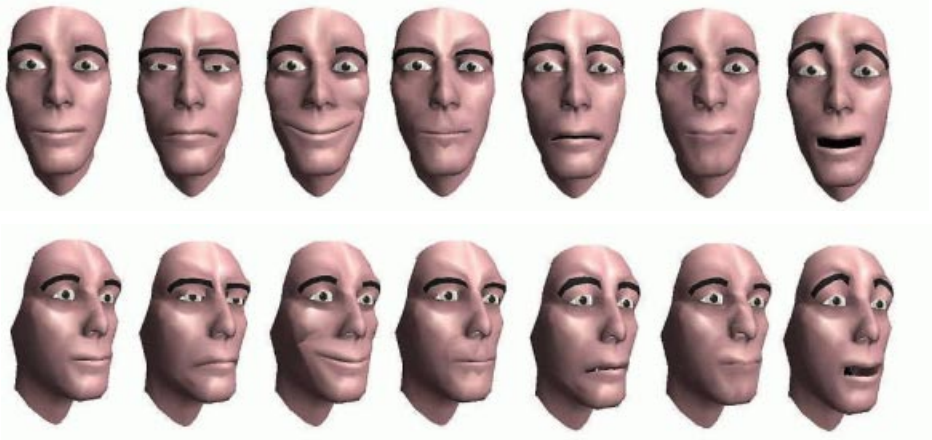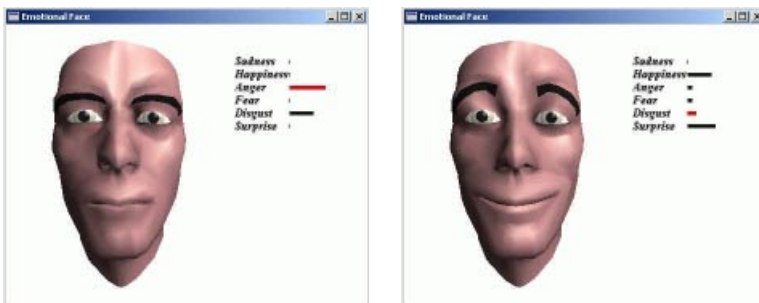
Fig. 5. Basic emotions: Neutral, Sadness, Happiness, Anger, Fear, Disgust, Surprise (from left to right)
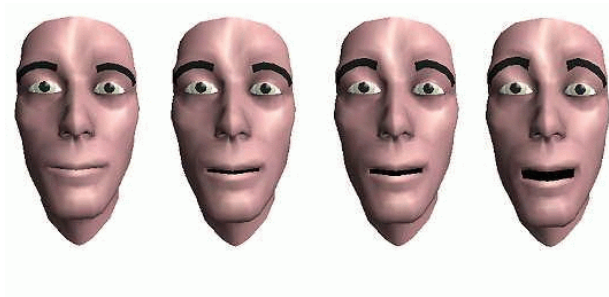
Fig. 6. Increasing surprise

Fig. 7. Blend of anger and disgust(left), happiness and surprise(right)

face generated by the system were shown to 20 people. The result of how they recognised these emotional expressions is summarized in table 4. As can be seen from the table, the generated emotion expressions are recognised as what they are intended to be by a large

percentage of the people. We also showed these people the picture of the generated blend expression of Anger and Disgust (figure 7 left) to see how good our blend expression generation is. The possible answers that people could choose from were expressions of: Sadness and Disgust, Anger and Disgust, Fear and Anger, Fear and Sadness, Sadness only, Disgust only, and Anger only. There were 11 people who recognised it as a blend expression of Anger and Disgust; 3 people recognised it as the expression of Sadness only; 1 person recognised it as an expression of Disgust only; and 5 people recognised it as an expression of Anger only. So about half of the people recognized it correctly. For an appropriate analysis of this result we need further questionaires comparing these with similar expressions generated by other systems or with the photos from Ekman and Friesen.

Table 4. Evaluation result on how people recognise generated facial expressions

	Intended	Neutral	Sadness	Happiness	Anger	Fear	Disgust	Surprise
Recognised as								
Neutral		95%	-	-	-	-	-	5%
Sadness		-	75%	-	15%	5%	5%	-
Happiness		-	-	100%	-	-	-	-
Anger		5%	5%		70%	10%	10%	-
Fear		-	15%	-	-	80%	-	5%
Disgust		-	5%	-	10%	5%	75%	5%
Surprise		-	-	-	5%	-	10%	85%

6   Conclusion and Future Research

In this paper, we have proposed a fuzzy rule based system to generate facial expressions from an agent's emotional state. With simple fuzzy rules, lifelike facial expressions are generated based on descriptions in the literature. The variations resulting from differences in the intensity of emotions are also successfully displayed.

The effect of the fuzzy membership function on the manner of expression is one of the issues for future research. In the next phase of the project, we intend to add a learning component into the system so that we can have slightly different way of expressing an emotion state for different agents. The display of emotions will have to be combined with other systems that influence what is shown on the face (like communicative signals and lip-movements). The expression mode selector will have to become more complex to take into account other factors besides intensity. Finally, other emotions besides the "universal" ones will have to be dealt with.

References

1. Bui, T.D (2001), Generation of facial expression from emotion using a Fuzzy Rule Based Sytem, Technical Report, Parlevink, Faculty of Computer Science, University of Twente.
2. Ekman, P. and W. Friesen(1975). *Unmasking the Face*. Prentice Hall.
3. Ekman, P., and W. Friesen, Facial Action Coding System. Consulting Psychologists Press, Inc., 1978.
4. Galernter, D.H. (1994), *The muse in the machine*. New York: Free Press.
5. Hayes-Roth, B. and R. van Gent(1997), Story-making with improvisational puppets, in *Proceedings of 9th Conference Uncertainty in Artificial Intelligence*, pp 122-127, San Francisco, Morgan Kaufmann.
6. J. Hendrix, Zs. Ruttkay, P. ten Hagen, H. Noot, A. Lelievre, B. de Ruiter (2000), A facial repertoire for avatars, Proceedings of the Workshop "Interacting Agents", Enschede, The Netherlands.
7. Izard, C.E. (1971), *The face of emotion*, New York: Appleton-Century-Crofts.
8. Izard, C.E. (1991), *The psychology of emotions*, New York: Plenum.
9. Izard, C.E. (1997), Emotions and facial expressions: A perspective from Differential Emotions Theory, in J.A. Russell and J.M. Fernandez-Dols (Eds.), *The Psychology of Facial Expression*, Maison des Sciences de l'Homme and Cambridge University Press.
10. Kesteren, A.J., R. Op den Akker, M. Poel and A. Nijholt (2000), Simulation of emotions of agents in virtual environments using neural networks. In: *Learning to Behave: Internalising Knowledge*. Proceeding Twente Workshops on Language technology 18 (TWLT18).
11. Lee, C.C. (1990), Fuzzy Logic in control systems: Fuzzy logic controller - parts i and ii. *Fuzzy Sets and Systems* 15:111-128,224-240
12. Nijholt, A., M. van den Berk and A. van Hessen (1998), A natural language web-based dialogue system with a talking face, *Proceedings Text, Speech and Dialogue*. Sojka et al (eds.), Brno, Czech republic, pp 415-420.
13. Nijholt, A., D. Heylen and R. Vertegaal (2000), Inhabited interfaces: Attentive conversational agents that help. In: Proceedings 3rd international Conference on Disability, Virtual Reality and Associated Technologies - CDVRAT2000, Alghero, Sardinia.
14. Paradiso, A. and M. L'Abbate (2001), A Model for the Generation and Combination of Emotional Expressions, In: *Proceedings of the Workshop on Multimodal Communication and Context in Embodied Agents*. C. Pelachaud and I. Poggi (eds.) - Autonomous Agents, Montreal, Canada.
15. Parke, F.I. and Waters, K. (1994) *Computer Facial Animation*, AK Peters. ISBN 1-56881-014-8.
16. Perlin, K., and A. Goldberg (1996), Improv: A system for scripting interactive actors in virtual worlds, in SIGGRAPH'96, *Proceedings of the 23rd Annual Conference on Computer Graphics*, pp 205-216, New York, ACM Press.
17. Perlin, K. (1997), Layered compositing of facial expression, SIGGRAPH'97 Technical Sketch, New York University Media Research Lab. Available: $< http : //mrl.nyu.edu/improv/sig97 - sketch/ >$.
18. F.Pighin, J.Hecker, D.Lischinski, R. Szeliski, and D.Salesin (1998), Synthesizing realistic facial expressions from photographs, in *SIGGRAPH 98 Conference Proceedings, pages 75-84*, ACM SIGGRAPH, July 1998.
19. Plutchik, R. (1962), *The emotions: Facts, theories, and a new model*, New York: Random House.
20. Prevost, S. and C. Pelachaud. (1995). Talking Heads: Physical, Linguistic and Cognitive Issues in Facial Animation. Course Notes for Computer Graphics International 1995, Leeds, UK

21. Russell, J.A. and J.M. Fernández-Dols (1997), The meaning of Faces, in J.A. Russell and J.M. Fernandez-Dols (Eds.), *The Psychology of Facial Expression*, Maison des Sciences de l'Homme and Cambridge University Press.
22. Stern, A., A. Frank, and B. Resner(1998), Virtual Petz: A hybrid approach to creating autonomous, lifelike Dogz and Catz, in *Proceedings of the Second International Conference on Autonomous Agents*, AGENTS98, pp 334-335, New York, ACM Press.