

Design of a Realistic Robotic Head Based on Action Coding System

Samuel Marcos¹, Roberto Pinillos¹, Jaime Gómez García-Bermejo², and Eduardo Zalama²

¹ Fundación Cartif.

Parque Tecnológico de Boecillo Parcela 205.
47011 Boecillo, Valladolid, Spain
`sammar@cartif.es, robpin@cartif.es`

² Instituto de las Tecnologías de la Producción, University of Valladolid.
Paseo del Cauce 59, 47011 Valladolid, Spain
`sammar@cartif.es, robpin@cartif.es`

Abstract. In this paper, the development of a robotic head able to move and show different emotions is addressed. The movement and emotion generation system has been designed following the human facial musculature. Starting from the Facial Action Coding System (FACS), we have built a 26 actions units model that is able to produce the most relevant movements and emotions of a real human head. The whole work has been carried out in two steps. In the first step, a mechanical skeleton has been designed and built, in which the different actuators have been inserted. In the second step, a two-layered silicon skin has been manufactured, on which the different actuators have been inserted following the real muscle-insertions, for performing the different movements and gestures. The developed head has been integrated in a high level behavioural architecture, and pilot experiments with 10 users regarding emotion recognition and mimicking have been carried out.

Keywords: FACS, mechatronic head, emotional expressions, pilot study

1 INTRODUCTION

The development of robots with social skills is one of the challenges to which more efforts has been devoted in the research centers, during the last years. These robots should be characterized by the ability of transmitting visual information through gestures, in order to achieve a natural human-to-robot communication similar to the human-to-human one.

However, researchers have not reached a general consensus about what a robot's appearance should be. There is a large number of particular complications associated to the design of more realistic social robots [2], and although a human appearance may improve the interaction so to overcome the limitations imposed by the Uncanny Valley Theory [3], there are always inconveniences associated with the behavioral expectations that an anthropomorphic robot creates in the users [4].

Regarding the degree of realism expected for a humanoid robot, in [6] the need to use more sophisticated ways to express emotions in virtual environments was denoted. Also, the results presented in [7] show that little expressive agents had more difficulties to attain an effective interaction. On the other hand, in [8] simple drawings of faces were able to create emotional feelings on the user. The face is a highly expressive element that humans tend to interpret in face to face communication, so highly realistic artificial faces may confuse the user if they are not designed properly [7]. According to [9], there is evidence that the use of exaggerated cartoonish faces can express emotions more accurately than realistic ones.

Since the emerge of humanoid robots, the use of unrealistic or cartoonish-like humanoid robots has been justified in terms of avoiding well-known uncanny valley. However, different studies show that user preferences in terms of appearance and behavior are not universal and may be evolving as the technology spreads. These preferences may differ depending on cultural and psychological aspects [10] [11] [12], as well as on age or gender (men vs. women, elderly vs. young people).

The first developments of expressive humanoid robots have been designed using a cartoonish simple approach [16]. One of the first developments is the Kismet robot [17]. This robot is equipped with an unrealistic face with large expressive eyes, eyebrows and mouth. Another example is the Robot WE-4R Waseda [18] which can output rich emotional expressions and behavior by using a simplistic face, waist, arms and neck. However, although these robots exhibit large expressiveness, they lack of the ability to express detailed emotional cues. It is necessary to incorporate movements that mimic human gestures in greater detail, as is evident in the development of robots with humanoid appearance, like the android HRP-4C [19] presented by the National Institute of Advanced Industrial Science and Technology (AIST) of Japan, or the DER2 and DER3 projects constructed by Osaka University.

There exist other approaches that try to emulate the human face structure at a greater detail. For instance, in the Geminoid robot [13] and [20] they use pneumatic actuators to display facial expressions. In [21] an interesting emotional architecture is presented, and in [22] the muscular structure of the human face is imitated with great detail. However, the obtained results either do not have a natural appearance or are very hard to implement.

The proposed approach described in this paper also looks for obtaining a realistic mechatronic face. We follow the biomechanic principles of the human face based on microexpressions. Our goal is to obtain visually realistic results, so the anatomical features of a real face are taken into account in the design of the robotic head. Our approach is similar to Hansons work [14],[15] as the generation realistic emotional expressions is based on the human muscular model, but while Hansons model uses 32 servo motors coupled to a mask to map the facial muscles, we have developed a simplified model using 22 servomotors that can be engaged and disengaged easily from the facial mask, thus allowing to use different masks

Table 1. Relevant muscles of the human face.

area	Most significant muscles
Forehead and eyebrows	Occipitalis and frontalis
Eyes and eyelids	Elevator palpebrae superioris, orbicularis oris, corrugator supercilii
Lips and mouth	Quadratus labii inferioris, Quadratus labii superioris, triangularis, zygomaticus major and minor, buccinator, orbicularis oris, caninus
Nose	Pyramidalis transverse nostril dilatator, elevator muscles, transverse
Ears	Superior and posterior auricularis

to the same cranial structure. Our approach also eases the integration of the robotic head with a high level control architecture developed in our lab.

2 The human facial musculature

The human skull includes 14 bones to give support to the face. The facial musculature, in turn, is related directly to the different expressions the face can show. The main face muscles involved into facial gestures, along with the face areas they belong to, are shown in table 1. The synchronized movement of these muscles allows the different gestures and emotions to be performed.

2.1 FACS - Facial Action Code System

The relationship between a given muscle contraction and the corresponding shape of the face surface can be described upon the muscular anatomy. However, from the viewpoint of social interaction, the existing relationship among the facial expressions, what those expressions mean and what others can interpret from them should be established. To this end, social psychology researchers have addressed the development of standards for coding and parameterizing the facial movements objectively. This way, emotions and facial expressions can be linked upon specific patterns of facial movements.

Some recognised standars are: EMFACS [23], *Monadic Phases* [24], *the Maximally Discriminative Facial Movement Coding System* [25], and the *Facial Action Coding System* [26].

The (*Facial Action Coding System*, FACS) [26] is one of the most widespread standard used for measuring and describing the face behavior.

The FACS is structured into *Action Units*, AU. An AU is a contraction or relaxation of a given set of muscles that results into a change in the face appearance. The FACS is not based on the activity of isolated muscles because certain muscles may actuate more than once during the production of a given gesture. Moreover, the face appearance changes are weakly dependent on the movement

Table 2. Common AUs involved in the universal expressions.

Universal Expression	AUs
Joy	6+12
Sadness	1+4+15
Surprise	1+2+5+26
Fear	1+2+4+5+20+26
Anger	4+5+7+23
Disgust	9+15+16

of an isolated muscle. There are more than 70 AUs involved in the movement of a human head and face. 46 AUs are linked specifically to the contraction of facial muscles, and 30 of them are linked to the contraction of specific face muscles: 12 for the upper face and 18 for the lower face.

According to Ekman, there are 6 universal expressions: happiness, disgust, sadness, anger, fear and surprise. Each expression is defined (although not completely, as we will discuss later) by a given number of AUs. The universal expressions and the most common AUs associated are listed in table 2.

3 Mechanical design

3.1 Full model

Figure 1 shows side and front views of the full model. A silicon skin is then added to this model, to complete the realistic look of the robotic head. For a complete description of the mechanical structure of the head please see [28].

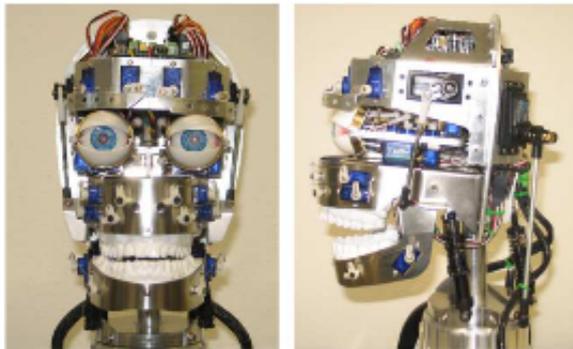
**Fig. 1.** Side and front view of the head.



Fig. 2. (a). Head with mask underlayer and silicon mask. (b) Molds employed for the mask construction.

3.2 Face mask

First, a plaster model of the desired head has been created. As our robot is intended to interact with aged people, a mature aged face has been selected. In addition, the wrinkles ease performing the different gestures. Then, a negative mold has been generated that will be used for producing the outer surface of the mask. A mold corresponding to the inner surface of the mask has been also generated. This mold provides the required thickness to the different areas of the mask according to the envisaged behavior of the skin during the servomotor actions. Both outer and inner molds can be seen in figure 2.

Figure 3 shows the final look of the head, with the skin on, after the painting the outer surface and including additional elements such as eyebrows, eyelashes and facial hair. An intermediate foam layer has been added between the mechanical structure and the mask in order to fill the structure gaps and facilitate the skin movements (see figure 2).



Fig. 3. Final look of the head with the silicon mask.

4 Emotion generation

The prominent biologist Charles Darwin was the first to suggest that the expressions corresponding to the main primary emotions can be easily perceived and categorized into different gesture patterns performed by the human face [1]. Following this work, Ekman [26] postulated that there are six universal expressions corresponding to the six primary emotions expressions: happiness, disgust, sadness, anger, fear and surprise. In the present work we have modeled the six emotional expressions as a linear combination of the involved AUs. Each AU j is given a *weight* value that defines its level of activation in the current gesture, according to the desired intensity of this gesture. Moreover, a gain term K_i is added which defines the intensity of the overall gesture.

$$ComplexExpression_i = K_i \sum (weight_j * AU_j) \quad (1)$$

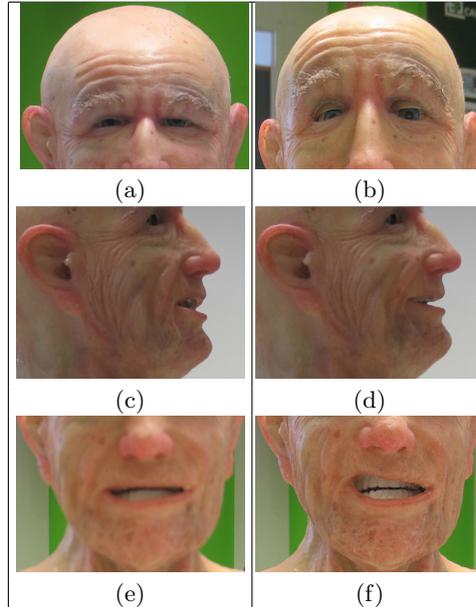


Fig. 4. Example of AUs
 (a) Neutral (b) AU4 (c) Neutral (d) AU12 (e) Neutral (f) AU5

Figure 4 shows the basic emotional expressions performed by the robot. The generation of the happiness expression involves the coordinated activation of AUs UA6+UA12+UA25+UA26 that correspond to lifting the cheeks, pulling up the mouth corners, separating the lips and lowering the jaw. The final result can be seen in figure 5.a.

For instance, the sadness expression involves the activation of AU1+AU17+AU54+AU55+AU63 that correspond to raising the inner part of the eyebrows, raising the chin and tilting forward the head and turning it left and moving the eyes up. (Figure 5.b).

The anger expression shown involves the activation of AU4+AU7+AU10+AU16+AU55+AU63 that correspond to lowering the eyebrows, narrowing the eyelid aperture, raising the upper lip, depressing the lower lip, tilting forward the head and slightly moving the eyes up. (Figure 5.c).

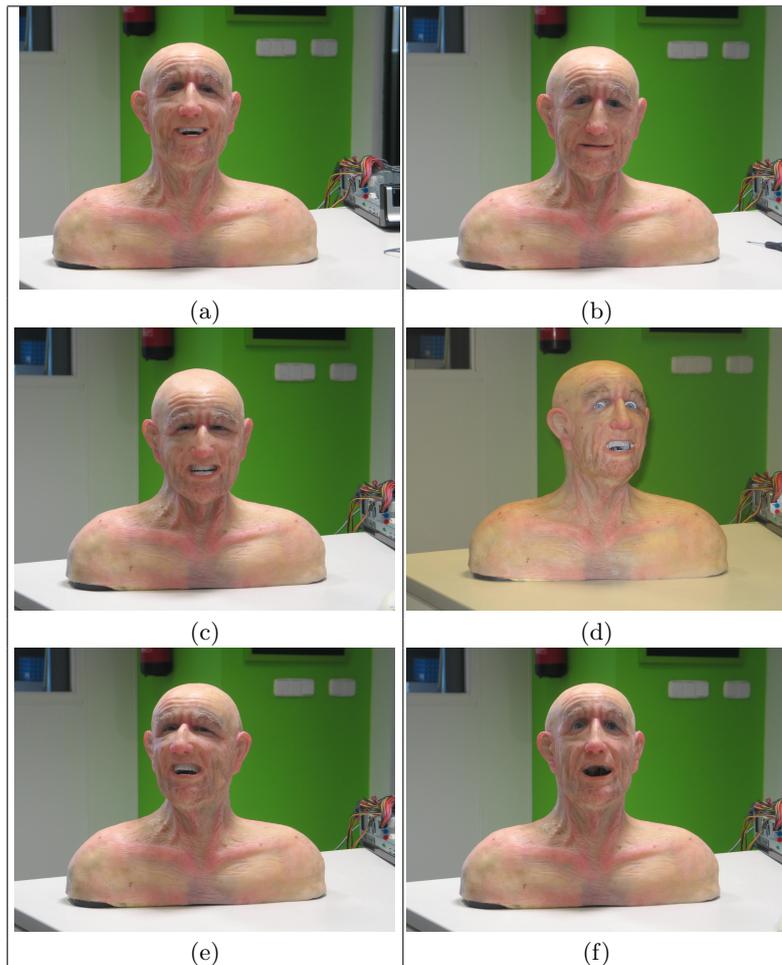


Fig. 5. Basic Emotions

(a) Joy (b) Sadness (c) Anger (d) Fear (e) Disgust (f) Surprise

5 Integration in a multimodal architecture

Currently there is a growing interest in modular architectures and component frameworks that encourage the creation of easily modifiable and scalable open systems. This kind of initiative is of particular interest in the field of mechatronics and service robotics, where systems are specifically designed for high level interactions and sought to be easily modified and improved to meet new requirements. The main trend these days in terms of development of robotic systems is the design of architectures based on components or frameworks. For this purpose rules for combining components and the interaction patterns between them are defined in order to achieve a flexible system. This flexibility is understood mainly in the following aspects:

- Easy replacement of components with different functionality.
- Reuse of different modules in other projects where the functionality of those particular modules is required.

The potential of the open modular architectures is huge, and leads to reconfigurable, portable and extensible systems. This philosophy has been adopted to develop the control software of different social robots in our laboratory. Based in the ROS framework, it consists of a set of layers with different interchangeable modules (see Figure 6). The interaction between layers and modules is well defined, so a new module can easily be integrated. It has to be noted that this architecture has not only been used in robots, but also in other agents such as animated faces [5]. Thanks to the AU based emotional generation approach, both animated and robotic faces can be controlled using the same upper level components.

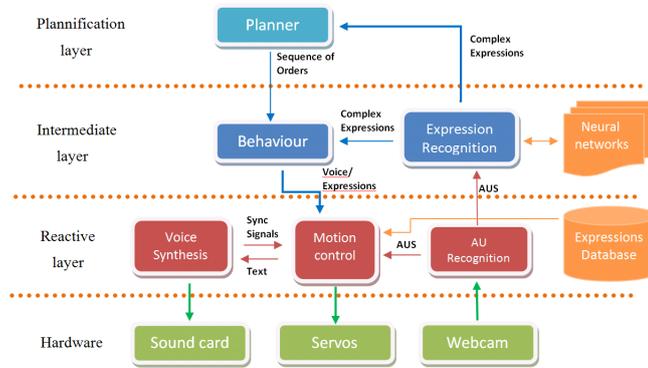


Fig. 6. Modular multimodal architecture

Making use of this architecture, the mechatronic head and its low level control drivers have been integrated within the down layer, taking advantage of

the upper layer services such as voice recognition, expression recognition or behavioural planner.

As can be seen in the video linked to this article (see Section 7 below), as an example the robotic head has been easily connected to our previously developed facial expression recognition system. This system is able to perform both: facial expression recognition and facial emotion recognition. Expression recognition refers to the detection and analysis of facial movements, and changes in facial features from visual information. On the other hand, emotional expression recognition also associates these facial displacements to their emotional significance, given that there will be certain movements associated with more than one emotion. The proposed method [5] is divided into two different stages: the first one focuses on the detection and recognition of small groups of facial muscle actions using a combination of Gabor filters and Active Shape Models. Based on the output of this first stage, the second determines the associated emotional expression using an habituation based network plus a competitive based network. Thanks to the implemented control architecture, it can be observed in the video how the mechatronic head is able to mimic the emotional expressions of the user using a regular webcam.

6 Pilot experiments in emotion recognition

The mechatronic head expression generation capabilities were tested in a pilot study with 10 participants which ages ranged between 28 and 40 years old. In order to see the influence of the intensity parameter k_i (see equation 1) in the generation of recognisable expressions, users were asked to recognize the six universal expressions under two conditions: half-intensity ($k_i = 0.5$), and full-intensity ($k_i = 1$). Expressions were randomly generated for each condition and users were asked to annotate which expressions they were able to recognize, as well as the order the mechatronic head had displayed them. The subjects answers in labeling an expression were scored as correct or wrong and results are shown in figure 7.

From the results it can be observed that the intensity has a great impact in the recognition of the majority of the expressions, specially in the anger expression. Apart from the joy expression, a greater intensity results in a higher recognition rate, although in expressions such as sadness or disgust the difference is not highly significant. It has to be noted that figure 7 just shows the correct answer ratio and not those answers in which users doubted between two expressions. This happened between the anger and disgust expressions and between the fear and disgust expressions. Users found it difficult to distinguish between them, and labeled the shown expressions as both anger and disgust or fear and surprise. It can be seen that this confusion is even more accentuated in the case of the half-intensity expressions. It has to be noted, however, that the small number of participants make some of the obtained results far from being significant. For instance, recognition rates of 3 out of 10 are close to randomness meaning that users were no able to recognize the expression. In order to that

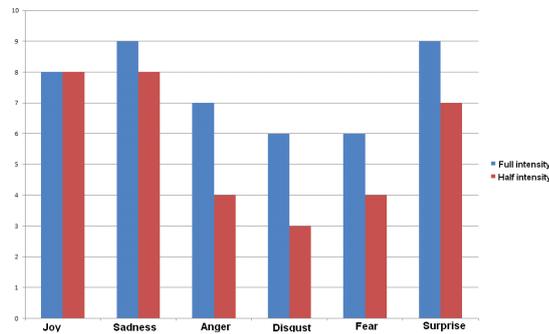


Fig. 7. Recognition rate out of ten for the six universal expressions

obtained results are in fact due to a lack of expressivity of our robotic head, a bigger experiment with more samples should be performed.

7 VIDEO

A video of the head construction, and the head performing different gestures and mimicking can be seen at:

<http://www.eii.uva.es/\%7Eeduza1/videos/HeadMask.mp4>

8 CONCLUSIONS

In this paper a methodology for developing a robotic head with large realism and expressiveness has been presented. Many researchers have suggest that the development of simplified facial animation models, even further away from human appearance in order to avoid creating false expectations on the interaction ability of the robots, should be preferred. Moreover, this would allow avoiding the adverse implications of the Mori's uncanny valley theory. However, the development of realistic robots that create the illusion of life is a challenge that contributes to raising the emotional communication ability of robots over that of simplified models.

We have addressed the development of a design methodology of realistic robotic heads, that is based on the neuromuscular human system and the FACS. On the one hand, we have undertaken a deep study of the different AUs and an analysis on which AUs are the most relevant for generating movement and emotions. This research has resulted into a refined model with 30 AUs, which can be implemented physically. On the other hand, we have designed an arrangement of servomotors that allow the facial muscles to be emulated. Also, a silicon skin has been designed and manufactured, on which the different actuators has been inserted in a muscle-insertion way, after a detailed analysis of the human anatomy. Also, the presence of wrinkles in this skin provides increased expressiveness.

An outstanding characteristic of the proposed design is that the skull, with the corresponding actuators and the associate electronics, is a separate element. In particular, the silicon skin can be easily removed away from the mechanical structure, which simplifies the assembly/ disassembly operations and thus, the robot tuning and subsequent maintenance. Even, silicon masks other than that designed in the present work could be mounted onto the same mechanical structure.

The current research work opens up some important horizons. One future research line will consist in evaluating more extensively how accurately (and pleasantly) the robot emotions are perceived by humans. Also, as there is evidence that men and women display differences in both cognitive and affective functions [29], our future experimentation should take into account those individual differences. Recent studies have examined the processing of emotions in males and females, Another interesting next-future line is to evaluate the communication skills of the robot with respect to previous researches that use virtual avatars [30] or non-realistic mechatronic heads [31].

References

1. Darwin, C.: *The Expression of Emotions in Man and Animals*.(872). John Murray, reprinted by University of Chicago Press (1965)
2. Duffy, B. R.: Anthropomorphism and the social robot. *Robotics and Autonomous Systems* 42, 177–190 (2003)
3. Mori, M.: The uncanny valley. *Energy* 7(4), 33–35 (1970)
4. Dautenhahn K., Werry I.: Towards interactive robots in autism therapy: background, motivation and challenges. *Pragmatics and Cognition* 12(1), 1–35 (2004)
5. Marcos, S., Gmez, J., Zalama, E., Lpez, J.: Dynamic Facial Emotion Recognition Oriented to HCI Applications. *Interacting with Computers* 27(2), 99–119 (2015)
6. Saugis, G., Chaillou, C., Degrande, S., Viaud, M.L., Dumas, C: A 3-d interface for cooperative work. In *In: Proceedings of the Conference on Collaborative Virtual Environments*, 17-19 (1998)
7. Paiva, A., Machado, I.: The child behind the character. *IEEE Journal on Systems Man and Cybernetics, Part A*, 31(5),361–368 (2001)
8. Nass, C., Reeves, B.: *The Media Equation, How People Treat Computers, Television and New Media Like Real People and Places*. Cambridge University Press (1996)
9. C. Bartneck. *An Embodied Emotional Character for the Ambient Intelligent Home*. PhD thesis, Technische Universiteit Eindhoven (1982)
10. Koda, T., Ishida, T.: Cross-cultural study of avatar expression interpretations. In: *International Symposium on Applications on internet*, 130-136 (2006)
11. Deng, Z., Hiscock, M., Yun, C.: Can local avatars satisfy a global audience? a case study of high-fidelity 3d facial avatar animation in subject identification and emotion perception by us and international groups. *Computers in Entertainment* 7(2), 1-26 (2009)
12. Kanda, T., Ishiguro, H., Hagita, N., Bartneck, C: Is the uncanny valley an uncanny cliff? In: *Proceedings of the 16th IEEE International Symposium on Robot and Human Interactive Communication, RO-MAN 2007*, 368-373 (2007)
13. Nishio, S., Hiroshi Ishiguro, I., Hagita, N.: *Humanoid Robots, New Developments*. I-Tech Education and Publishing, 343–352 (2007)

14. Mazzei, D., Lazzeri, N., Hanson, D., De Rossi, D.: Hefes: An hybrid engine for facial expressions synthesis to control human-like androids and avatars. In: 4th IEEE RAS and EMBS International Conference on Biomedical Robotics and Biomechatronics (BioRob), 195–200 (2012)
15. Hanson, D. F.: U.S. Patent No. 8,594,839. Washington, DC: U.S. Patent and Trademark Office (2013)
16. Beira, R., Lopes, M., Praga, M., Santos-Victor, J., Bernardino, A., Metta, G., Becchi, F., Saltaren, R.: Design of the robot-cub (icub) head. IEEE International Conference on Robotics and Automation 51, 94100 (2006)
17. Breazeal, C.: Sociable Machines: Expressive Social Exchange Between Humans and Robots. PhD thesis, MIT (2000)
18. Itoh, K., Miwa, H., Nukariya, Y., Zecca, M., Takanobu, H., Roccella, S., Carrozza, M. C., Dario, P., Takanishi, A.: Mechanisms and functions for a humanoid robot to express human-like emotions. In: Proceedings of the 2006 IEEE International Conference on Robotics and Automation (2006)
19. Kaneiro, F., et. al.: Cybernetic human hrp-4c: A humanoid robot with human-like proportions. Springer Tracts in Advanced Robotics, 70, 301314 (2011)
20. Kobayashi, H., Hashimoto, T., Hiramatsu, S.: Development of face robot for emotional communication between human and robot. Proceedings of IEEE International Conference on Mechatronics and Automation 54, 2530 (2006)
21. Berns, K., Hirth, J., Schmitz, N.: Emotional architecture for the humanoid robot head roman. Proceedings of IEEE International Conference on Mechatronics and Automation 53, 2150–2155, (2007)
22. Sadoyama, T., Sugahara, T., Hashimoto, M., Yokogawa, C.: Development of a face robot imitating human muscle structures. Journal of Robotics and Mechatronics 19, 324330 (2007)
23. Friesen, W. V., Ekman, P.: Facial Action Coding System. Palo Alto, CA: Consulting Psychologists Press (1978)
24. Tronick, E., Als, H., Brazelton, T.B.: Monadic phases: A structural descriptive analysis of infant–mother face to face interaction. Merrill-Palmer Quarterly of Behavior and Development 26(1), 3-24 (1980)
25. Izard, C.: Innate and universal facial expressions: Evidence from developmental and cross-cultural research. American Psychological Association 115, 288–299 (1994)
26. Ekman, P., Friesen, W. V.: Facial action coding system: A technique for the measurement of facial movement. Consulting Psychologists Press (1978)
27. Marcos, S., Gómez-García-Bermejo, J., Zalama, E.: A realistic, virtual head for human-computer interaction. Interacting with Computers 22, 176–192 (2010)
28. Loza, D., Marcos, S., Gómez-García-Bermejo, J., Zalama, E.: Application of the FACS in the Design and Construction of a Mechatronic Head with Realistic Appearance. Journal of Physical Agents 7(1), 30-37 (2013)
29. Montagne, B., Kessels, R. P. C., Frigerio, E., de Haan, E. H. F., Perrett, D. I.: Sex differences in the perception of affective facial expressions: Do men really lack emotional sensitivity? Cognitive Processing 6(2), 136-141 (20015)
30. Pierce, B., Kuratate, T., Vogl, C., Cheng, G.: Mask-Bot 2i: An active customisable Robotic Head with Interchangeable Face. In: 12th IEEE-RAS International Conference on Humanoid Robots, 520-525 (2012)
31. Cid, F., Moreno, J., Bustos, P., Nez, P.: Muecas: A Multi-Sensor Robotic Head for Affective Human Robot Interaction and Imitation. In: Sensors 14(5), 7711-7737 (2014)