EMYS—Emotive Head of a Social Robot

**Jan Kędzierski · Robert Muszyński · Carsten Zoll ·
Adam Oleksy · Mirela Frontkiewicz**

Abstract This paper presents the design, control, and emotion expressions capabilities of the robotic head EMYS. The concept of motion control system based on FACS theory is proposed. On the basis of this control system six basics emotions are designed for EMYS head. The proposed head shapes are verified in experiments with participation of children aged 8–12. The results of the experiments, perception of the proposed design, and control system are discussed.

1 Introduction

The robotic head EMYS (EMotive headY System) has been designed and built within the EU FP7 LIREC project [33].

J. Kędzierski · R. Muszyński (✉) · A. Oleksy · M. Frontkiewicz
Institute of Computer Engineering, Control and Robotics,
Wrocław University of Technology, ul. Janiszewskiego 11/17,
50-372 Wrocław, Poland
e-mail: robert.muszynski@pwr.wroc.pl

J. Kędzierski
e-mail: jan.kedzierski@pwr.wroc.pl

A. Oleksy
e-mail: adam.oleksy@pwr.wroc.pl

M. Frontkiewicz
e-mail: mirela.frontkiewicz@pwr.wroc.pl

C. Zoll
Group for Interdisciplinary Psychology, University of Bamberg,
Kapuzinerstrasse 16, 96045 Bamberg, Germany
e-mail: carsten.zoll@uni-bamberg.de

The LIREC's objective is to provide a technology of designing robots—human companions. A robotic companion is a social robot whose aim is to accompany people in various places and situations. A crucial capability of social robots is that of interacting with humans [20, 32]. From the human-robot interactions (HRI) experiments point of view, for expressive social robots their motion should be human-friendly, legible, and expressive [44, 50, 52, 54]. The overall impression of motion coordination and smoothness have primary significance for the perception of a social robot. Most of all, a social robot should be able to communicate in human way, i.e. to use human-specific verbal and non-verbal communication means, in particular to receive and express emotions. In the process of robot-human communication, the robot face plays a vital role [15], and the facial expressions are natural and intelligible means of expressing emotions. Beyond that, the social robot should possess a character and personality, noticeable by humans.

Commercial products of this kind have recently appeared on the market (iCat [49], Aibo [55], Wakamaru [41]). Sophisticated humanoid robots like Asimo [43], MDS [38], WE-4RII [42], IURO [1], Actroid DER2/DER3 [31] may ultimately find their way to our homes. Advanced prototypic social robots have been built at MIT, mainly dedicated to the research on human-robot interactions. The social robots Kismet [9], Leonardo [8], and Mertz [2] can make a short conversation with a human, recognise their owner, and notice and record people appearing in their neighbourhood. Another construction of this type is the robotic head Samuel [11] built at Wrocław University of Technology and serving HRI experiments. The exemplary heads of social robots are shown in Fig. 1. Their essential features are summarised in Table 1.

The head EMYS is a mechanoidal type robot, which has been designed for HRI experiments (see Table 1 for its basic
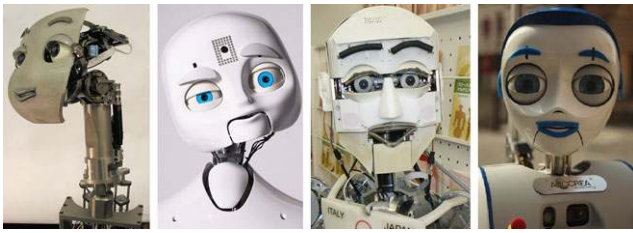
Fig. 1 Social robot heads: Mertz, MDS Nexi, WE-4RII, and IURO

Table 1 Exemplary social robots heads characteristics

Robot	Feature	Description
Mertz	Total DOF	12
	Neck DOF	2
	Camera	2 × Point Grey OEM Dragonfly
	Other sensors	GN Netcomm array mic
Nexi	Total DOF	19
	Neck DOF	4
	Camera	2 × colour CCD
		Active 3D IR
	Other sensors	speech mic
		4 × localisation mic
WE-4RII	Total DOF	27
	Neck DOF	4
	Camera	2 × colour CCD
	Other sensors	2 × mic
		26 × force sensors
		temperature sensor
		smell sensor
IURO	Total DOF	24
	Neck DOF	3
	Camera	2 × stereo CCD
	Other sensors	Kinect
EMYS	Total DOF	11
	Neck DOF	3
	Camera	colour CMOS
	Other sensors	speech lapel mic
		Kinect

parameters). The head was constructed with the aim of being mounted on top of a wheeled balancing platform,[1] equipped with two arms and hands WANDA [28], which all together constitutes the social robot FLASH—the flagship robotic companion of Wrocław University of Technology [27, 29, 34]. An overall view of FLASH is presented in Fig. 2.

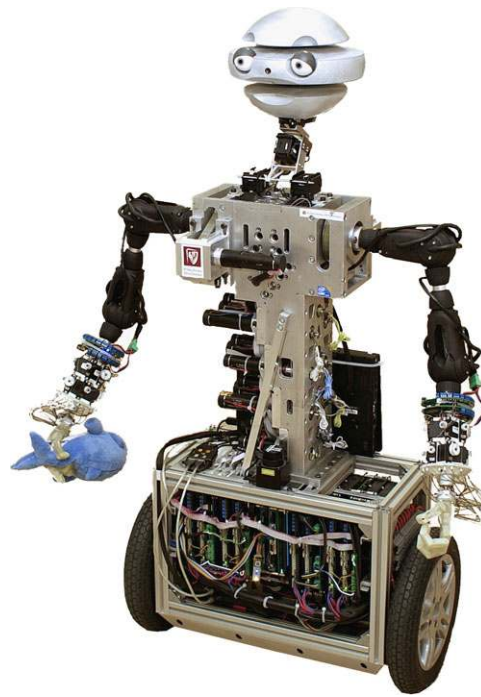[1]However it can serve as a stand alone social robot as well.

Fig. 2 FLASH: overview

This paper presents the design, control system, basic expression capabilities with overall functionality, and experimental evaluation of EMYS. On the design stage of the head, the perception of its construction by humans, the state of the art (affective computing, uncanny valley phenomenon), and the necessity of providing required functionality were taken into account. The EMYS's control system was designed to be compact, modular, and flexible, and to make the head functionalities feasible. It complies with the paradigm of the three-level control architecture [3, 10, 22], provides a hardware abstraction interface to the head, and is realised in the form of a complete operating system. Its lowest level realises the basic hardware abstraction, integrates the low-level motion controllers, the sensor systems and the feed system. The middle level implements robot competencies. In particular, to enable task-oriented control the Facial Actions Coding System (FACS) [18] has been applied here for describing EMYS's facial expressions. The high level, depending on specific needs, may incorporate a dedicated decision system, state automaton, or a comprehensive program system simulating some human mind functionalities, and thus work autonomously. Nonetheless, this level not always needs to act autonomously, sometimes being assisted by a human. The overall head functionality and distinctiveness of emotions has been evaluated in experiments.

The paper is composed in the following way. Section 2 introduces the concept of EMYS and sketches its design and functionality. Section 3 is devoted to the EMYS's motion control system. Section 4 describes HRI experiments with EMYS. Section 5 concludes the paper.

2 Design and Functionality

2.1 Head Characteristics

The head EMYS is a three discs construction, equipped with a pair of movable eyes with eyelids, and mounted on a movable neck. These elements, together with a speaker, constitute the EMYS's system of actuation on the environment. EMYS uses the speaker for speech purposes, which can be synthesised or prerecorded and replayed—this allows the head to speak with different voices. To perceive the environment, the head has been endowed with a colour CMOS camera Logitech Sphere AF [35], Kinect sensor [39], and a lapel microphone for speech recognition purposes. This allows the head for visual perception of the environment, eye-tracking of objects and humans, establishing and maintaining eye-contact with humans, paying attention, expressing emotions, speech recognition and speaking. All these functionalities are controlled by the head control system, and can be executed autonomously.

2.2 Concept

On the stage of the concept elaboration of a new robotic head EMYS, several head's layouts have been analysed, with different designs and mobilities. In each layout the head was mounted on a neck and equipped with eyes with eyelids. Occasionally, additional elements like a tongue or lips were considered. Each concept of the head was evaluated by means of computer graphics visualisation, taking into account its perception by humans, expression capabilities, and functionality. Finally, on this basis, we have chosen the head of a turtle-like appearance, consisting of three movable discs, since it was the best perceived construction, concurrently allowing to maintain desirable functionality. This design has been inspired by characters from cartoon and movie series *Teenage Mutant Ninja Turtles*[2] (see Fig. 3) [30].

2.3 Head Structure

To make the project realisable, it has been decided to limit the head complexity and to equip it with the total number

Fig. 3 Teenage Mutant Ninja Turtle and EMYS

of 11 joints (3 in the neck, 2 in the eyes, 4 in the eyelids, and 2 in the upper and lower discs). The final head's layout revealing its movement capabilities is displayed in Fig. 4. The head's joints deployment is shown in Fig. 5.

The main movable elements of EMYS are its upper and lower disks. They are supposed to imitate the human raising eyebrows and dropping jaw, respectively. Each of them has 1 DOF. The middle disk, hosting a vision camera, is not movable independently. From the viewpoint of facial expressions the eyes need to be perceived together with eyelids and eyebrows. In the design of EMYS the eyelids are mounted on the eyeballs, and can open and close (1 DOF each). The eyeballs and eyelids can turn around the horizontal axis (1 DOF each) that intensifies essentially the ex-

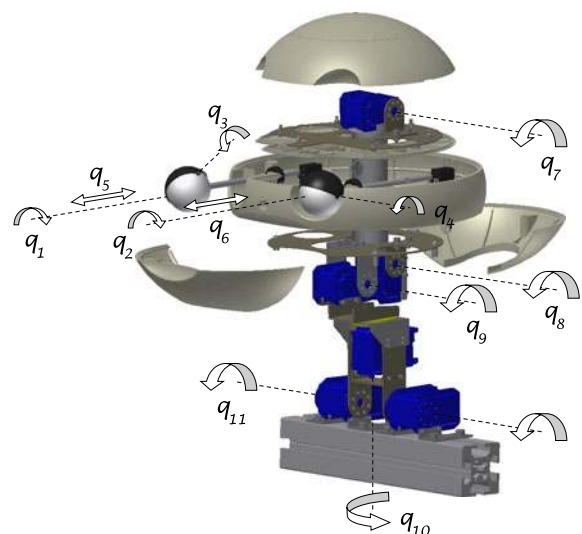**Fig. 4** EMYS: snapshots with movement capabilities

Fig. 5 EMYS: deployment of joints

pressed emotions. This means, for instance, that to express sorrow or sadness they are turned outside, whereas when expressing anger or frustration they are turned inside. Another remarkable ability of pulling out the eyeballs (1 DOF each) enhances the head's extensibility when showing surprise.

For the reason of not to perturb excessively the balancing motion of the FLASH platform, the head should be possibly light weight. Also, for aesthetic reasons, its size should fit the remaining FLASH's components. This being so, the head's supporting construction has been made of aluminium, and all the shells of the head have been printed with the use of the rapid prototyping SLS technology. The head's bearer has the form of a pipe, holding all three disks. The bearer is screwed to a 3 DOF tilt-pan-tilt mechanism constituting the neck.

2.4 Movements Representation

To construct the high level interface of a robotic head, one needs a method for coding the facial expressions. Several such methods are available, see e.g. [4, 18, 48]. In EMYS the FACS system [18] was chosen, dedicated to the description of human facial expressions, widely used by psychologists and animators. The application of FACS system allows EMYS to follow biological rules applied by humans. The advantages of FACS are confirmed by its applications in control of other social robots heads [5].

The work of Ekman and Friesen [18] describes, in terms of action units of the face, how humans universally express the six basic emotions, along with psychological and physical descriptions of their reflection in the human beings. Other researchers have investigated these basic emotions in robots and virtual agents as well [6, 7, 20]. The six basic emotional expressions considered by Ekman and Friesen are the following: surprise, disgust, fear, anger, happiness and sadness. After analysing the background that specifies the recognition of emotions in both, humans and cartoon animations, we reduced the necessary expressive features into a more simple set, that can be easily mapped into EMYS. It was a requisitive step to be performed, since our head suffer from the lack of complexity: the human face is much more complex than EMYS, however the use of FACS is still possible here. We will use the concept of expressive effectors to refer to the physical means and degrees of freedom of the embodiment that can be used for expressing emotions.

In accordance with FACS, each facial expression is decomposed into specific Action Units (AUs). Each of the AUs originally define elementary movements of a single muscle or a group of them, taking values from the range [0, 1], participating in a change of the countenance. Obviously, when applied to a robotic face, the list of activated AUs serves solely as a description of the facial expressions, without attributing to them any explanatory power.

Table 2 EMYS: Action Units with associated joints movements

Action Unit	Movement	Joint
AU(1 + 2)	up	Upper Disc
AU(4)	down	
AU(43)	close	Eyelids
AU(45)	blink	
AU(46)	wink	
AU(1)	turn outward	Eyebrows
AU(2)	turn inward	
AU(5)	pull out	Eyes Trans
AU(17)	up	Lower Disc
AU(25)	down	
AU(51)	turn left	Neck Pan
AU(52)	turn right	
AU(53)	turn back	Neck Lower Tilt
AU(54)	turn fore	
AU(57)	lower fore, upper back	Neck Lower &
AU(58)	lower back, upper fore	Upper Tilt

Coding the EMYS's expressions by means of FACS leads to phrasing them as sets of AUs. This means that each individual AU has to be mapped into the head's joints movements. This task may not be easy, mainly because in comparison with the human head the EMYS's movement capabilities are very restricted on one hand, whereas they include certain movements that cannot be performed by a human, on the other hand. For this reason, individual AUs identification needs to be conducted for EMYS. Table 2 contains all AUs identified for EMYS, together with a list of associated head's joints movements. All these AUs can be naturally divided into two groups: the AUs executed identically by EMYS and the human head (like head turning, nodding, and eyes blinking), and thus being easily interpreted, and the AUs that have to be interpreted separately (like pulling out the eyes interpreted as the eyes opening wide or lifting the upper disc interpreted as lifting eyebrows). All EMYS single activations of AUs are illustrated in Fig. 6.

Having defined the AUs for EMYS, we have described all EMYS's facial expressions as combinations of them. We focus on seven basic facial expressions described in [17], i.e. neutral, anger, disgust, fear, joy, sadness and surprise. Their definitions in terms of AUs for EMYS are the following:

1. neutral—AU(0),
2. anger—AU(2 + 4 + 43 + 57),
3. disgust—AU(4 + 17 + 43),
4. fear—AU(1 + 2 + 17 + 58),
5. joy—AU(1 + 2 + 25),
6. sadness—AU(1 + 4 + 54),
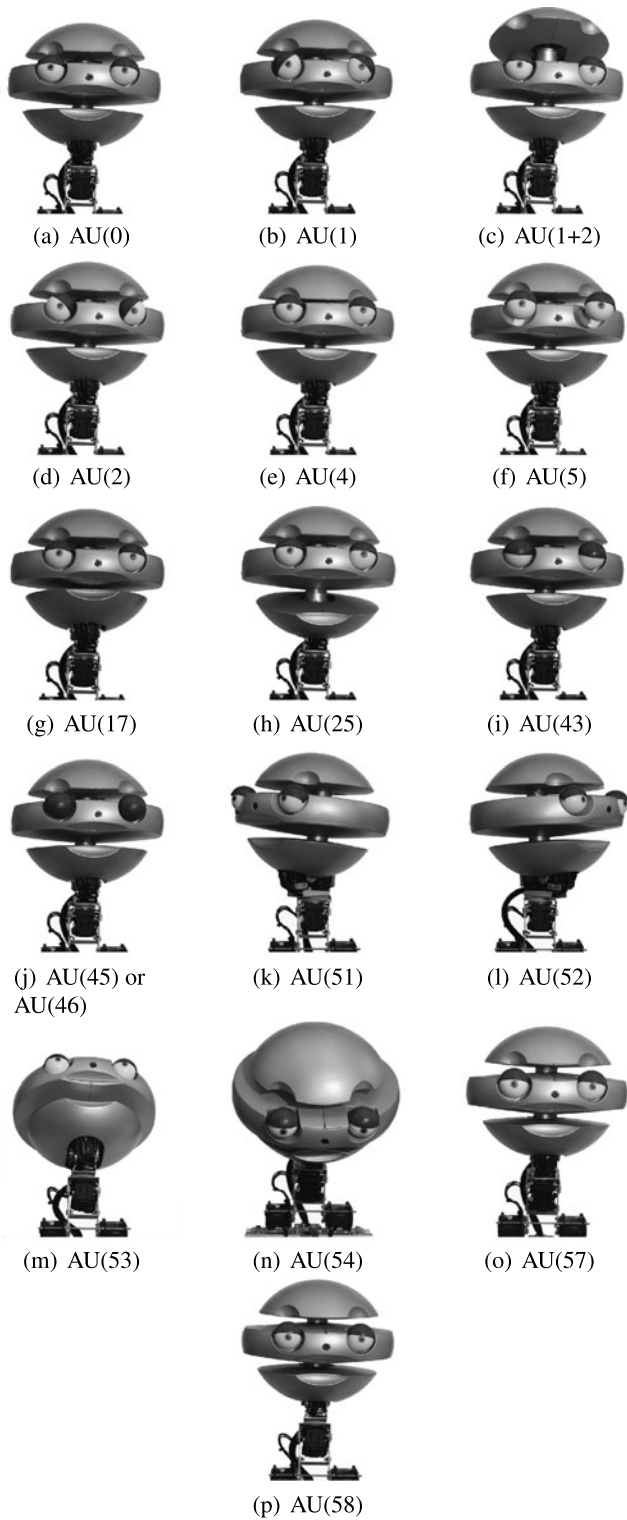7. surprise—AU(1 + 2 + 5 + 25 + 58).

(a) AU(0) (b) AU(1) (c) AU(1+2)

(d) AU(2) (e) AU(4) (f) AU(5)

(g) AU(17) (h) AU(25) (i) AU(43)

(j) AU(45) or AU(46) (k) AU(51) (l) AU(52)

(m) AU(53) (n) AU(54) (o) AU(57)

(p) AU(58)

Fig. 6 EMYS: Action Units

The snapshots of all EMYS's basic facial expressions are demonstrated in Fig. 7. It should be borne in mind that these static pictures do not fully reflect the actual expressiveness of EMYS, since it can be quite difficult to correctly recognise facial expressions outside a context on the basis of a
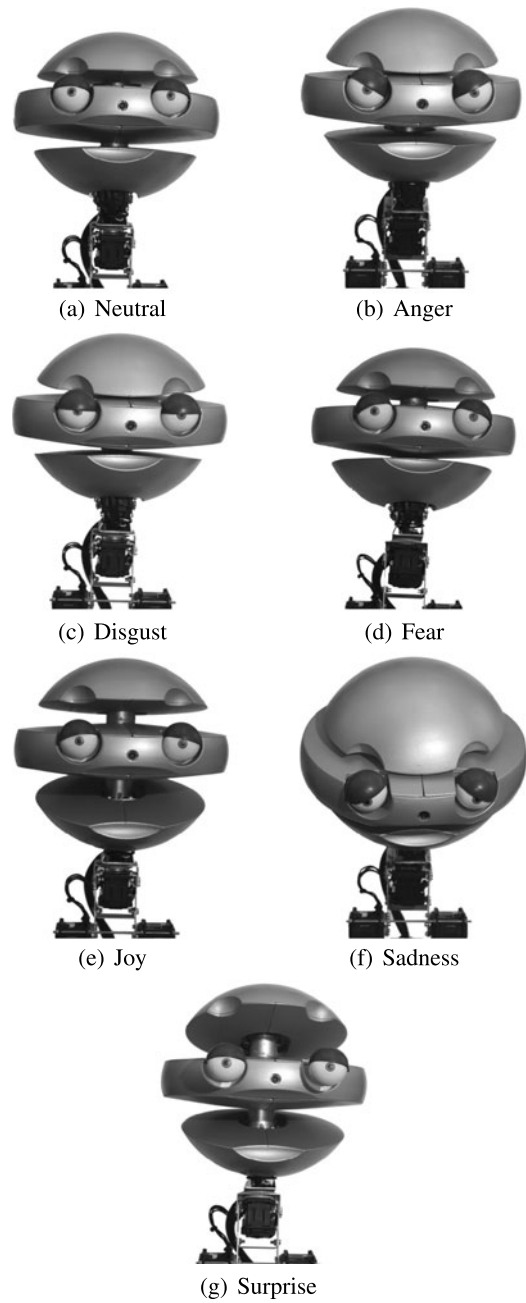

(a) Neutral (b) Anger

(c) Disgust (d) Fear

(e) Joy (f) Sadness

(g) Surprise

Fig. 7 EMYS: basic facial expressions

snapshot only. Each expression execution utilises the multiphase gesture generator model [56]. In this approach gesture expressions are divided into three phases: preparation, stroke, and retraction. In Fig. 7 the expressions are shown in the most meaningful and effortfull phase: the stroke phase.

3 Drive and Control

3.1 Drive System and Hardware Layer Controller

EMYS's neck movements need to be smooth and not very fast. As we have already said, the head should addition-

ally be able to realise the functions of maintaining the eye-contact with a human, eye-following objects, looking around etc. Leaving apart the other functions we shall concentrate on expressing emotions by EMYS. This feature is obtained by utilising 4 high quality Robotis digital servomotors provided by Dynamixel. Two RX-64 servomotors cooperatively drive the neck lower tilt axle, another realises the pan movement, and the 4th, smaller RX-28 servomotor is responsible for the neck upper tilt, see Fig. 5. Two other RX-28 servos drive the upper and lower disks. The servo communication is based on the Dynamixel protocol [51]. Each servo is identified by its individual ID number. Their basic movement parameters (maximum torque, speed, range of movement) are configurable. Also, the servos' actual state (temperature, overdriving or overheating alarm, power supply voltage) can be accessed via the protocol.

Contrary to the neck, the movements of eyeballs and eyelids need to be rapid. For this reason, in order to open/close eyelids and turn the eyeballs, the high performance analog micro servomotors Hitec HS-65HB have been employed, controlled by PWM signals. The eyeballs pull out has been accomplished with use of Alps high-speed, motor driven slide potentiometers.

Since EMYS's joints are driven by different types of servomotors requiring different control methodology, in order to facilitate the control, all the control inputs have been uniformly defined at some level of the device independent abstraction. To this objective, an additional module unifying the controls has been introduced. As a result, all the servos, no matter which type, can be accessed using the same Dynamixel protocol scheme. This control function is accomplished by the hardware layer controller which steers both the micro servomotors and the potentiometers and implements the Dynamixel protocol. The controller is based on Freescale HC9S12A64 microcontroller [21] which generates control signals for all 4 micro servomotors and 2 potentiometers in the PID control loop. The block diagram of the controller hardware is depicted in Fig. 8.

The head's control is accomplished by sending appropriate Dynamixel commands from the main computer to the head's drives. On the basis of the data provided by the gesture generator, the computer calculates the head's joint trajectories, and after expressing them in terms of the Dynamixel protocol commands, sends them to the servomotors. These operations are performed on a PC computer. Here a PC computer equipped with Intel Core i7-2640M processor (2.8 GHz, 8 GB RAM, 4 MB Cache and 120 GB SSD), running under Linux Ubuntu 10.04 or MS Windows 7 operating systems is utilised. Such hardware resources allow to run all vision and audio system components, together with the movement control. As a data carrier between the PC computer and the servomotor controllers the serial interface RS-485 is applied.
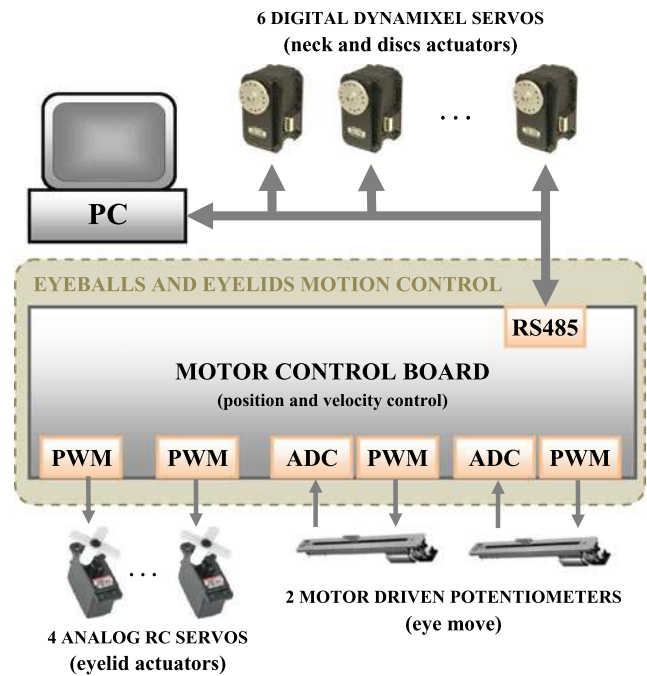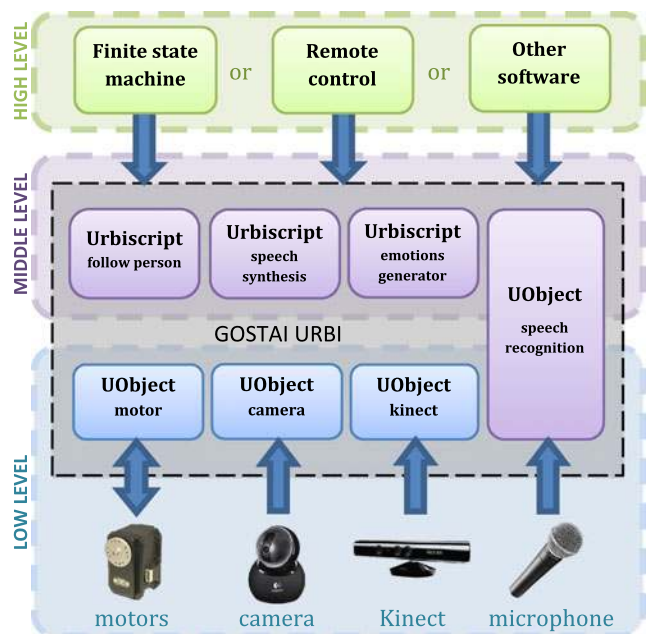

Fig. 8 Hardware layer controller

Fig. 9 EMYS's control system architecture

3.2 Integration Layer Controller

An integration layer of EMYS's control system architecture is the open source Urbi software [24], created by Gostai. It is employed on the lowest levels of the three-level architecture (see Fig. 9). Urbi SDK is a fully-featured environment to orchestrate complex organisations of robot components. It relies on a middle-ware architecture that coordinates com-
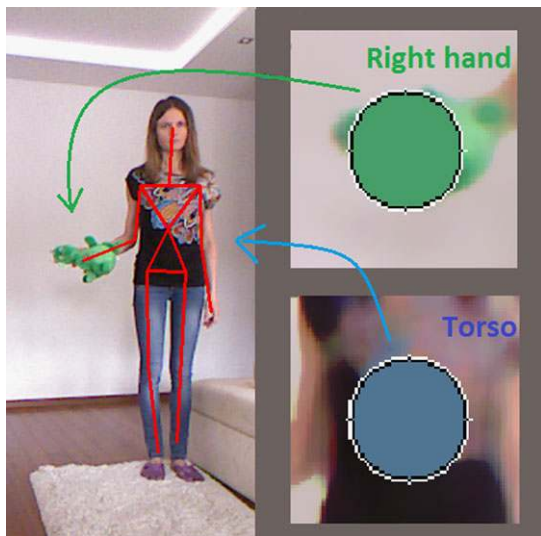
Fig. 10 Exemplary image segmentation with Kinect

Fig. 11 Head camera image processing examples

ponents named UObjects. This software allows for the dynamic uploading of modules, that enable accessing the robot hardware or realise certain competencies. URBI provides communication means and integrates all of them.

The robot can be programmed in the script language urbiscript by uploading instructions to the Urbi engine through a client application. Urbiscripts supports and emphasises parallel and event-based programming. A dedicated collection of modules and scripts has been used in order to develop an integrated programming interface for EMYS.

Communication with EMYS's motors is achieved by means of a module able to communicate through serial ports. This module realises the communication using the Dynamixel protocol. Additionally, a module relying on SDL library [53] has been provided, enabling to remotely control the robot with the help of a joystick. To further advance the control process, the individual joint movements can be grouped and preprogrammed as the components of facial expressions, what in EMYS case has been achieved by means of the FACS system [18], as described in the next subsection.

The realisation of competencies referring to visual image processing is based on OpenNI [47] and OpenCV [46] libraries. The former makes possible to utilise the Microsoft Kinect device [39], playing the role of an advanced motion sensor. The developed Urbi module, exploiting Kinect functionalities, enable detection of the human silhouette as well as provide information on distances between prescribed elements of the processed image. Such data allows to localise a human in the robot neighbourhood, define positions of his/her extremities, and perform the image segmentation (see Fig. 10). OpenCV library has been used to create a number of image processing modules based on image from the camera mounted in EMYS. It allows to detect: human faces,
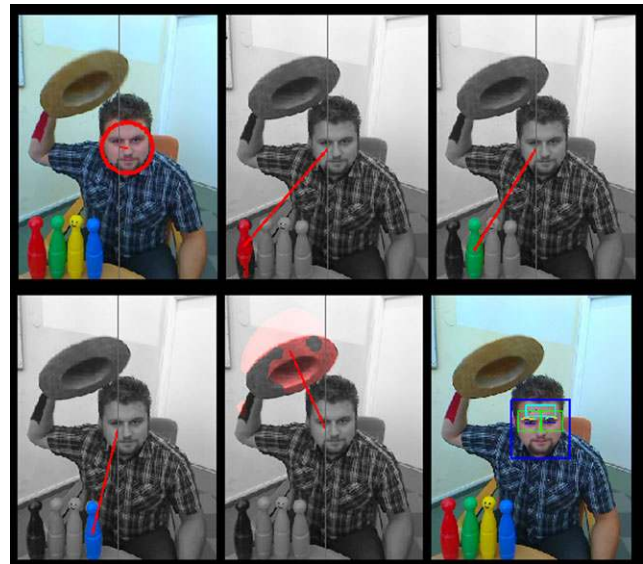
colours, movement or face features (see Fig. 11). Beside the mentioned, a collection of modules realising basic operations on the image has been provided.

The auditory competencies, based on Loquendo [36], Microsoft SAPI [40] and SDL library, have been implemented as three separate modules responsible for speech recognition, speech synthesis and replaying of audio files. These modules establish one of the most important human-robot communication channels, by speech and sound.

Also a learning competency has been developed, relying on algorithms coming from OpenCV library. The learning module implements the k-nearest neighbours classifier. This module can serve for information acquisition and classification of the robot environment. In conjunction with other available modules, it is possible to teach the robot a favourite colour of its user dress. All these modules have been provided in the form of dynamically loaded plugins, complemented by scripts written in the language urbiscript, that facilitate the employment of the modules in the process of creating scenarios of robot actions. Specific functions, competencies and parameters can be accessed by API (Application Programming Interface) as the structure `robot`, with exemplary fields listed bellow:

```
robot.body.neck.head.ExpSmile();
robot.body.neck.head.Say("hello world");
robot.video.camera.image;
robot.video.humanDetector.position;
robot.audio.recognition.result;
robot.ml.colorLearning.LearnFromRightHand();
```

3.3 AU Controller

To exploit the head movement capabilities, it is advantageous to equip it with a higher level controller aimed pri-
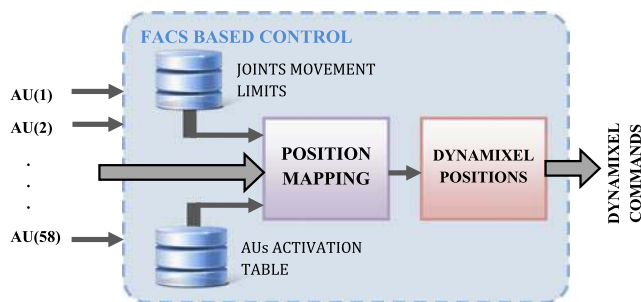
Fig. 12 AU controller

Fig. 13 Exemplary FSM for an EMYS experiment

marily at the expression of emotions. The task of such a controller consists in providing a kind of abstraction separating a head animator from the movement execution issues. Instead of dealing with the problem how to realise in very detail a required movement, thanks to this controller the animator is allowed for a task-oriented control of the head. Such an interface facilitates the description of emotions to be expressed, and enables the application of ready-to-use methodologies for face animation [5, 45, 48].

To this aim we have elaborated a FACS based controller, called AU controller. This controller (Fig. 12) transforms activated AUs, which are the controller input, to corresponding head's joints positions. It is done on the basis of the joints movements limits and the AUs activation table, which are implemented in the controller. The overall functionality of this controller for the EMYS head can be characterised by a set of functions, i.e.

$$q_{ud} = f(AU(1+2), AU(4))$$
$$q_{ld} = f(AU(17), AU(25))$$
$$q_{el} = f(AU(43), AU(45), AU(46))$$
$$q_{eb} = f(AU(1), AU(2))$$
$$q_{ep} = f(AU(5))$$
$$q_{np} = f(AU(51), AU(52))$$
$$q_{nlt} = f(AU(53), AU(54), AU(57), AU(58))$$
$$q_{nut} = f(AU(57), (AU(58)),$$

where q_{ud}, q_{ld}, q_{el}, q_{eb}, q_{ep}, q_{np}, q_{nlt}, q_{nut} are the joint positions of the upper and lower discs, the eyelids and eyebrows, the eyes protrude, the neck pan, and the neck lower and upper tilts, respectively. Since the full form of the above functions is rather long and complicated, below only an exemplary function (for the upper disc movement) is displayed

$$q_{ud} = q_{ud\_nom} - (q_{ud\_nom} - q_{ud\_max})AU(1+2)$$
$$- (q_{ud\_nom} - q_{ud\_min})AU(4),$$

where q_{ud_nom}, q_{ud_min}, q_{ud_max} are the nominal, minimal, and maximal values of the described joint, respectively. After calculating the joints positions they are transferred to the low level controller by means of Dynamixel protocol.
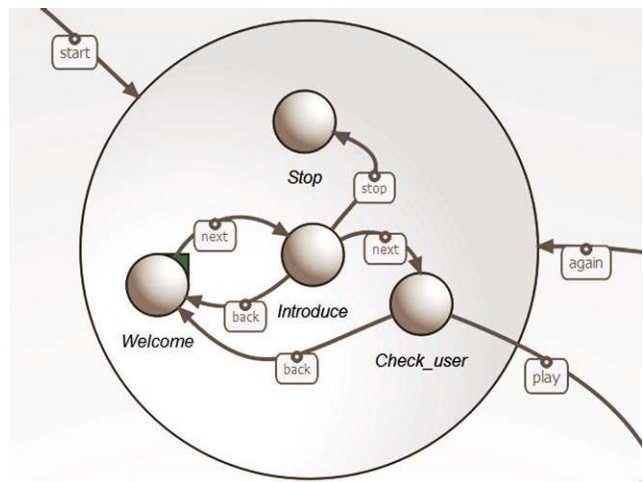
3.4 Task-Oriented Layer Controller

Depending on specific needs the task-oriented layer controller may incorporate a dedicated decision system, a state automaton, or a comprehensive program system simulating some human mind functionalities. Substantially, this should be a controller working autonomously, but in general, there is not always such a need—in some cases the controller can be assisted by a human. For this purpose, Gostai Studio software [23] can be applied. This software has served as a tool for the implementation of a hierarchical finite state machine. The graphical interface of this application is aesthetic, intuitive, and easy to use. It allows for a relatively quick implementation of scenarios not requiring long term simulation of processes running in human minds.

Simple robot behaviour is created as a node. Connection between behaviour can be done by creating transitions between nodes. All transitions include conditions for the changes in behaviour. It is also possible to create nodes inside another nodes. In this way even most complex behaviour graphs, such as an experiment scenario, is easy to create. Exemplary finite state machine for an experiment with EMYS implemented in Gostai Studio is shown in Fig. 13.

4 Experiments

4.1 Description of the Study

In order to examine both children's engagement in the interaction with EMYS and whether the children are able to decode the intended expressed emotions correctly, an experiment was conducted. The design of the experiment resulted from the analysis of the results from two pilot experiments, which were conducted to gain insights into the interaction behaviour, that can be expected by the child subjects and

Fig. 14 Affect matching task

whether the children show a significant variance in emotion recognition rates, which could give hints for the further development of EMYS's ability to display certain emotions. We have chosen child instead of adult subjects out of three reasons. Firstly, it can be expected that children show lower emotion recognition rates than adults. Thus, if the rates are sufficient for children, they will also be for adults. Secondly, it is known that children of different age also show different emotion recognition abilities [26]. If we find corresponding results for emotion recognition abilities regarding EMYS this can account for high validity of the experimental procedure. Thirdly, since children tend not to mask or cover their emotions during interactions, we expected to get more and more valid results on EMYS's ability to engage human users in interaction (which has to be confirmed in subsequent studies with adult subjects, of course).

The experiment was conducted in a primary school in a small village near Wrocław/Poland and involved originally 48 schoolchildren of which 3 had to be removed completely from the final sample because of malfunctions of EMYS.[3] The remaining 45 subjects (Ss) were aged 8 to 12 years (AM = 9.9 years, SD = 1.41). Overall, the sample consists of 18 boys and 27 girls.

The robotic head was programmed to operate autonomously and to provide two game scenarios—each subject went through both scenarios. In the first one, called "imitation task", EMYS showed emotional facial expressions and asked the children to repeat them. Overall six basic emotions have been used: anger, disgust, fear, happiness, sadness, and surprise. In the second scenario ("affect matching task"), the robot expressed the same six emotions in a different order and asked the children to show a toy coloured corresponding to the expression (see Fig. 14). The coloured toys were stored in different boxes on which the respective emotion was written.

With its implemented vision system, EMYS was able to recognise the colour of the toy and to react accordingly, i.e. praising or dispraising. Since EMYS was only able to detect four different colours with sufficient security, we decided to assign three colours to different emotions (red: anger, blue: sadness, green: joy) whereas the fourth toy (yellow) was used as "none of the other emotions". After the two scenarios an interview was conducted, which implied a third task: The children watched the video-taped interaction of the imitation task scenario and were asked to name the emotions shown by EMYS ("affect description task"—for further information on affect description and affect matching tasks see [26]). The duration of the interaction experiment with a single child was about 5–8 minutes. During the experiments EMYS talked a male voice.

4.2 Description of the Psychological Analyses

Five types of data were gathered:

1. assessments of EMYS' facial expressions of emotions (see above),
2. interview data of the children collected after the interaction,
3. interaction data (video-taped),
4. a "Big Five" personality self assessment for the child subjects [37],
5. a self-developed test of emotional face recognition.[4]

Interaction data were analysed with the help of a video analysing form in which several behavioural variables could be noted by the observers: body posture, number of verbal and non-verbal utterances and their direction (towards EMYS or towards experimenter), emotional expressions, activation and gaze direction. For validation, 8 subjects were coded by three different observers and interrater reliability turned out to be good (75 %) [25].

The behavioural variables were subsequently used to calculate measures for the child's engagement in the interaction. We calculated different engagement scores: A score for positive engagement (engagement score pro) was calculated out of the values for communication towards EMYS, activation and body posture. Furthermore a score for negative engagement (engagement con) results from the sum of measures of communication towards experimenter and gazing elsewhere (e.g. to the camera or the experimenter). Finally, a score for overall engagement was calculated as the difference between the former two.

Note, that positive engagement does not mean that the interaction with EMYS is an overall emotionally positive

---

[3]In some cases, data from the remaining 45 subjects is missing, mainly due to malfunctions of EMYS, but the subjects were kept within the sample as long as this had no obvious impact on the interaction. Thus, sample size will differ a little for different statistical analyses.

[4]Which was conducted in order to get information on how the children were able to recognise the six emotions expressed by humans. To achieve this, we took 12 (6 × 2) pictures from [17] and presented them to the children as a multiple-choice questionnaire with the distractors being the other five emotions used in the study.

Table 3 Perceived personality of EMYS (from −1 = "not at all" to 1 = "absolutely")

Extroversion	0.96
Agreeableness	0.51
Conscientiousness	0.53
Emotional Stability	0.80
Openness	0.87

Table 4 Emotion recognition rates

Emotion	Affect description	Affect matching	Affect description & matching[a]	Ekman Emotion recognition questionnaire
Anger	97.8 %	97.6 %	**97.7** %	**93.0** %
Surprise	**46.7** %			**92.1** %
Happiness	15.9 %	22.5 %	**19.2** %	**91.9** %
Sadness	91.1 %	95.3 %	**93.2** %	**79.7** %
Disgust	**13.6** %			**78.1** %
Fear	**68.9** %			**64.0** %

[a]Mean where feasible

experience, but that the subject is very focused on the interaction with EMYS. Correspondingly, negative engagement doesn't necessarily mean that the subject experiences mainly negative emotions, but that it is not solely focused on EMYS, but also on the environment (e.g. the experimenter).

4.3 Results and Discussion

The experiment results have been analysed and discussed considering the aspects formulated below as questions to be answered.

How do Subjects Assess EMYS and the Interaction with It? Regarding sex, all subjects thought that EMYS is male. By far the most of the subjects made their decision because of EMYS's voice. 96 % children reported their desire to interact again with the robot. 33 % of the subjects thought that EMYS has emotions, 21 % didn't know, while 46 % didn't think so. The subjects were asked to rate EMYS's personality according to the "Big Five" personality factors [12] at a "yes"/"no"/"neither nor"-level in a child-appropriate form. Table 3 shows the overall results (mean values over all subjects). The results show that EMYS is perceived as extremely extroverted and open, as agreeable and as conscientious and emotionally very stable. This rather positive personality assessment seems especially important for our experimental purpose, since humans tend to exhibit rather negative attitudes when thinking of social robots. EMYS's perceived personality might have facilitated the interaction with it for the child subjects [19].

To Which Degree are the Ss Able to Recognise the Emotions Shown by EMYS and do Differences Exist Between Single Emotions? During the experiment, EMYS showed six different emotions to the child subjects in each task. Table 4 shows the mean recognition rates for each and over both tasks. Note, that surprise, disgust and fear do not appear in the affect matching task since they were assigned the same yellow toy. The last column in the table shows the recognition rates for all six emotions in the self-developed emotion recognition questionnaire based on the Ekman & Friesen photographs [17].

Are There Differences Between the Recognition Rates of Emotions Shown by EMYS and Humans? Regarding the Ekman based questionnaire it is notable that recognition rates are quite high in general (with the exception of fear),

which can be interpreted as a sign for a high validity of the questionnaire.

Anger is the best recognised emotion both when interacting with EMYS and in the Ekman questionnaire. While a literature review [26] shows that the best recognised emotions by children are happiness, sadness, and anger, in the case of our experiment happiness was difficult for the children to recognise when interacting with EMYS. This is probably due to the fact that EMYS is not able to show the most salient expression for happiness—raise mouth corners—and children tend to evaluate expressive information from the mouth region first (see [13]). It can be assumed that the low recognition rates for disgust are—apart from the fact that this is a more complex emotion—also due to the limited abilities of EMYS to modify its emotional expression in both the mouth and the nose region. A salient sign for disgust among humans is wrinkling the nose (see [17]) which cannot be performed by EMYS in its current state.

On the other hand, EMYS's expression of sadness was recognised extremely well by the children, significantly better than in case of the Ekman & Friesen examples. This could be due to EMYS's cartoonlike exaggerated facial expression facilitates, especially its ability to lower the eye lids strikingly. The recognition rate for fear corresponds to the one found in the Ekman questionnaire and surprise was recognised significantly worse when displayed by EMYS.

Which Factors Influence the Overall Recognition Rates of Emotions Shown by EMYS? While gender specific significant correlations could not be observed, the age of the subjects has an influence on success in decoding EMYS's facial expressions ($r = 0.36$, $p = 0.02$). The older the subjects are, the better they are in decoding EMYS's emotional expressions. The ability to discriminate different emotions is a developmental process (e.g. [16]). Surprisingly, in our case no connection between age and success in decoding human emotional expressions (Ekman questionnaire) could be observed, which could be due to the overall high recognition rates in the Ekman questionnaire.

Engagement, activation and the subjects' emotions during the interaction have no impact on the emotion recognition rate. This could be due to the fact that emotion recognition seems to depend mainly on EMYS's "physiognomy" and is less dependent from other factors, which can be interpreted as a sign of high validity of the experimental design.

Regarding the personality of the subjects, two of the "Big Five" factors correlate with the recognition rate: agreeableness and neuroticism. The negative correlation with agreeableness ($r = -0.36 p = 0.02$) could be due to the fact, that agreeable Ss are more focused on the demand characteristics of the experimental situation compared to EMYS, and thus have difficulties in recognising its emotional expressions. The connection with neuroticism ($r = 0.30$, $p = 0.04$) means that Ss with higher values on the neuroticism scale are more successful in decoding EMYS's emotional expressions. This—on the first sight surprising—result could be due to the higher irritability of neurotic Ss. Maybe neurotic Ss are emotionally affected more easily by emotional expressions of EMYS which could in turn lead to more appropriate affective empathic reactions. These processes might gain additional importance in the current experimental setting since Ss have no context information, that could be used for inferring the emotion state of EMYS (cognitive empathy—for the difference between affective and cognitive empathy see [14]).

Does Interest in Robots Influence the Interaction and the Ss Experience of Interaction with EMYS? Positive correlations exist between interest in robots and the positive engagement score ($r = 0.34$, $p = 0.02$) as well as with the overall engagement score ($r = 0.41$, $p = 0.00$). In the same way, a positive correlation between the question "Did you have fun playing with EMYS?" (answer: "yes") in the questionnaire and the overall engagement score ($r = 0.31$, $p = 0.04$) exists. Children interested in robots (thus also in EMYS) and having fun by playing with EMYS were more willing to interact with it and hence showed more (positive) engagement in interaction with EMYS. Being interested in robots is also associated with higher activation ($r = 0.38$, $p = 0.01$). This can be explained with higher curiosity of interested users with reference to the robots capabilities. Being interested also correlates negatively ($r = -0.30$, $p = 0.05$) with gazes at experimenter and positively with having fun ($r = 0.43$, $p = 0.00$) during interaction with EMYS.

Is There an Impact of the Ss' Personality on the Interaction with EMYS, the Experience of Interaction with and/or the Attitude Towards EMYS? Subjects with higher values on the neuroticism scale look more often at the camera, which is probably due to nervousness and insecurity in the experimental situation ($r = 0.30$, $p = 0.05$). The more subjects are open for experiences, the more fun they report interacting with EMYS ($r = 0.54$, $p = 0.00$). Subjects with higher values on the scale openness utter more often the wish to interact with EMYS again ($r = 0.38$, $p = 0.01$). The more subjects are agreeable, the more fun they report interacting with EMYS ($r = 0.50$, $p = 0.00$) and the more they would like to interact with the robot again ($r = 0.38$, $p = 0.01$). This could be explained by the psychological concept of social desirability: Agreeable subjects try to act/answer according to the expectations of experimenter. Most importantly, subjects' personality has no measurable impact on engagement during interaction with EMYS.

5 Conclusion

In this paper we have described the design, functionality, and control of the robotic head EMYS. To facilitate the head utilisation and allow for its control on some abstraction level the AU controller was designed, enabling the programming and control of the EMYS's facial expressions via FACS coding the EMYS's movements.

The experiments has confirmed the usefulness of the proposed design. The friendliness of the construction, its movement and behaviour capabilities encouraged both children and adults to interact with the robot. Beside the simplicity of the construction it displays more than expected expressiveness and emotion. The emotion expressions were recognised basically good—the low recognition level of happiness and disgust can be assumed to be due to the missing abilities of EMYS to raise and lower the mouth corners and wrinkle the nose. However, at the same time, it can be expected that displaying these emotions in a context will increase the recognition level, what should be investigated.

The design and construction methodology utilised in the EMYS head makes it a compact, self-contained device, thus easily adaptable to different research tasks and fully autonomously operating, if needed. Despite the fact, that this is a prototype, EMYS appeared to be a very reliable, solid, and safe machine. The evaluation shows the potential of the EMYS head for the use with social companions.

The future work should include the investigation of emotion recognition displayed in a context, the influence of perspicuity and dynamism of facial expressions on the recognition level, as well as the influence of other factors (maintaining the eye-contact, paying attention). Simultaneously, the control system of the head should be extended with new components, particularly with regard to perceiving the environment, what its modularity makes possible and rather simple. Considering utilisation of EMYS head as a part of a social companion, the methods of its work synchronisation with the operation of the whole system should be elaborated. Next, its utilisation as a tool in occupational therapy may be assessed.

References

1. Accrea Engineering (2012) Interactive Urban Robot. http://www.iuro-project.eu/
2. Aryananda L (2007) A few days of a robot's life in the human's world: toward incremental individual recognition. PhD thesis, Massachusetts Institute of Technology
3. Aylett R et al (2010) Updated integration architecture. LIREC GROUP Deliverable 9-2
4. Aylett R et al (2011) Facial and body expressions for companions. LIREC GROUP Deliverable 3-5
5. Berns K, Hirth J (2006) Control of facial expressions of the humanoid robot head ROMAN. In: IEEE/RSJ international conference on intelligent robots and systems, Beijing, China, pp 3119–3124
6. Breazeal C (2003) Emotion and sociable humanoid robots. Int J Hum-Comput Stud 59(1–2):119–155
7. Breazeal C (2009) Role of expressive behaviour for robots that learn from people. Philos Trans R Soc Lond B, Biol Sci 364(1535):3527–3538
8. Breazeal C, Brooks A, Chilongo D, Gray J, Hoffman G, Kidd C, Lee H, Lieberman J, Lockerd A (2004) Working collaboratively with humanoid robots. In: IEEE/RAS fourth international conference on humanoid robots, Los Angeles, USA, pp 253–272
9. Breazeal CL (ed) (2002) Designing sociable robots. MIT Press, Cambridge
10. Brooks R (1986) A robust layered control system for a mobile robot. IEEE J Robot Autom 2(1):14–23
11. Budziński R, Kędzierski J, Weselak B (2010) Social robotic head Samuel. In: Scientific papers on electronics, vol 175. Warsaw University of Technology, Warsaw, pp 185–194, in Polish
12. Costa P, McCrae R (1992) Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Psychological Assessment Resources
13. Cunningham J, Odom R (1986) Differential salience of facial features in children's perception of affective expression. Child Dev 57(1):136–142
14. Davis M (1994) Empathy: a social psychological approach. Social psychology. Brown & Benchmark Publishers, Madison
15. DiSalvo C, Gemperle F, Forlizzi J, Kiesler S (2002) All robots are not created equal: the design and perception of humanoid robot heads. In: 4th conference on designing interactive systems, London, pp 321–326
16. Durand K, Gallay M, Seigneuric A, Robichon F, Baudouin J (2007) The development of facial emotion recognition: the role of configural information. J Exp Child Psychol 97(1):14–27
17. Ekman P, Friesen W (1975) Unmasking the face: a guide to recognizing emotions from facial clues. Prentice Hall, New York
18. Ekman P, Friesen W, Hager J (2002) Facial action coding system. Research Nexus Division of Network Information Research Corporation
19. Enz S, Diruf M, Spielhagen C, Zoll C, Vargas P (2011) The social role of robots in the future—explorative measurement of hopes and fears. Int J Soc Robot 3:263–271
20. Fong T, Nourbakhsh I, Dautenhahn K (2003) A survey of socially interactive robots. Robot Auton Syst 42(3–4):143–166
21. Freescale (2012) MC9S12DJ64 device user guide, document number 9s12dj64dgv1/d. http://www.freescale.com
22. Gat E (1998) On three-layer architectures. In: Artificial intelligence and mobile robots, pp 195–210
23. Gostai (2012) Gostai Studio. http://www.gostai.com/products/studio/
24. Gostai (2012) The Urbi Software Development Kit. http://www.urbiforge.org
25. Greve W, Wentura D (1997). Scientific observation: an introduction. PVU/Beltz, in German
26. Gross A, Ballif B (1991) Children's understanding of emotion from facial expressions and situations: a review. Dev Rev 11(4):368–398
27. Kędzierski J, Janiak M (2012) Construction of the social robot FLASH. In: Scientific papers on electronics, vol 182. Warsaw University of Technology, Warsaw, pp 681–694, in Polish
28. Kędzierski J, Janiak M, Małek L, Muszyński R, Oleksy A, Tchoń K, Wnuk M (2009) Foundations of embodied companions. LIREC GROUP Deliverable 6-2
29. Kędzierski J, Małek L, Oleksy A (2012) Application of opensource software for robot companions control system. In: Scientific papers on electronics, vol 182. Warsaw University of Technology, Warsaw, pp 671–680, in Polish
30. Kinde M (1991) Playing with power in movies, television and video games: from Muppet Babies to Teenage Mutant Ninja Turtles. In: Nonverbal communication. University of California Press, Berkeley
31. Company Ltd K (2011) Actroid DER2/DER3. http://www.kokoro-dreams.co.jp/english/robot/act/index.html
32. Li J, Chignell M (2011) Communication of emotion in social robots through simple head and arm movements. Int J Soc Robot 3:125–142
33. LIREC Project (2012) Project website. http://www.lirec.eu
34. LIREC Project (2012) Robot FLASH. http://www.flash.lirec.ict.pwr.wroc.pl/
35. Logitech (2012) Sphere AF—roduct information. http://www.logitech.com/en-us/support/webcams/quickcam-sphere-af?osid=s14&bit=32
36. Loquendo (2012) Loquendo—We speak. We listen We understand. http://www.loquendo.com/en
37. Mackiewicz M, Cieciuch J (2011) The picture-based personality survey for children—a new instrument to measure the big five in childhood. In: 11th European conference on psychological assessement, poster
38. Massachusetts Institute of Technology (2012) Mobile/Dexterous/Social platform. http://robotic.media.mit.edu/projects/robots/mds/overview/overview.html
39. Microsoft (2012) Kinect technology for Xbox 360. http://www.sxbox.com/Kinect
40. Microsoft (2012) Microsoft Speech Application Interface. http://msdn.microsoft.com/en-us/library/ms723627(v=vs.85).aspx
41. Mitsubishi Heavy Industry Co (2011) Wakamaru. http://www.mhi.co.jp/en/products/detail/wakamaru.html
42. Miwa H, Itoh K, Matsumoto M, Zecca M, Takanobu H, Rocella S, Carrozza MC, Dario P, Takanishi A (2004) Effective emotional expressions with expression humanoid robot WE-4RII: integration of humanoid robot hand RCH-1. In: Intelligent robots and systems, Sendai, Japan, vol 3, pp 2203–2208
43. Mutlu B, Osman S, Forlizzi J, Hodgins J, Kiesler S (2006) Perceptions of ASIMO: an exploration on co-operation and competition with humans and humanoid robots. In: ACM/IEEE international conference on human-robot interaction, New York, USA, pp 351–352
44. Nakatsu R, Tosa N (2000) Active immersion: the goal of communications with interactive agents. In: Fourth international conference on knowledge-based intelligent engineering systems and allied technologies, vol 1, pp 85–89

45. Niewiadomski R, Bevacqua E, Mancini M, Pelachaud C (2009) Greta: an interactive expressive ECA system. In: 8th international conference on autonomous agents and multiagent systems, Richland, USA, vol 2, pp 1399–1400
46. OpenCV (2012) OpenCV Wiki. http://opencv.willowgarage.com
47. OpenNI (2012) Introducing OpenNI. http://openni.org
48. Osipa J (2010) Stop staring: facial modeling and animation done right. Wiley, New York
49. Philips, Holland (2011) Research robot iCat. http://www.research.philips.com/technologies/syst_softw/robotics/index.html
50. Ribeiro T, Leite I, Kędzierski J, Oleksy A, Paiva A (2011) Expressing emotions on robotic companions with limited facial expression capabilities. In: 11th international conference on intelligent virtual agents, Reykjavik, Iceland, pp 466–467
51. ROBOTIS CO, LTD (2012) User's manual RX-64. http://www.robotis.com
52. Schulte J, Rosenberg C, Thrun S (1999) Spontaneous, short-term interaction with mobile robots. In: IEEE international conference on robotics and automation, pp 658–663
53. SDL (2012) Simple DirectMedia Layer. http://www.libsdl.org
54. Shibata T, Ohkawa K, Tanie K (1996) Spontaneous behavior of robots for cooperation. Emotionally intelligent robot system. In: IEEE international conference on robotics and automation, vol 3, pp 2426–2431
55. Sony, Japan (2011) Toy AIBO. http://support.sony-europe.com/aibo/
56. Wachsmuth I, Kopp S (2002) Lifelike gesture synthesis and timing for conversational agents. In: Gesture and sign language in human-computer interaction. Lecture notes in computer science, vol 2298. Springer, Berlin, pp 120–133

Jan Kędzierski M.Sc., he is a Research Assistant at Wroclaw University of Technology (WRUT), Poland. From 2008 to 2012 he was a member of LIREC ("Living with Robots and Interactive Companions") EU FP7 project team. His research interests are focused on design, control, programming of and experimenting with social robots. He is the main designer of the robotic companion FLASH (with the robotic head EMYS). Now he is developing a uniform control system of WRUT social robots. Kedzierski is an author of 15 publications.

Robert Muszyński received M.Sc. and Ph.D. degrees (in control engineering & robotics) from Wrocław University of Technology, Poland, in 1992 and 1996, respectively. From 1994 to 1996 he was an invited researcher at Research Institute for Symbolic Computation, Johannes Kepler University, Linz, Austria, where he worked on a research project "Hybrid Evolutionary Programming Techniques" in the frame of "Real World Computing Partnership" sponsored and coordinated by Japanese Ministry of Science. From 2008 to 2012 he was a researcher in LIREC ("Living with Robots and Interactive Companions") EU FP7 project. He is currently an Assistant Professor at Institute of Computer Engineering, Control and Robotics Wrocław University of Technology. His research interests include robot manipulator control, modelling, kinematic singularities of robotic manipulators, social robotics, and symbolic computations.

Carsten Zoll the leader of the Group for Interdisciplinary Psychology (GRIP) at the University of Bamberg, Germany, holds a master's degree in psychology and computer science. Before, he worked at the Department for General Psychology at the University of Bamberg and the professorship for economic and social psychology at the Technical University of Chemnitz, Germany, where he received his Ph.D. on financial decision-making. His scientific work focuses on simulation of humanlike behaviour, modelling the human mind, empathy, evaluation, and behavioural finance.

Adam Oleksy received M.Sc. degree (in control engineering & robotics) from Wrocław University of Technology, Poland, in 2009. From 2008 to 2012 he was a member of LIREC ("Living with Robots and Interactive Companions") EU FP7 project team, where he was responsible for designing and developing software for social robots. He is currently a software developer in Nokia Siemens Networks.

Mirela Frontkiewicz received M.Sc. degree (in control engineering & robotics) from Wrocław University of Technology, Poland, in 2012. During the studies she gained an internship at LIREC ("Living with Robots and Interactive Companions") EU FP7 project. She is currently a PhD student at Department of Fundamental Cybernetics and Robotics of WRUT. Her research interests cover control algorithms for stationary and mobile manipulators, as well as social robotics issues. She is an author of 3 publications.