MADRID 2018
IROS

October, 1-5, 2018 Madrid, Spain

Exhibitors Job Openings

IROS

October, 1-5, 2018, Madrid, Spain

2018 IEEE/RSJ International Conference on Intelligent Robots and Systems

Towards a Robotic Society

▶ Play Video

Copyright

Copyright and Reprint Permission

Abstracting is permitted with credit to the source. Libraries are permitted to photocopy beyond the limit of U.S. copyright law for private use of patrons those articles in this volume that carry a code at the bottom of the first page, provided the per-copy fee indicated in the code is paid through Copyright Clearance Center, 222 Rosewood Drive, Danvers, MA 01923. For reprint or republication permission, email to IEEE Copyrights Manager at pubs-permissions@ieee.org.

All rights reserved. Copyright ©2018 by IEEE.
IEEE Catalog Number: CFP18IRO-USB
ISBN: 978-1-5386-8093-3

Author Responsibilities

The IEEE distributes its technical publications throughout the world and wants to ensure that the material submitted to its publications is properly available to the readership of those publications. Authors must ensure that their Work meets the requirements as stated in section 8.2.1 of the IEEE PSPB Operations Manual, including provisions covering originality, authorship, author responsibilities and author misconduct. More information on IEEE's publishing policies may be found at
http://www.ieee.org/publications_standards/publications/rights/pub_tools_policies.html
Authors are advised especially of IEEE PSPB Operations Manual section 8.2.1.B12: "It is the responsibility of the authors, not the IEEE, to determine whether disclosure of their material requires the prior consent of other parties and, if so, to obtain it." Authors are also advised of IEEE PSPB Operations Manual section 8.1.1B: "Statements and opinions given in work published by the IEEE are the expression of the authors."

Author/Employer Rights

If you are employed and prepared the Work on a subject within the scope of your employment, the copyright in the Work belongs to your employer as a work-for-hire. In that case, the IEEE assumes that when you sign this Form, you are authorized to do so by your employer and that your employer has consented to the transfer of copyright, to the representation and warranty of publication rights, and to all other terms and conditions of this Form. If such authorization and consent has not been given to you, an authorized representative of your employer should sign this Form as the Author.

Contact Information

For technical inquiries about the content of these proceedings, please contact:
PaperCept, Inc.
contact@papercept.net

Hamad, Mahdi	Mr
Weiss, Jakob	Tech. Univ. München
Eslami, Abouzar	Carl Zeiss Meditec AG
Huang, Kai	Sun Yat-Sen Univ
Maier, Mathias	Klinikum Rechts Der Isar Der TU München
Lohmann, Chris P.	Klinikum Rechts Der Isar Der TU München
Navab, Nassir	TU Munich
Knoll, Alois	Tech. Univ. Muenchen TUM
Nasseri, M. Ali	Tech. Univ. Muenchen

WeCTS6   Room 1.L3

Human-Robot Interaction VII   Regular session

Chair: Padir, Taskin   Northeastern Univ

Co-Chair: Magnusson, Martin   Örebro Univ

12:30-12:33, Paper WeCTS6.1

📄 🔗 I Can See Your Aim: Estimating User Attention from Gaze for Handheld Robot Collaboration

Stolzenwald, Schachar Janis Immanuel	Univ. of Bristol
Mayol, Walterio	Univ. of Bristol

12:33-12:36, Paper WeCTS6.2

📄 Recursive Bayesian Human Intent Recognition in Shared-Control Robotics

Jain, Siddarth	Northwestern Univ. Rehabilitation Inst. of Chicago
Argall, Brenna	Northwestern Univ

12:36-12:39, Paper WeCTS6.3

📄 🔗 A Novel Shared Position Control Method for Robot Navigation Via Low Throughput Human-Machine Interfaces

Sinyukov, Dmitry	Worcester Pol. Inst
Padir, Taskin	Northeastern Univ

12:39-12:42, Paper WeCTS6.4

📄 🔗 Robot Identification and Localization with Pointing Gestures

Gromov, Boris	IDSIA
Gambardella, Luca	USI-SUPSI
Giusti, Alessandro	IDSIA Lugano, SUPSI

12:42-12:45, Paper WeCTS6.5

📄 Establishing Appropriate Trust Via Critical States

Huang, Sandy H.	UC Berkeley
Bhatia, Kush	Univ. of California Berkeley
Abbeel, Pieter	UC Berkeley
Dragan, Anca	Univ. of California Berkeley

12:45-12:48, Paper WeCTS6.6

📄 🔗 Learned Hand Gesture Classification through Synthetically Generated Training Samples

Lindgren, Kyle	Univ. of Washington
Kalavakonda, Niveditha	Univ. of Washington
Caballero, David	Univ. of Washington

Huang, Kevin	Trinity Coll
Hannaford, Blake	Univ. of Washington

12:48-12:51, Paper WeCTS6.7

[PDF] Interaction System Based on an Avatar Projected on a Pyramidal Display

Loza Matovelle, David César	Univ. De La Fuerzas Armadas ESPE
Marcos, Samuel	CARTIF Foundation
Zalama, Eduardo	Inst. De Las Tecnologías Delaproducción(itap).univ. Va
Gomez Garcia Bermejo, Jaime	Univ. of Valladolid

WeCTS7	Room 2.L3
Autonomous Vehicles II	Regular session
Chair: Topp, Elin Anna	Lund Univ. - LTH
Co-Chair: Stachniss, Cyrill	Univ. of Bonn

12:30-12:33, Paper WeCTS7.1

[PDF] Multimotion Visual Odometry (MVO): Simultaneous Estimation of Camera and Third-Party Motions

Judd, Kevin Michael	Univ. of Oxford
Gammell, Jonathan	Univ. of Oxford
Newman, Paul	Oxford Univ

12:33-12:36, Paper WeCTS7.2

[PDF] Underwater Surveying Via Bearing Only Cooperative Localization

Damron, Hunter	Univ. of South Carolina
Quattrini Li, Alberto	Dartmouth Coll
Rekleitis, Ioannis	Univ. of South Carolina

12:36-12:39, Paper WeCTS7.3

[PDF] Ego-Motion Estimate Corruption Due to Violations of the Range Flow Constraint

Monaco, Chris	Pennsylvania State Univ
Brennan, Sean	Penn State Univ

12:39-12:42, Paper WeCTS7.4

[PDF] Semi-Supervised SLAM: Leveraging Low-Cost Sensors on Underground Autonomous Vehicles for Position Tracking

Jacobson, Adam	Queensland Univ. of Tech
Zeng, Fan	Queensland Univ. of Tech
Smith, David	Caterpillar
Boswell, Nigel	Caterpillar
Peynot, Thierry	Queensland Univ. of Tech. (QUT)
Milford, Michael J	Queensland Univ. of Tech

12:42-12:45, Paper WeCTS7.5

[PDF] An Automatic Tracked Robot Chain System for Gas Pipeline In-Spection and Maintenance Based on Wireless Relay Communica-Tion

Zhao, Wen	Waseda Univ
Kamezaki, Mitsuhiro	Waseda Univ
Yoshida, Kento	Waseda Univ
Konno, Minoru	Tokyo Gas Co. Ltd

Interaction system based on an avatar projected on a pyramidal display

David Loza Matovelle, Samuel Marcos, Eduardo Zalama, Jaime Gómez García Bermejo

Abstract— In this paper an interaction system based on a three dimensional virtual head projected onto a pyramidal display is proposed. The proposed system makes use of a social robot behavioral architecture already developed in our lab, which allows us to interchange developments between our robotic realizations and the 3D avatar. The overall system is divided into two parts: back projection subsystem and expression generator subsystem. The back projection subsystem projects a three-dimensional avatar onto a pyramidal structure in order to achieve a sensation of depth and realism. The expression generator subsystem carries out the avatar animations using shape keys and bones, following the Facial Action Coding System (FACS). The system consists in several nodes that are integrated in ROS middleware (Robotic Operating System), and includes a user interface that makes the avatar teleoperation easier (the package is avaible in github public respository). In order to evaluate the expressiveness of the system, two sets of experiments have been performed: one to analyze the avatar's gestural ability, that is, its capability to perform expressions that can be identified by an observer, and a second experiment to measure the emotion displaying ability in terms of valence and arousal.

I. INTRODUCTION

Human robotic interaction (HRI) is becoming more common. Due to recent developments in artificial intelligence and the increased level of interaction required in many difficult tasks. HRI has changed from being a mere requisite to a need to develop tasks in a better way. In order to do this, we must switch from complicated procedures to more natural forms of interaction and communication.

One of the ways to make communication more fluid is to improve the anthropomorphism of robots. For this reason, systems with high gestural abilities have been built and used in the HRI. However, there are significant limitations regarding the development of systems that can have high interaction and gestural abilities. Indeed, developing such systems can be excessively expensive and also have problems related to the appearance and naturalness. According to Zlotowski *et al* [12] the impact of anthropomorphism on HRI is not developed, for this reason is necessary working in a robot to maintain its *"human-like impression"*.

In the present paper we propose an intermediate solution which lies in between avatars and robots. As such, an animated avatar is projected onto a pyramidal display and

David Loza M.; DISA, University of Valladolid, Valladolid - España; Universidad de las Fuerzas Armadas ESPE, Sangolqui - Ecuador; e-mail: (dcloza@espe.edu.ec)

Samuel Marcos; CARTIF Foundation, Valladolid - España; e-mail: (sammar@cartif.es)

Eduardo Zalama; ITAP-DISA, University of Valladolid, Valladolid - España; e-mail: (ezalama@eii.uva.es)

Jaime Gómez García Bermejo; ITAP-DISA, University of Valladolid, Valladolid - España; e-mail: (jaigom@eii.uva.es)

endowed with a ROS based interaction system which is easy to replicate and can reach high interaction abilities. The goal is to develop a 3D environment that improves the interaction level and the presence sensation, adding a dimension to the common avatars at a relatively low cost. The present work is aimed at improving interaction through the development of virtual avatars based on a robotic architecture.

II. HUMAN EXPRESSION AND INTERACTION SYSTEMS

One of the main communication form of human beings is gestures; human face produces the most obvious and immediate signals to convey a message. There are multiple studies about human expression, such as EMFACS [3](Ekman and Friesen, 1982), Monadic Phases [11](Tronick, Als, and Brazelton, 1980) and the Maximally Discriminative Facial Movement Coding System [4] (MAX: Izard, 1979). In the present work we used Eckman and Friesen approach based on Facial Action Coding System (FACS) and action units [2]. Action Units (AUs) are the fundamental actions of individual muscles or groups of muscles which produces small discriminable changes in the face. FACS describes about one hundred units taking into account movements of the eyes, eyebrows, eyelids, neck, among others, that combined produce basic expressions.

On the other hand, transferring a system like FACS to a robot or avatar with gestural capacities is a difficult task. Many times we have seen robots and avatars with a group of coordinated animations that do not take into account other elements such as flexibility for an adequate interaction. Many examples of systems with high gestural capacities but lacking the necessary realism and adequate interaction have been developed, many of which include elaborate routines, making them predictable and unappealing.

Hence, to generate systems with high gestural capacities, it is important to take into account the analysis of expressions for a better understanding of interaction. Part of the present study analyses the implementation of FACS in the model, as well as its execution to obtain a high degree of anthropomorphism.

III. BACK PROJECTION SYSTEM

As starting point we have worked with a 3D model obtained from side and frontal pictures from a young girl, the model of this virtual head was developed in [8]. This model has been chosen because it contains an adequate appearance together with a reduced number of polygonal surfaces with respect to other head models such as [5] which facilitates manipulation. This model has been obtained by the deformation of a generic model (photographs and
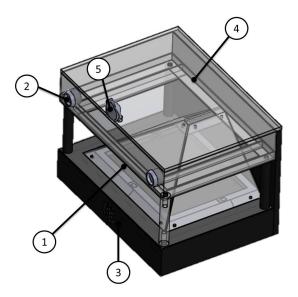
Fig. 1. CAD model of the system 1. Pyramid 2. Speakers 3. Microphones 4. Display 5. Video camera

a base model). As a modeling tool, Blender project has been selected because it provides an open and multi-platform software, has a large tool range and an active development community, among others.

The back projection system is based on the three views of the avatar on the pyramid. Fig. 1 shows the CAD model of the pyramid and its different elements. General dimensions are $60 \times 50 \times 45$ cm. The pyramid (Fig. 1-1) is truncated in the back side to increase the projection area of the avatar. Three views of the avatar are projected on the pyramid from a flat screen (Size: 17 inches, and 1024×768 pixel resolution) disposed on the upper side. Other elements, such as microphones, speakers, and web-cam are part of the equipment (Fig. 1-2, 1-3, 1-4) and are used as communication elements.

Three animated copies of the same avatar which move simultaneously have been created generating three views. Each view is projected over one face of the pyramid (front and two side views). Rear face is not projected since it is not accessible to the user because there is a dark panel to increases the contrast and to limit the incidence of external light. Side projections are linked to the actions of the front projection. Fig. 2 shows a screenshot of the avatar projected over the pyramid.

IV. GESTURE GENERATOR SYSTEM

The gesture-generating system incorporates the FACS into the Avatar to give it gesture and movement ability.

A. Application of the FACS for the animation.

FACS is a widely used system to systematically describe human facial expressions. FACS analyses the classification of individual face movements, based on the activation of different muscles. The different movements are compiled in action units. Due to the great quantity and variety of units described by FACS, only the 11 most significant action

Fig. 2. Pyramidal display

TABLE I
SELECTED ACTION UNITS

Action Unit	Name
AU1	Inner Brow Raiser
AU2	Outer Brow Raiser
AU4	Brow Lowerer
AU5	Upper Lid Raiser
AU6	Cheek Raiser
AU7	Lid Tightener
AU9	Nose Wrinkler
AU10	Upper Lip Raiser
AU12	Lip Corner Puller
AU15	Lip Corner Depressor
AU26	Jaw Drop

units have been taken for the development of expressions in the current work (Table I). Ten additional units have been added to complement the previous ones, to implement the movement of face and eyes (Table II) and provide further naturalness.

According to FACS, there are six basic expressions (sadness, fear, anger, happiness, surprise and disgust) and they can be understood as the combination of action units. Table III summarizes the combination of action units used for each basic expression. Fig. 3 shows the resulting avatar displaying the six basic expressions.

B. Face animation through shape keys and bones

Shape keys are particular deformations in the model mesh, interpolated based on the base expression. These shape keys

TABLE II
COMPLEMENTARY ACTION UNITS

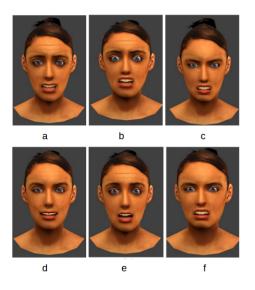| Unidad de acción	Nombre original
AU45	Blink
AU51	Head Turn Left
AU52	Head Turn Right
AU53	Head Up
AU54	Head Down
AU61	Eyes Turn Left
AU62	Eyes Turn Right
AU63	Eyes Up
AU64	Eyes Down
AU0	Posición neutra

Fig. 3. Basic expressions a. Sadness, b. Fear, c. Anger, d. Happiness, e. Surprise, f. Disgust

TABLE III

ACTION UNITS IN BASIC EXPRESSIONS

Emotion	Action units in the FACS	Action units implemented
Happiness	6+12	6+12
Sadness	1+4+15	1+4+15
Surprise	1+2+5+26	1+2+5+26
Fear	1+2+4+5+7+20+26	1+2+4+5+7+26
Disgust	9+15+16	9+15
Anger	4+5+7+23	4+5+7

are easy to record and can be edited, scaled and mixed with one another, in different percentages. Other form of animation is though armatures. Armatures are based on bones added by the user, which generate the different degrees of freedom of the model. An armature is the equivalent to a skeleton and is used to create animations quickly, but without the details provided by shape keys.

In this work an accurate combination of bones and shape keys have been used and stored in action units. The obtained model contains actions for each selected action unit, as shown in Tables I and II.

C. ROS – Avatar Connection

The proposed animation arquitecture (Fig. 4) has been implemented using ROS (a framework for software development for robots with many applications). The idea of the work is to use the capacities of ROS and connect it with Blender to provide benefits to the system and ease the future integration of enhanced interaction nodes. Three independent modules or ROS nodes have been implemented: ROS-Blender Bridge, Gesture Manager and Behaviours.

To connect ROS and Blender, a server in charge of compiling and executing simultaneously all action units sent from ROS has been programmed. The actions units were processed and transformed in animations. In brief, the node receives a large group of action units resulting in an avatar animation based on shape keys and bones.
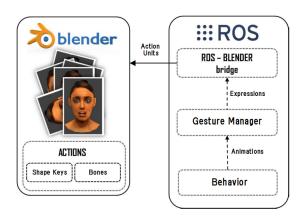

Fig. 4. Basic structure of animation in ROS

TABLE IV

DESCRIPTION OF ACTIVATION ELEMENTS IN AN ACTION UNIT

Element	Description
Name	Action unit identity in Blender.
Intensity	Upper and lower limits (start – end), which determine the animation intervals that will be executed.
Blendin	Percentage that can be combined with other action units, speed of each AU.
Velocity	Speed with which the action unit is executed, which then becomes the execution time of the action unit.

The next node in the animation system is the ROS-Blender Bridge, which ensures the connection with the avatar. This node is in charge of sending action units to Blender with the expression as a vector defined in formula 1. Table IV describes components of an action: name, intensity, velocity and blendin.

$$action = [Name, Intensity, Blendin, Velocity] \quad (1)$$

The gesture manager is responsible for modeling emotional expressions from the combination of different action units as shown in expression 2. Each action unit is associated with an individual intensity (AU_n , i_n). In addition, each expression contains a parameter of total intensity (i_t) that regulates the intensity of the expression as a whole (for example we can tune a 60% happiness or a 30% sadness). Finally, The total time (t_t), indicates how long will take to perform the expression. Those expressions can be defined as a combination of action units that can be created, recorded, and used expressions that can be called through labels (sadness, fear, sleepiness, and so on).

$$expresion = \begin{bmatrix} i_t \begin{bmatrix} AU_1 & i_1 \\ AU_2 & i_2 \\ ... & ... \\ AU_n & i_n \end{bmatrix}, t_t \end{bmatrix} \quad (2)$$
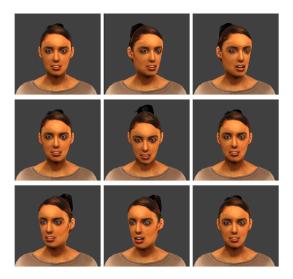
Fig. 5.   Animation sequence with key frames

	Happ	Sad	Surp	Fear	Dis	Ang	Percent
Happiness	21	2	2	0	0	0	84%
Sadness	0	11	0	8	6	0	44%
Surprise	0	1	18	6	0	0	72%
Fear	0	0	8	15	2	0	60%
Disgust	0	0	0	1	14	10	56%
Anger	0	0	0	1	10	14	56%

	Happ	Sad	Surp	Fear	Dis	Ang	Percent
Happiness	22	0	0	1	2	0	88%
Sadness	0	15	7	2	1	0	60%
Surprise	0	0	20	4	1	0	80%
Fear	0	0	8	12	5	0	48%
Disgust	0	0	0	3	14	8	56%
Anger	0	0	0	0	4	21	84%

The behavior node handles a group of key frames defined in time and intensity. The idea of a key frame is similar to the ones used in animation programs. Key frames are constitutive elements in a sequence and are structured by a basic expression (defined as 2), to which extra action units are added. For instance, eye closure and head tilt can be added to the happiness expression by adding the corresponding action units.

Besides the basic expression plus the extra action units, the total intensity, expression time, and total time are added to define a key frame. These parameter correspond to the final expression intensity, the expression duration and the waiting time to exchange key frames, respectively.

The animation process consists in determining different key frames that are defined and interpolated. An example of a sequence carried out by the animation process is presented in Fig. 5. Such sequences can be saved and the operator can recall them through the user interface.

V. EXPERIMENTATION AND RESULTS

In order to evaluate the expressiveness of the system, two sets of experiments have been performed: one to analyze the avatar's gestural ability, that is, its capability to perform expressions that can be identified by an observer, and a second experiment to measure the emotion displaying ability in terms of valence and arousal.

The system has been placed on a supporting column at 1 m height Users have been suggested to place (initially) about 0.4 m away from the front face of the pyramid, seeing it at a viewing angle about 60 degrees. Moreover, users have been allowed to move freely around the pyramid, although most of them have remained about the initial distance. Room lighting has been dimmed to allow a proper visual contrast.

The experiments have been conducted as follows. In order to evaluate the emotional behavior of the developed avatar, we have performed two sets of tests that are reported in Table V. The first set (denoted as T1, T2 and T3) is oriented to measure the avatar's gestural ability, that is, its capability to perform expressions that can be identified by an observer. On the other hand, the second set (T4-1 and T4-2) aims to measure the emotional displaying ability of the head, which refers to the ability to perform emotional expressions which valence and arousal can be correctly perceived by an observer. The performed tests are based on other evaluations [9], [10] oriented to compare the expressiveness of a robotic development with the expressiveness of a reference set of emotions (e.g. a set of static images) or with other robots endowed with similar interaction capabilities. In addition, for the categorization of the valence and arousal of the performed expressions the "circumplex model of affect", proposed by James A.Russell [7] has been used in our experiment.

A. Tests of gestural ability

The first set of experiments aims to measure the differences, if any, in terms of gestural ability of the avatar. To achieve this, 75 participants took part in tests T1, T2 and T3. We divided participants into heterogeneous subgroups of 25 persons each, Each group was asked to identify the emotional expression showed by the avatar under three conditions: static avatar displayed on a computer screen (T1), animated avatar displayed on a computer screen (T2) and animated avatar displayed on the pyramidal display (T3). The six basic emotional expressions shown in Fig. 3 (sadness, fear, anger, happiness, surprise, and disgust) where randomly presented to the participants who were asked to identify them.

In the case of T1, the images shown to the participants correspond to a capture of the avatar displaying each emotional expression at maximum intensity. On the other hand, for T2 and T3, the avatar performed dynamically the six emotional expressions from a neutral face to the maximum expression intensity.

Tables V, VI and VII show the results for the T1, T2 and

TABLE VII

Evaluation T3 - Identifying expressions in the pyramid (animated)

	Happ	Sad	Surp	Fear	Dis	Ang	Percent
Happiness	22	0	0	3	0	0	88%
Sadness	1	17	2	4	1	0	68%
Surprise	1	0	22	2	0	0	88%
Fear	0	0	7	11	4	3	44%
Disgust	1	0	1	1	19	3	76%
Anger	0	2	0	0	5	18	72%

TABLE VIII

Evaluation T4-1 - Valence: Number of users who rated each expression in each category (from 1 to 9)

	1	2	3	4	5	6	7	8	9
Happiness	1	2	6	4	6	1	5	0	0
Sadness	5	3	8	5	3	1	0	0	0
Fear	3	2	3	8	6	2	1	0	0
Anger	1	5	8	7	2	2	0	0	0
Surprise	1	1	3	6	6	0	6	1	1
Disgust	2	4	15	4	0	0	0	0	0

TABLE IX

Evaluation T4-2 - Arousal: Number of users who rated each expression in each category (from 1 to 9)

	1	2	3	4	5	6	7	8	9
Happiness	4	3	4	4	3	6	1	0	0
Sadness	2	0	3	2	5	3	4	3	3
Fear	0	0	1	4	2	2	6	3	7
Anger	0	2	6	1	4	4	3	4	1
Surprise	0	0	3	0	3	1	5	7	6
Disgust	0	0	0	0	3	2	5	7	8

T3 respectively. It can be observed that dynamic expressions are more easily recognized than those displayed statically, which is aligned with previous research [1]. Also, we can see that the overall identification of expressions improves (around 20%) when displaying them on the pyramidal display (Table VII).

For all conditions, the expressions with higher recognition scores are happiness and surprise. The least clearly recognized expression is fear, whereas sadness and disgust are about a medium recognition level, in coherence with previous studies where the most difficult expressions to identify were anger and fear [6].

B. Tests of emotional expression ability

To evaluate the emotional expression ability of the prototype, the circumplex model of affect, proposed by James A. Russell [7] has been used. This model assesses emotions as the association of one or more dimensions in different levels. It is worth mentioning that there are not precise values for one or another expression in this study; rather, expressions are categorized in zones where they should be shown according to the intensity that we have.

For the present case, two dimensions of interest have been taken: valence and arousal. Valence is understood as the attractiveness (positive valence) or aversion (negative valence) of an event, object or situation (the expressions shown on the pyramid, in our case). On the other hand, the arousal can be understood as the value necessary for activating the different expressions.

The circumplex model of affect is shown in Fig. 6-a. Arousal represents the vertical axis and valence represents the horizontal axis, while the center of the circle represents a neutral valence and a medium level of arousal. The emotional states or expressions can be represented at any level of valence and arousal, or at a neutral level of one or both of these factors. Expressions such as alarmed, excited, and happy have high valence activation levels. On the other hand, fatigued and sleepy have low activation levels. In addition, expressions with positive valence values (pleasant) are found on the right side of the circumplex, consent and calm among them. In contrast, on the other side (unpleasant) expressions such as miserable, frustrated, distressed, etc. can be found. According to the model, the emotions decrease in correlation as one moves away from another in the circumplex.

This context was explained to the users, and they were shown the six expressions and asked to describe, in terms of valence and arousal, the sensation "conveyed" by the avatar. A moderator registered the level of arousal and valence (positive or negative). The 25 users were asked to score from 1 to 9 the six emotional expressions in terms of valence and arousal. Previously they were informed about the meaning of the two variables. The results are shown in table VIII (valence) and table IX (arousal). The valence was evaluated from 1 to 9, 0 corresponding to the most unpleasant value and 9 to the most pleasant one; and the arousal (different to that of the FACS's) was evaluated in the same range, 1 for a moderate "effort" and 9 for the maximum one. Tables show the number of subjects reporting each possible value.

Averaged values of valence and arousal for the different expressions from tables VIII and IX are represented in Fig. 6-b. Figure shows the areas in which the different expressions produced by the pyramid are present. The majority of expressions are shown in the second quadrant, which is appropriate (4 out of 6 expressions fall in this region: sadness, fear, disgust, anger) in comparison with the model of Fig. 6-a. Also, the activation level is appropriate: the majority of expressions require positive activation levels. Also there is a high consensus on arousal and valence. However, happiness is correctly located in valence, but don't have appropriate intensity to be more evident. Particularly the case of the smile depends on context (sometimes is called the "fake smile"), in the present case being an evaluation without any context, can give the user the wrong feeling. In general the emotional expressions evaluated agree with the theoretical results in terms of valence and intensity, which means that the expressions carried out inside the pyramid are appropriate and allow for the context to be conveyed for each expression.

VI. Conclusion and future work.

In this paper, an interaction system based on a pyramidal display has been presented. The different elements have been described, covering the modeling, structure and construction
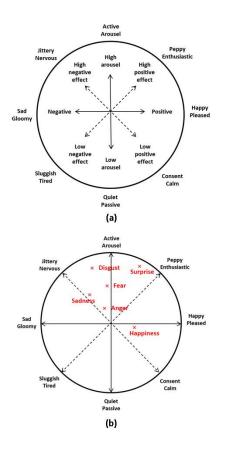
Fig. 6. Circumplex model of affect (a) Theoretical [7] (b) Obtained from experimentation

REFERENCES

[1] Zara Ambadar, Jonathan W. Schooler, and Jeffrey F. Cohn. "Deciphering the Enigmatic Face". In: *Psychological Science* 16.5 (2005), pp. 403–410.

[2] P. Ekman & W. Friesen. "Facial Action Coding System: A Technique for the Measurement of Facial Movement". In: *Consulting Psychologists Press* (1978).

[3] P. Friesen W. & Ekman. "EMFACS-7: Emotional Facial Action Coding System." University of California, California. 1983.

[4] Carroll Ellis Izard. *The Maximally Discriminative Facial Movement Coding System*. Ed. by University of Delaware. University of Delaware, 1983.

[5] Samuel Marcos, Jaime Gómez-García-Bermejo, and Eduardo Zalama. "A realistic, virtual head for human–computer interaction". In: *Interacting with Computers* 22.3 (2009).

[6] Samuel Marcos-Pablos et al. "Virtual Avatar for Emotion Recognition in Patients with Schizophrenia: A Pilot Study". In: *Frontiers in Human Neuroscience* 10 (2016), p. 421.

[7] J.A. Russell. "A circumplex model of affect". In: *Journal of personality and social psychology* 39.6 (1980), pp. 1161–1178.

[8] Cristina Saldaña Burgos Irene & Del Rio Blanco. "Modelado y animación facial de un avatar virtual realista mediante formas clave". MA thesis. Universidad de Valladolid - UVa, 2013.

[9] Jelle Saldien et al. "Expressing Emotions with the Social Robot Probo". In: *International Journal of Social Robotics* 2.4 (2010), pp. 377–389.

[10] Stefan Sosnowski et al. "Design and Evaluation of Emotion-Display EDDIE." In: *IROS*. IEEE, 2006, pp. 3113–3118. ISBN: 1-4244-0259-X.

[11] Tronick E & Als H & Brazelton TB. "A structural descriptive analysis of infant-mother face to face interaction." In: *Merrill-Palmer Quarterly of Behavior and Development* 26(1) (1980), pp. 3–24.

[12] Jakub Zlotowski et al. "Anthropomorphism: Opportunities and Challenges in Human–Robot Interaction". In: *International Journal of Social Robotics* 7.3 (2015), pp. 347–360.

of the pyramid. The interaction system has been discussed, as well as its node-based architecture and the way the avatar is controlled though ROS. The different expressions have been generated using the FACS, and to this end, the number of action units has been refined to perform the six basic emotions. The animations are built by using shape keys and bones, grouped in actions. The system have been developed following a methodology with a clear robotic orientation based on the use of (virtual) bones-type actuators and the FACS human expression model, as opposed to the usual approaches used in animation. Therefore, the current approach allows the results to be directly applied to the development of physical robotic heads.

The experiments carried out in this work have covered two main aspects. On the one hand, the recognition of emotions by the users has been analyzed. It has been shown that the emotional expressions generated by the 3D avatar are easier to recognize and preferred by the users over static and animated images. On the other hand, the emotional ability has been analyzed using the circumplex model. Major finding has been that the pyramidal display helps perceiving the system expressions and, therefore, provides a better interaction with the user.

We are currently working on integrating further interaction abilities (face recognition, lip synchronization, speech, visual