A New Paradigm for Learning Affective Behaviors: Emotional Affordances in Human Robot Interaction

F. Cid, A. J. Palomino and P. Núñez

***Abstract*—In the last decade, affordances have been successfully used in robotics for learning by imitation. From robotics perspective, affordances have been powerfull since represent relationships between objects, effects and opportunities of observer's actions. In Human Robot Interaction, emotions play an important role when the communication participants try to show their intentions. The proposed paper introduces the concept of affective affordances as an extension of the classical perceptual affordances, where now represent the relation between affective elements, such as the objects and the emotional states, effects and oportunities for the robot's reactions. Affective affordances are also related to intentions and are presented here as part of an more complex affective behaviors learning. The major purpose of this paper is to communicate the proposed conceptual idea to the robotics community working in affective Human-Robot Interaction.**

***Index Terms*—Affordances, Human-Robot-Interaction, Learning, Social Robots.**

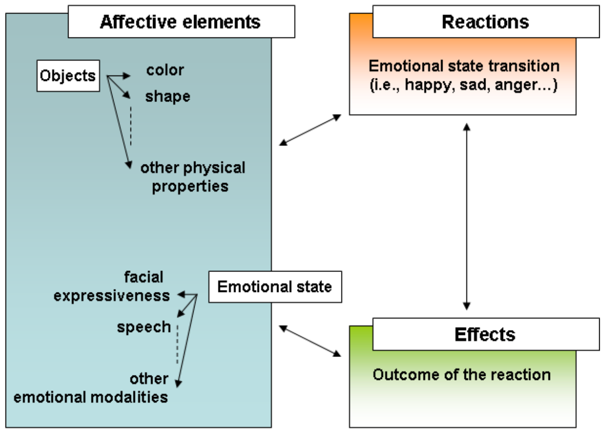

Fig. 1. Emotional affordances: a relationship between stimulus (*affective elements*), observer's reactions and effects.

I. Introduction

Probably, in a near future robots will help humans in their daily life. In this context, in that robots work in cooperation with people in real scenarios like offices or homes, it is essential that these robots have social skills. Social robots follow behaviors similar to humans: they interact and communicate with humans by following a set of social rules, *e.g.,* by using means of communication also used in human-human interaction (such as speech, facial expressiveness or body language). How these robots autonomously learn these social behaviors is an interesting working area and source of different research lines in the last decades. And how this social behaviors configure an affective Human Robot Interaction (HRI) is the final goal of the ongoing work presented in this paper.

Some classical psychological theories, such as Piaget's theory [1], describe different stages in the development of human behaviors. In the beginning, Piaget suggests that newborns have no cognitive structures. Instead they have reex structures for sucking, grasping, and crying (*e.g.,* to close their hands when their palms are touched). Social behaviors appear during the second or third month of life and most of them are learned by imitation.

Related to this theory, the ecological psicologist Gibson defines the concept of *Affordances* [2] as "the affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill", that is action opportunities available at an environment or at an object to an observer, thus depending on its action skills. This ecological theory has been used in learning research in the last decades. In robotics, for instance, affordances are powerfull since: a) robots must be able to deal with unknown objects; and b) humans need ways of instructing robots. The classical use of affordances for action learning by imitation was based on interacting with physical objects characterized by their color, size or shape and analyzing the final effect [3].

In this paper, an extension to the concept of these perceptual affordances is proposed for learning affective behaviors. In a similar way that affordances define connection between objects and opportunities for actions, the emotional affordances in robotics extend this concept by including the relationship between affective elements (*stimulus*), effects and possibilites for observer's *reactions* (see Fig. 1). In the context of affective HRI, affective elements are related to not only physical characteristics of the objects, but also the environment or the own emotional state of the communication participants. Similar to perceptual affordances, emotional affordances can be used to predict outcome of an emotional reaction, to plan a reaction to achieve a specific goal or to select affective elements to produce an effect.

The approach presented in this work is part of a more complex cognitive robotic architecture for affective HRI, which is also outlined in this paper. Although there is not yet experimental results, the conceptual result of ongoing work, in conjunction with a detailed description of the basic skills of the robot are addressed. In order to proof this new paradigm for learning affective behaviors, the existing formalization for modelling affordances is used [4].

This paper is organized as follows: after discussing previous

F. Cid and P. Núñez are members of Robotics and Artificial Vision Lab. *Robolab* Group, University of Extremadura, Spain. A. Palomino is with *ISIS* Group, University of Málaga, Spain.
E-mail: fecidb@alumnos.unex.es

works in the literature related to cognitive architectures and perceptual affordances for learning by imitation, in Section II, Section III presents the concept of emotional affordances. An overview of the proposed learning algorithm for affective behaviors is presented in Section IV, where the affective HRI scenario used for the learning is also described. Finally, Section V summarizes the conclusions and future works of the approach.

II. Related Works

As defined by Gibson three decades ago [2], affordances are agent-dependent object usages. Since the formulation of this idea, many robotics research groups have done a great effort in investigating what is the best model for affordances and how affordances can be learned by robots. Modelling these perceptual affordances by learning has been studied in different works. In [5], the authors propose to learn affordances using a self-organizing maps (SOM). Other approaches were proposed in [6], where authors describe a general model based on Bayesian networks coupling actions, effects and objects features. This formulation is also used in other interesting works, such as [7] or [8], where the authors extend affordances by taking the environmental context. Also in Morie *et al.*'s work [9] the perceptual affordances were extended including affective elements. However, these emotional affordances were used in the creation of more emotionally affective Virtual Enviroments.

Development of Cognitive architectures has been studied for many researchers in the last decades. These general cognitive systems should consider both, sensory motor and learning aspects. An interesting review of cognitive architectures is found in [10]. Particularly, Scheutz *et al.*[12] presented an architecture for complex affective HRI, where the *affect* plays an important role in the integration of the subsystems. More recently, the two major aspects that should be simultaneously present in affective cognitive architectures was presented in [11]: *intentions* and *emotions*. This recommendation has been used for the development of the emotional cognitive architecture for affective HRI addressed in this paper, where affective affordances are the basis of the architecture.

III. Emotional affordances Modelling

The proposed algorithm for affective behaviour learning is built on Gibsons affordances, by expanding this concept to include the relationship between affective elements, effects and possibilites for observer's reactions. For instance, if someone gives to a person a bunch of flowers and does it with a kind smile, this person not only knows that can grasp it (perceptual affordances), but also, his logical reaction is a possitive emotional state. In this paper, affordances can be viewed as a continuum to illustrate the complementary and overlapping domains of perception and emotion [9] (see Fig. 2).

In order to model the affective affordances, the formalization from [3] has been adopted which is based on using a Bayesian Network. First, the robot is assumed to be endowed with a set of basic skills for perceiving both, the world

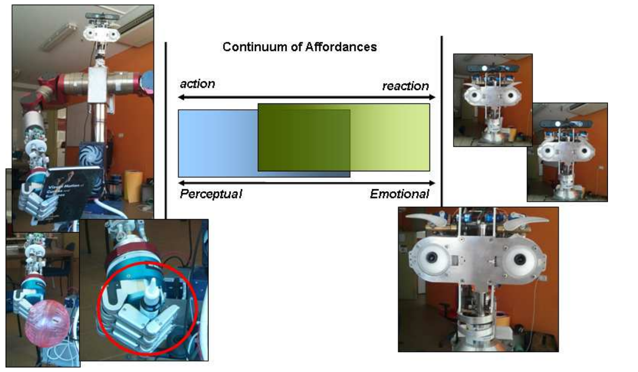

Fig. 2. Continuum of affordances from perceptual to emotional. On the left, different objects (size, color and shape) infer different plans and actions (perceptual affordances). On the right, these same objects and the emotional state of the communication participants also infer a reaction in the robot.

surrounding and the human emotional state. Details on the implementation of these basic skills are given in Section IV.

Let the discrete random variable $R = \{r_i\}$ represent the execution of a robot reaction. Affective elements (objects properties and the human emotional state) and effects during the HRI are also modeled using discrete random variables. Let F_{aff} be the human emotional state estimate according to a multimodal emotion recognition system (*e.g.,* facial expressiveness, speech or body gestures). Besides, let $F_o = \{F_o(1), ..., F_o(n_o)\}$ be object features extracted for object (*e.g.,* color, size, simetry, or shape). Finally, let $E = \{E(1), ..., E(n_e)\}$ represent the effects detected by the robot in the human emotional state after executing a reaction. The set of nodes G is formed by the discrete variables A, F_{aff}, F_o and E, $Z = (A, F_e, F_{aff}, F_o, E)$.

The graphical model of this Bayesian Network $B = (G, \Delta)$ over a set of variables Z = $\{Z_1, ..., Z_n\}$ is illustrated in Fig. 3. In this representation, the nodes of the graph G represent the random variables Z and the arcs are associated to conditional independence assumptions. In an affective HRI, the effect in the human emotional state, which is evaluated by the robot assuming a multimodal emotion recognition system, is dependent of the properties of the affective elements and also of the robot's reaction. Let $\Delta = \{\delta_i\}$ define the conditional probability distribution $p(Z_i|Z_{P_a(Xi)}, \delta_i)$ of each node in G. Thus, similar to [3], affective affordances are represented by the arrows between nodes and the parameters (casuality property). As this Bayesian Networks encodes the relationship between reactions, affective elements and effects, it is easy to compute the distribution of a variable or a group of variables given the values of the others (*ι.e.* to plan a robot's reaction given an intention, or to estimate an effect after a reaction).

IV. Learning affective behavior by imitation: a new perspetive

In order to develop cognitive architectures to be capable of generating affective behaviors similar to humans, it is important to consider the role of these emotional affordance. How these emotional affordances can be learned by robots is the essential of this work. The proposal is based on the work of [4], and assumes that the robot has learned a suitable set

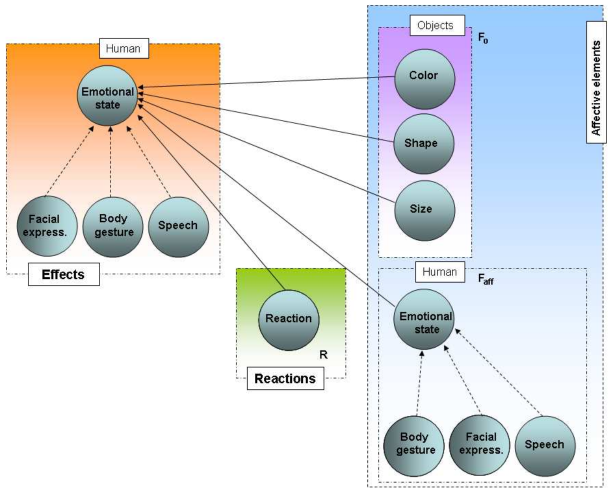

Fig. 3. Bayesian network model to represent the affective affordances. See the text for more details.

TABLE I
LEARNING STAGES OF THE ONGOING APPROACH

Basic skills	1: learn basic affective skills 2: Develop an active visual perception system for extracting phisical properties of objects
Affective HRI scenario	3: Perception of effects and categorization 4: Learn affective affordances 5: Prediction and planning affective skills
Imitation	6: Perform imitation tasks

of elementary actions and reactions to explore the world and estimate the human emotional state. In Table I, a description of the learning stages is illustrated.

Next, following subsections explain with details three main parts of the learning algorithm proposed in this paper.

A. Basic skills for an affective interaction

1) The Attention Model: In this subsection, an object-based model of visual attention for a social robot which works in a dynamic scenario [13] is proposed as a basic perception skill. In the last years, computer vision researchers have been trying to take advantage of biological visual systems, which are able to filter out the irrelevant information in the scene to focus all its resources in processing only relevant parts. The psychological basis to develop artificial visual attention systems are mainly two complementary theories: Treisman's *Feature Integration Theory* [14] and Wolfe's *Guided Search* [15]. The first one suggests that the human vision system detects separable features in parallel in an early step of the attention process (the *pre-attentive* stage, which is totally task-independent) to finally integrate them thorough a bottom-up process into a single saliency map. Several years later, Wolfe proposed that a top-down component in attention can increase the speed of the process giving more relevance to those parts of the image corresponding to the current task. Furthermore, attention theories introduce another important concept: the *Inhibition of Return*. This mechanism implies that an already attended object should not be selected again until some time later. Otherwise, the most relevant object would be always selected.

The used attention system integrates task-independent bottom-up processing and task-dependent top-down selection. In this model, the units of attention are the so-called *proto-objects* [16], that are defined as units of visual information that can be bound into a coherent and stable object. On one hand, the bottom-up component determines the set of proto-objects present in the image, describing them by a set of low-level features that are considered relevant to determine their corresponding saliency values. On the other hand, the top-down component weights the low-level features that characterize each proto-object to obtain a single saliency value depending on the task to perform.

An overview of the system is shown in fig. 4. In the *pre-attentive* stage, the different proto-objects present in the image are extracted, using a perceptual segmentation algorithm based on a hierarchical framework [17]. Then, the relevance of each proto-object is computed taking into account different low-level features. The features involved in saliency computation are: colour contrast, intensity contrast, proximity, orientation contrast, roundness, symmetry, similarity with red colour, similarity with blue colour, similarity with green colour, similarity with yellow colour and similarity with skin colour. These features are weighted by a set of *perception parameters* (λ_i) stored in a *Perception-Modulation Memory* (PMM). Depending on the value of these parameters, the system is able to modify the influence of each low-level feature in the global saliency computation. As a result of this stage, a set of proto-objects ordered by their saliency is obtained.

The next stage, the *semi-attentive* stage, deals with the management of the *Working Memory* (WM) and the *Inhibition of Return* (IOR). The WM establishes the maximum number of attended elements that can be maintained at once. It is a short-term memory where the system stores the recently attended objects and it has a reduced capacity, up to 5 elements [18]. Each proto-object in WM is characterized by a set of descriptors: its saliency value, its position in the image, the different low-level features values and a *time-to-live* value which establishes the maximum time that the proto-object can stay in WM. A proto-object's saliency also depends on this last parameter, so the longer an element is kept in WM, the lower its saliency will be. A new proto-object get into the WM if and only if it has bigger saliency than the currently stored elements. If the memory is full, the less salient element is dropped out. Regarding the IOR, a tracker module keeps permanently updated the position of each element in WM. Thereby, it is avoided to attend an already selected proto-object.

Both WM and PMM are the interface between early attention stages and the rest of the system, including the deliberative level. This interface includes a categorizer which is able to classify the perceived proto-objects into categories corresponding to high-level predicates. Besides, the PMM has been modified to translate high-level instructions into a new set of perception parameters λ_i, so it is allowed to change the way the vision system perceives the world in terms of a high-level decision. Fig. 5 illustrates the image processing of the attentional perception system.

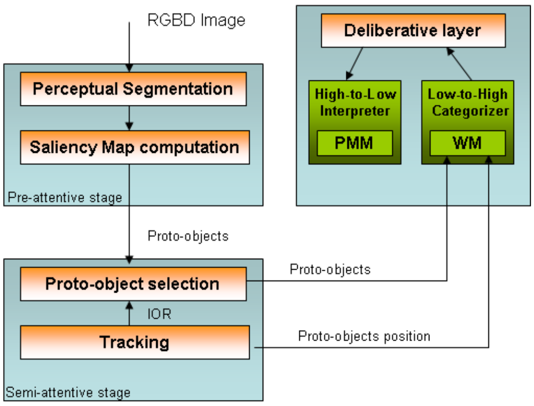

Fig. 4. Overview of the Object-Based Attention Model.

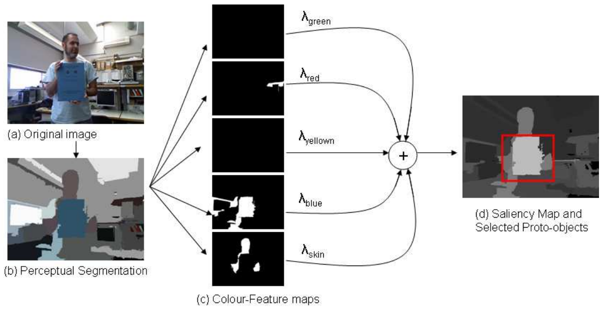

Fig. 5. The attention model described in this work as basic skill.

2) Multimodal Emotion recognition system: The multimodal emotion recognition system is presented in this subsection. The framework consists of three subsystems running in parallel that estimate the human emotional state using three independent Dynamic Bayesian Networks (DBN). Each subsystem deals with an information channel or *mode* associated to the facial expressiveness, the speech and the body language. First, the facial expression recognition subsystem is based on a modified version of the authors' previous work [20]. In this proposal, the robot acquires the information using a RGB-D camera and extracts the Action Units[1] from a Candide-3 reconstruction model. The same RGB-D sensor is used to estimate a human emotional state, extracting invariant features from the body analysis. In the third subsystem, users speech in the conversation is analyzed in order to extract a set of independent descriptors too.

In Fig. 6, an overview of the Multimodal emotion recognition system is illustrated. Four different emotions as well as non-emotional state are detected using this Bayesian approach (happiness, sadness, anger, fear and neutral). Finally, the proposed system integrates the information associated with all methods in a fourth DBN, which estimates the final user emotion (fusion engine). Each subsystem is briefly described below.

- *Emotion recognition system from facial expressiveness*: The proposed methodology consists of a robust feature extraction algorithm, which uses the Candide-3 reconstruction model described in [21]. These facial features extracted from the face mask are a combination of independent and antagonistic distortions of the face and constitute the input of a Dynamic Bayesian Network (DBN) used as classifier [20]. The facial expressiveness descriptor $d_f = \{EB, Ch, LE, LC, CB, MF, MA\}$ is a vector containing a set of normalized distances calculated using Actions Units values as is described in Table II. In Fig. 7b, the candide-3 face model is drawn over the human face. Extracted features are shown in Fig. 7c. Finally, three consecutive time intervals of the DBN are shown in Fig. 7d.
- *Emotion recognition system from body gestures*: The Body language plays an important role in affective state recognition. Body gestures are analyzed according to a set of invariant features, which are extracted from the human skeleton. Table III illustrates the methods used for extracting the body gestures features vector $d_g = \{QoM^{lh}, QoM^{rh}, Vel, Acc, c_i\}$, where QoM^{lh} QoM^{rh} represent the *Quantity of Motion* for the left and right hands, respectively. Vel, Acc and c_i describe the velocity, acceleration and the *contraction index*[2] of the human body. Finally, d_g is used as the input vector in a DBN classifier. Fig. 8a shows the 3D positions of hands and head during a human motion (N frames) in an emotional state of happiness (Fig. 8b), x^{lh}, x^{rh} and x^h, respectively. These 3D positions are the basis of the description vector defined in Table III.
- *Emotion recognition system from speech*: The last sub-

[1]Similar to the most of facial expression recognition system, the proposed work use the Facial Action Coding System (FACS) proposed by Ekman et al. [19]

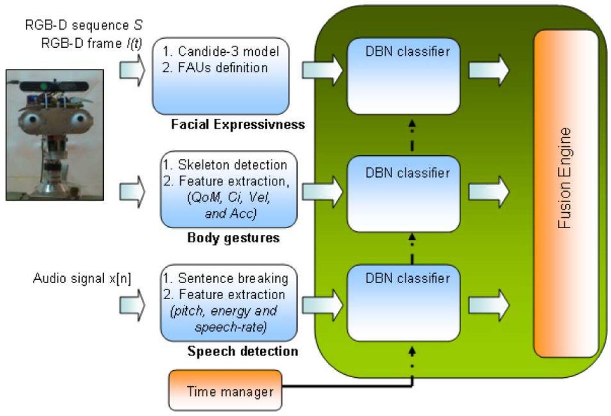

Fig. 6. Overview of the Multimodal emotion recognition system.

TABLE II
FACIAL EXPRESSIVENESS VECTOR d_f

EB	Eye-Brows	$\{AU1, AU4, none\}$
Ch	Cheeks	$\{AU6, none\}$
LE	Lower Eyelids	$\{AU7, none\}$
LC	Lips Corners	$\{AU12, AU15, none\}$
CB	Chin Boss	$\{AU17, none\}$
MF	Mouth's Form	$\{AU20, AU23, none\}$
MA	Mouth's Aperture	$\{AU24, AU25, none\}$

[2]C_i is a value for the contraction degree of the body, based on the relationship of the chest and the hands position. This relationship has been carried out by the area of the triangle defined by the three points, u, v and w, and the perimeter s.

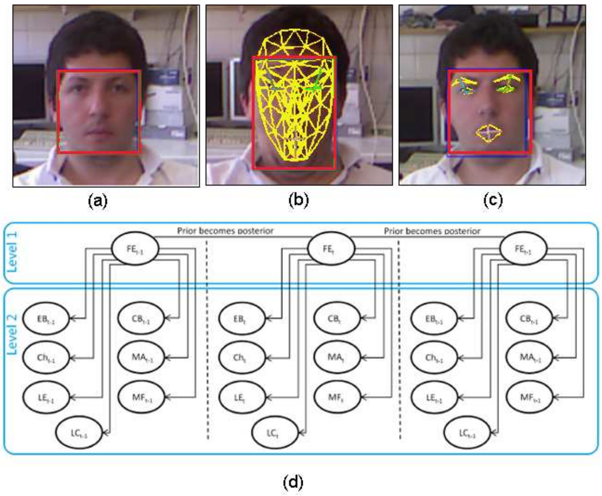

Fig. 7. a) RGB Image acquired by the sensor; b) Candide-3 recontruction model; c) Features extracted from the face mask; and d) Dynamic Bayesian Network, three time intervals are shown.

TABLE III
BODY GESTURES FEATURES VECTOR d_g

QoM^{lh}	Quantity of Motion	$QoM_i^{lh} = \frac{1}{N} \cdot \sum_{k=0}^{N} x_i^{lh} - x_{i-1}^{lh}$
Vel	Velocity	$Vel_i = \frac{1}{N} \cdot \sum_{k=0}^{N-1} \frac{(x_k^{lh} - x_{k-1}^{lh})}{N}$
Acc	Acceleration	$Acc_i = \frac{Vel_i - Vel_{i-1}}{t}$
C_i	Contraction index	$C_i = \sqrt{s \cdot (s-u) \cdot (s-v) \cdot (s-w)}$

system uses a similar structure to the methods aforementioned. Pitch, Energy and Speech-rate has been used for emotion recognition from speech. First, the audio signal $x[n]$ is pre-processed on-line in order to detect the presence or absence of speech (see Fig. 9a). Then, a speech features vector d_s= $\{Pt, En, Sr\}$ is calculated as shown in Table IV, considering N samples. This descriptor is also the input of the third DBN classifier, as is illustrated in Fig. 9b.

TABLE IV
SPEECH FEATURES VECTOR d_s

Pt	Pitch	$max\left\{Y(\omega) = \prod_{r=1}^{i=R} \|X(\omega r)\|^2\right\}$
En	Energy	$E = \frac{1}{N} \cdot \sum_{x=0}^{x=i} x[i]^2$

- *Fusion Engine:* Through the use of a multimodal system, it is possible to eliminate errors in the detection, by checking the results with those obtained through another modality. Thus, the decisions of each emotion recognition subsytem are combined in a fused decision vector that is analysed by the Bayesian network of three levels (see Fig. 10). The Time Manager in Fig. 6 synchronizes directly the output of all the classifiers.

B. *Affective Human Robot Interaction Scenario*

In this section, the scenario used for the affective behaviors learning is described. This scenario consists of a basic HRI,

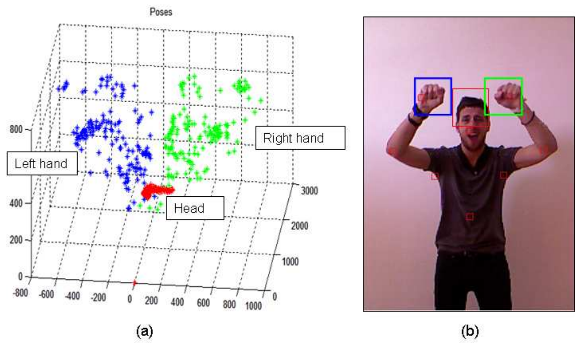

Fig. 8. a) In a emotional state of happiness, the 3D positions for the left hand, head and right hand are drawn in blue, red and green, respectively; and b) Given the human skeleton, a set of invariant features are extracted.

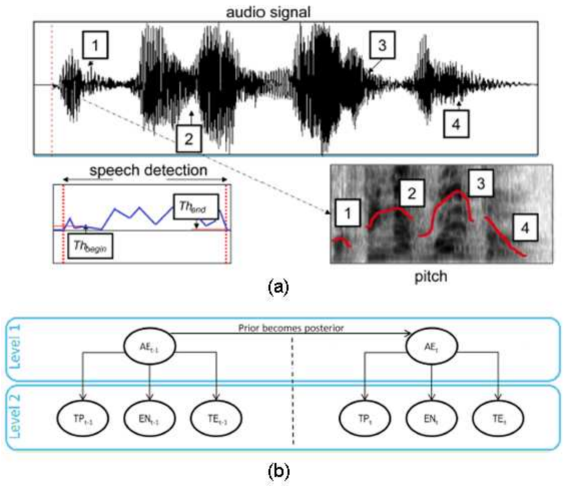

Fig. 9. a) The audio signal is processed in order to detect the beginning and ending of the sentences, and then, a set of features are extracted (*e.g., pitch*); and b) Two time intervals for the Dynamic Bayesian Network classifier.

in which a person interacts with the robot (or with other person). During this interaction, the robot builds the structure of the Bayesian Network, G, using Markov Chain Monte Carlo (MCMC) [4] and the parameter of each node in the graph. To do that, several experiments are conducted to illustrate the capability of the system to discover affective affordances associated with reactions applied to affective elements (objects and emotional states).

The proposed interaction consists of presenting a set of objects, with different colors, shapes and sizes. Besides, the human expresses different emotional states. The main intention is to generate a transition between two emotional states (*e.g.,* from sadness to happiness). In a first stage of the learning process, the person's reactions are analyzed by the robot as an observer. After, the own robot is part of the learning process, repeating the experiments. The main elements included in this HRI scenario are detailed below:

- Robot Loki: Loki is an autonomous mobile manipulator endowed with two arms (Each arm has 7 degrees-of-freedom (DOF)) and a expressive head named Muecas[3] (13 DOF), designed for researching in social robotics and HRI. It has learned basic skills for perceiving the world

[3]www.iadex.es

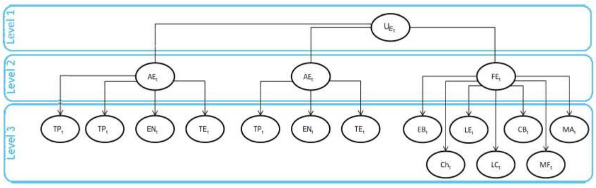

Fig. 10. The last DBN classifier combines the decision of each modality.

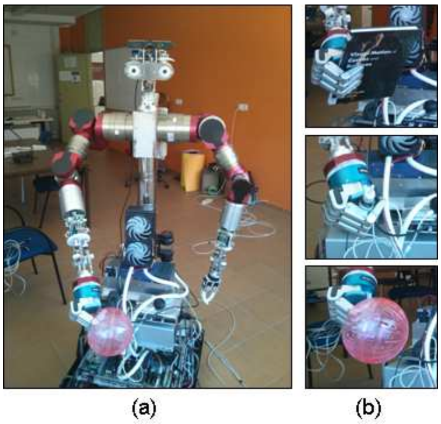

Fig. 11. The Loki robot, an AMM built as a collaboration among several Spanish universities and companies, following design and coordination by RoboLab.

and recognizing human emotions (see Fig. 11a).

- Objects: A set of different objects are present in the scenarios. The objects have different shapes, color and sizes. All of them can be grasped by Loki as is illustrated in Fig. 11b.
- Emotional States: Four different emotional states (*i.e.,* happiness, fear, anger and sadness) in conjunction with the neutral state are used as affective elements. Besides, these same emotional states are estimate with the multimodal emotion recognition system described in this paper.

V. Discussion

This paper has described the ongoing work of autonomous learning of affective behaviors. In order to socially interact with human, the robot should be able not only to understand users behaviours and intentions, but also to generate its own emotional state. Thus, affective affordances have been defined here as an extension of perceptual affordances. Affective affordances represent the relation between affective elements, effects and oportunities for robot's reactions. In this work, the basic concept has been explained, and also has been posed the idea of using the affective affordances for learning affective behaviors. In addition, the main stages of an affective affordances learning have been summarized.

In the proposal, the affective Human-Robot Interaction scenario for affective affordances learning has been described, as well as the basic skills that the Robot Loki has learned. Both basic skills, active perception and human emotional state recognition, play an important role in the design of learning strategies.

Future works will be focused on implementing the full learning algorithm. It will spend a considerable time studying, for example, the structure of the Bayesian Networks or choosing a representative set of objects and emotions. However, the presented idea of affective affordances is, without doubts, beneficial for other researchers in the robotic community working in affective HRI.

Acknowledgment

This work has been partially supported by the MICINN Project TIN2012-38079-C03-01.

References

[1] J. Piaget, ”The origins of intelligence in children”. Intl. Univ. Press. 1952.
[2] J. J. Gibson, ”The ecological approach to visual perception”. Boston: Houghton Mifin. 1979.
[3] M. Lopes, F. S. Melo and L. Montesano, ”Affordance-based imitation learning in robots”. In *Proceedings of the IEEE/RSJ International Conference on Robots and Systems*, pp. 1015-1021, 2007.
[4] S. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor, ”Learning object affordances: From sensorymotor coordination to imitation”. In *IEEE Transactions on Robotics*, Vol. 24, pp. 15-26, 2008.
[5] B. Ridge, D. Skocaj, and A. Leonardis, ”A System for Learning Basic Object Affordances using a Self-Organizing Map”. In *Proceedings of the 2008 International Conference on Cognitive Systems*, University of Karlsruhe, Karlsruhe, Germany, 2008.
[6] S. Montesano, M. Lopes, A. Bernardino, and J. Santos-Victor. ”Modeling Affordances using Bayesian networks”. In *Proceedings of the 2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2007.
[7] P. Osório, A. Bernardino, R. Martinez-Cantin, and J. Santos-Victor, ”Gaussian mixture models for affordance learning using Bayesian Networks”. In *Proceedings of the 2010 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pp. 4432-4437, 2010.
[8] M. Kammer, M. Tscherepanow, T. Schack, and Y. Nagai, ”From Affordances to Situated Affordances in Robotics - Why Context is Important”. Frontiers in Computational Neuroscience, 2011.
[9] J. Morie, J. Williams, A. Dozois, and D. Luigi, ”The Fidelity of ”Feel”: Emotional Affordance in Virtual Environments”. In *Proceedings of the 11th International Conference on Human-Computer Interaction*, 2005.
[10] Toward a Comparative Repository of Cognitive Architectures, Models, Tasks and Data in *http://bicasociety.org/cogarch/*.
[11] I. Infantino, ”Affective Human-Humanoid Interaction Through Cognitive Architecture”. In *The Future of Humanoid Robots - Research and Applications*, In-Tech, Edited by Dr. Riadh Zaier, pp. 147-164. 2012.
[12] M. Scheutz, P. Schermerhorn, C. Middendorff, J. Kramer, D. Anderson and A. Dingle. ”Toward Affective Cognitive Robots for Human-Robot Interaction”. In *Proceeding of 17th Innovative Applications of Artificial Intelligence Conference*, pp.1737-1738, 2005.
[13] A. J. Palomino, R. Marfil, J. P. Bandera and A. Bandera, ”A novel biologically inspired attention mechanism for a social robot”. In *Journal on Advances in Signal Processing*, 2011.
[14] A. Treisman and G. Gelade. ”A feature integration theory of attention”. In *Cognitive Psychology*, Vol. 12, no. 1, pp. 97136, 1980.
[15] J. Wolfe, K. Cave, and S. Franzel. ”Guided search: An alternative to the feature integration model for visual search”. In *Journal of Experimental Psychology: Human Perception and Performance*, Vol. 15, no. 3, pp. 419-433, 1989.
[16] F. Orabona, G. Metta and G. Sandini, ”A photo-object based visual attention model”. In *WAPCV 2007. LNCS (LNAI)*, L. Paletta and E. Rome, Eds., vol. 4840. Heidelberg: Springer, pp. 198-215, 2007.
[17] R. Marfil, L. Molina-Tanco, J. Rodríguez and F. Sandoval, ”Real-time object tracking using bounded irregular pyramids”. In *Pattern Recognition Letters*, No. 28, pp. 985-1001, 2007.
[18] C. Bundesen, T. Habekost and S. Kyllingsbaek, ”A neural theory of visual attention and short-term memory (ntva)”. Neuropsychologia, vol. 49, No. 6, pp. 1446-1457, 2011.
[19] P. Ekman, WV. Friesen, JC. Hager, ”Facial Action Coding System FACS”, The manual, 2002.
[20] F. Cid, J. A. Prado, P. Manzano, P. Bustos and P. Núñez, ”Imitation System for Humanoid Robotics Heads”. In *Journal of Physical Agents*, Vol. 7, pp. 22-29, 2013.
[21] J. Ahlberg, ”CANDIDE-3 - an updated parameterized face”. Report No. LiTH-ISY-R-2326, Dept. of Electrical Engineering, Linkping University, Sweden, 2001.