Face Recognition for Social Interaction with a Personal Robotic Assistant

R. Barea, L. M. Bergasa, E. López, M.S. Escudero and C. León

Abstract — **This work describes the main characteristics of a "social interactive" robot for which social human-robot interaction is important. The main goal of this work is the development of a robotic assistant (called SIRA) for elderly which serves several primary functions such us tele-presence, tele-medicine, intelligent reminding, safeguarding, mobility assistance and social interaction. We will focus our attention in to equip an intelligent robot assistant with a robust user-machine interface, which enable this robot to communicate, interact, and collaborate with human users in a natural and intuitive way. The social interaction module developed will be commented and finally, an application of the social interaction will be explained.**

Keywords — **socially interactive robot, robotic assistant, social interaction, human-robot interface, face recognition.**

I. Introduction

IN the last years, the number of elderly in need of care is increasing dramatically. Therefore, the society needs to find new technologies and alternative ways of providing care to the elderly, where the need for personal assistance is larger than in any other age group. Several factors suggest that now is the time to establish new applications in the home-care sector: Firstly, we actually have the technology (internet) to develop new applications in home health-care sector. Secondly, at the currently, the robots exhibit the necessary robustness, reliability, and level of capability. Thirly, the need for cost-effective solutions in the elderly care sector is larger than ever before. Aware of this necessity, nowadays exists several research groups working in this area and some important projects such us "Nursebot" [1] and "Morpha" [2].

On the other hand, from the beginning of robots, researchers have been fascinated by the possibility of interaction between a robot and its environment, by the possibility of robots interacting with each other and with humans. In this paper, we use the term "socially interactive robots" to describe robots for which social interaction plays a key role [3], this is, social learning and imitation, gesture and natural language communication, emotion, and recognition of interaction partners are all important factors. In this paper, we focus on peer-to-peer human–robot interaction. Specifically, we will describe a robotic assistant that exhibit the following "human social" characteristics: express and/or perceive emotions, communicate with high-level dialogue, learn/recognize models of other agents, establish/maintain social relationships, use natural cues (gaze, gestures, etc.), exhibit distinctive personality and character, and may learn/develop social competencies.

Socially interactive robots can be used for a variety of purposes: as research platforms, as toys, as educational tools, or as assistant aids. The common, underlying assumption is that humans prefer to interact with machines in the same way that they interact with other people. A survey of current applications is given in [4].

In order to offer a solution for described problem, a research group of Electronics Department at the University or Alcalá is working in two projects called TELEASISNET [5] (spanish acronym for Teleassistance System) and "SIRAPEM" (spanish acronym for Robotic System for Elderly Assistant) [6]. The main goals of this projects are the development of personal robotic aids that serves several primary functions such us tele-presence, tele-medicine, intelligent reminding, safeguarding, mobility assistance and social interaction [7].

This paper presents an initial application of our system in social interaction recognizing the user.

II. Social Interaction

Since already we have commented, in this work we will focus our attention in the task of social interaction between human and robots. When the user lives alone is deprived of social interaction, therefore, social engagement can significantly delay the deterioration and health-related problems. Robots cannot replace humans but can help to endure solitude by providing a communication interface with relatives and friends increasing the user contact with the outside world.

We have implemented an assistant robotic based on a commercial platform (PeopleBot robot of ActivMedia Robotics [8]) and a face with facial animation using OpenGL library based on the model developed by Keith Waters [9] and a synthesis/recognition system based on IBM Viavoice and Cloudgarden libraries using JAVA language. Figure 1 shows the social interaction interface developed.

The autors would like to express their gratitude to the Ministerio de Ciencia y Tecnología (MCYT) for their support through the project SIRAPEM (DPI2002-02193), to the Comunidad de Madrid for TEAPEM project (GR/SAL/0860/2004) financing.

R. Barea et al, are with the Electronics Department, University of Alcalá, Spain (phone: +34918856574; fax: +34918856591; e-mail: barea@depeca.uah.es; http://www.depeca.uah.es/personal/barea.

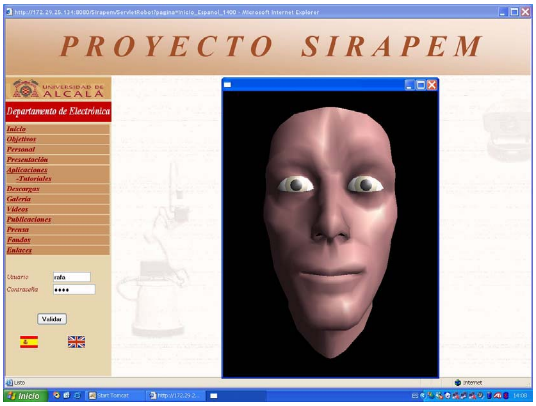

Fig. 1. Social interaction interface.

We are working to develop a robot that permits a natural human-robot interaction. To do this, we have focused our work in develop a dialogue as a join process of communication among human and robot. This involves sharing information using "spoken dialogs" or "pressing the screen", transmission of emotions and perception of the world [3].

A. Spoken dialog management

It has proven that an automatic speech recognition has made possible more natural human-robot interaction. Human speech can be noisy and ambiguous, therefore, any system that manages human-robot dialogs must be able to perform reliably even with noisy and stochastic speech input.

To solve the problem of spoken dialog management we use a Partially Observable Markov Decision Process (POMDP) proposed by Pineau and Thrun in [10], because a POMDP is a natural way of modelling dialog processes, especially when the state of the system is viewed as the state of the user. The partial observability capabilities of a POMDP policy allows the dialog planner to recover from noisy or ambiguous utterances in a natural and autonomous way. At no time does the machine interpreter have any direct knowledge of the state of the user, i.e, what the user wants. The machine interpreter can only infer this state from the (noisy) speech of the user. The POMDP framework provides exactly the right mechanism for modelling uncertainty about what the user is trying to accomplish. [11]. We have modelled the possible states, actions, transition probabilities, rewards, observations and observation probabilities that define the POMDP.

B. Emotions

Emotions play a significant role in human behaviour, communication and interaction. A good review of "basic emotions" is [12]. Artificial emotions are used in social robots for several reasons. The primary purpose, of course, is that emotion helps facilitate believable human–robot interaction [13]. Artificial emotion can also provide feedback to the user, such as indicating the robot's internal state, goals and (to an extent) intentions [14]. Lastly, artificial emotions can act as a control mechanism, driving behaviour and reflecting how the robot is affected by, and adapts to, different factors over time [15].

In our case, our robot can transmit emotions by mean of facial expressions, speech and body language.

B.1. Facial expression

To transmit facial expression we have modified the face model developed by Keith Waters [9]. Because our face is graphically rendered, many degrees of freedom are available for generating expressions. We can control eyebrows, eyeballs, eyelids, mouth with two lips, jaw and a pan/tilt head (figure 2).

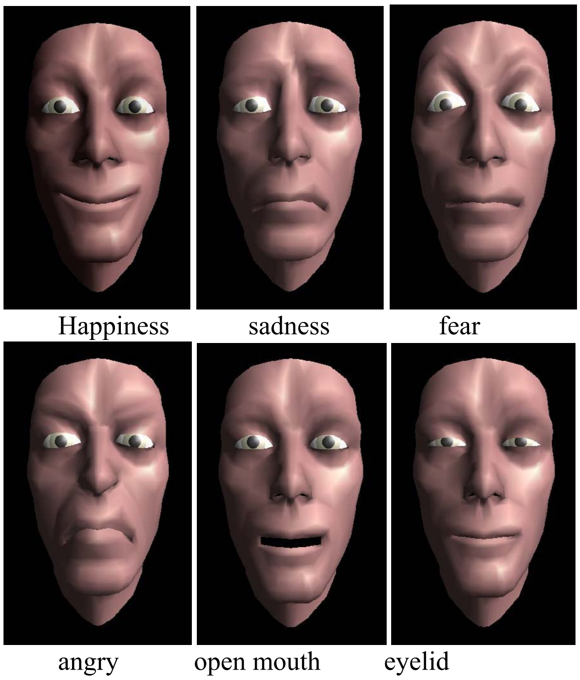

Fig. 2. Facial expressions.

B.2. Speech

Speech is a highly effective method for communicating emotion. The primary parameters that govern the emotional content of speech are loudness, pitch (level, variation, range), and prosody. To date, emotional speech has been used in few robot systems, probably because The quality of synthesized speech is significantly poorer than synthesized facial expression and body language [16].

Our synthesis/recognition system allows us to modulate the vocalization system changing the users (child, young or elderly, man or woman) depending on to whom it is relating and controlling the voice level and picth. For example, when the robot is frustrated, voice level and pitch are increased and if it is frightened, voice level is decreased. Another example is if the robot detects that is speaking with a boy, use a child voice.

B.3. Body language

In addition to facial expressions, non-verbal communication is often conveyed through gestures and body movement [16]. Over 90% of gestures occur during speech and provide redundant information [17]. We have described body movements for a number of basic emotions. For example, if the emotion detected by the robot is "fear", the robot moves away from the user, do slowly motions, put a fear facial expression and decrease voice level and speak slowly. On the contrary, the emotion is "happiness", put a "happy" facial expression, approaches to the user and speaks quickly and increase its voice level.

C. Perception of the world

To interact meaningfully with humans, social robots must be able to perceive the world as humans do, i.e., sensing and interpreting the same phenomena that humans observe. This means that, in addition to the perception required for conventional functions (localization, navigation, obstacle avoidance), social robots must possess perceptual abilities similar to humans. To obtain this, we are working in order that the robot is capable of tracking human features (faces, bodies, hands). It also need to be capable of interpreting human speech including affective speech, discrete commands, and natural language. Finally, it must have the capacity to recognize facial expressions, gestures, and human activity. To do it we process the image obtained by the camera using computer vision techniques.

III. Face Recognition In Social Interaction

The problem of user identification consists of the following: a PTZ camera placed in the robot acquire images of people in the work environment; an automated system extracts faces from these images and quickly identifies them using a database of known individuals. The system must easily adapt as people are added or removed from its database, and the system must be able to recognize individuals in the room. This section focuses on the face recognition technology that is required to address this real-world task. Face recognition has been actively studied [18], in this domain, techniques based on Principal Components Analysis (PCA) popularly termed eigenfaces [19], have demonstrated excellent performance. In our case, we use a simple, memory-based PCA algorithm for face recognition.

A. Database and Preprocessing

Our system use human face images from a database composed by 20 tightly-cropped images of different individuals with only minor variations in pose (∀20º) and facial expression. The faces are consistently positioned in the image frame, and very little background is visible. Figure 3 shows several faces of a person and figure 4 shows faces of some individuals.

For detecting faces, the system acquires an image and detects where exits some faces. To do it we use a PCA algorithm as can be see in figure 5.

After that, using the previous database compares the face obtained with faces stored in database using PCA. With this information the user is recognized. If the system does not recognize the user, ask for his name and introduced a set of faces of him in the database and the database is trained again. This way, the system stores all users and the database is increased on-line.

Fig. 3. Images of a user.

Fig. 4. Images from different users.

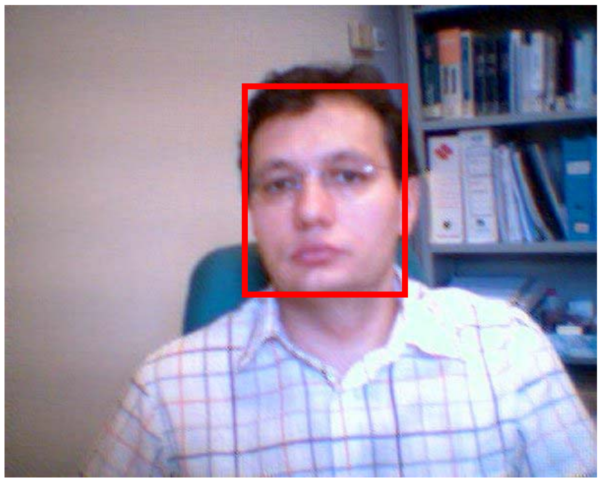

Fig. 5. Face detection.

B. Principal Components Analysis (PCA)

One of the most widely used baseline for face recognition is *eigenfaces* [20]. It employs Principal Components Analysis (PCA) which is based on the discrete Karhunen-Loeve (K-L), or Hotelling Transform, and is the optimal linear method for reducing redundancy, in the least mean squared reconstruction error sense. PCA has become popular for face recognition with the success of *eigenfaces*.

PCA algorithm projects points in R^d into R^m, (where m#d, and typically $m<<d$). For face recognition, given a dataset of N training images (X), it is create a N d-

dimensional vectors (X_1, X_2, … X_N). The principal components of this set of vectors is computed to obtain a d x m projection matrix, W.

Now, the image X may be compactly represented as *weights*, $2_i = (2_{i1}, 2_{i2}, ...2_{im})^T$, such that $Z_i = :+W2_i$ approximates the original image, where : is the mean of the X_i and this reconstruction is perfect when m = d. The columns of W form an orthonormal basis for the space spanned by the training images.

Each training image is first projected into the eigenspace, and represented as a weight vector $2_i = W^T(X_i-:)$. The centroid of the weight vectors for each person's images in the training set is computed and stored.

When a test image is presented to the system, it is first projected into the eigenspace and its weight vector 2_{new} is computed. 2_{new} is then compared against the stored weight vectors, Θ, and the 2_K that is closest 2_{new} is located. The label of 2_{best} is returned as the identity of the face represented by 2_{new}.

C. Results

The system implemented allows to explore the environment using a vision module for detecting users. When a user is detected, the system asks him if he wants to realize some activity of that it has programmed and depending on the response of the user it realizes a certain action. Because the user is recognized it is possible to know information about his preferences and likes and therefore the social interaction can be more natural.

For example, figure 8 shows a dialogue corresponding to the accomplishment of a task of guided in an environment by means of the acoustic interface. It is possible to see the conversation and an image of the graphical interface.

Robot: Hello.I'm the SIRA robot. Do you want to start the guidance?
Usuario: Yes
Robot: What is the new destination room?
Usuario: Office 4
Robot: Do you want to go to office 4?
Usuario: Yes

(...when reaches the destination)
Robot: OK. This is the destination room. What is the new destination room?

Fig. 8. Guidance example.

IV. CONCLUSIONS

This work reports the initial design and results of a social robotic assistant for elderly and/or sick people. This system could be implanted in medical centers and at homes of elderly and/or sick people y/o and it will provide several primary functions such us tele-presence, tele-medicine, intelligent reminding, safeguarding, mobility assistance and social interaction. The results obtained using users recognition improve the social interaction with user because it is possible to know some information about these users.

REFERENCES

[1] Nursebot Project. http://www-2.cs.cmu.edu/~nursebot
[2] Morpha Project. http://www.morpha.de
[3] T. Fong, I. Nourbakhsh and K. Dautenhahn. "A survey of socially interactive robots". Robotics and Autonomous Systems 42. 143–166. 2003.
[4] M. Lansdale, T. Ormerod, Understanding Interfaces, Academic Press, New York, 1994.
[5] TELEASISNET Project. "Sistema de teleasistencia a través de internet". Departamento Electrónica. Universidad de Alcalá. 2003.
[6] SIRAPEM Project. "Sistema Robótico para Asistencia a Personas Mayores". Departamento Electrónica. Universidad de Alcalá. 2003.
[7] N. Roy, G. Baltus, D. Fox, F. Gemperle, J. Goetz, T. Hirsch, D. Margaritis, M. Montemerlo, J. Pineau, J. Schulte and S. Thrun. Towards Personal Service Robots for the Elderly. Baltus et al. Computer Science and Robotics. Carnegie Mellon University.
[8] ActivMedia Robotics. http://www.robots.activmedia.com
[9] F. Parke and K. Waters. "Computer Facial Animation". AK Peters. 1996.
[10] J. Pineau and S. Thrun. "Hierarchical POMDP Decomposition for A Conversational Robot". Carnegie Mellon University. Robotics Institute.
[11] N. Roy, J. Pineau and S. Thrun. "Spoken DialogManagement for Robots". Robotics Institute, Carnegie Mellon University.
[12] P. Ekman, "Basic emotions", in: T. Dalgleish, M. Power (Eds.), Handbook of Cognition and Emotion, Wiley, New York, 1999.
[13] T. Ogata, S. Sugano. "Emotional communication robot: WAMOEBA-2R emotion model and evaluation experiments". Proceedings of the International Conference on Humanoid Robots, 2000.
[14] C. Bartneck, M. Okada. "Robotic user interfaces". Proceedings of the Human and Computer Conference, 2001.
[15] F. Michaud, et al.,. "Artificial emotion and social robotics". Proceedings of the International Symposium on Distributed Autonomous Robotic Systems, 2000.
[16] C. Bartneck. "eMuu: an emotional embodied character for the ambient intelligent home". Ph.D. Thesis, Technical University Eindhoven, The Netherlands, 2002.
[17] D. McNeill, Hand and Mind: What Gestures Reveal About Thought, University of Chicago Press, Chicago, IL, 1992.
[18] T. Sim, R. Sukthankar, M. Mullin and S. Baluja. "Memory-based Face Recognition for Visitor Identification". Proceedings of International Conference on Automatic Face and Gesture Recognition, 2000.
[19] R. Gonzales and R. Woods. "Digital Image Processing". Addison-Wesley, 1992.
[20] M. Turk and A. Pentland. "Eigenfaces for recognition". Journal of Cognitive Neuroscience, 3(1), 1991.