Universita Degli studi di Genova

A dialog management system for a proactive virtual assistant in smart environments

Literature Review

Polaka Surendra - 4846909

Supervisors : Luca Buoncompagni,Marta Cristofanini ,Fulvio Mastrogiovanni ,Alessandro Sperindè

1 Motivation

In recent times, there is a rise in the population of elderly individuals. This idea motivates the technology companies research community to provide, at home, healthcare services for the elderly, such that they can live safely and independently for longer periods. The ability to perform Activities of Daily Living without assistance from other people can be considered as a reference for the estimation of the independent living level of elderly individuals. Nowadays caregivers and family members need to take time to support elder people. In the future caregivers' demand may increase and it will create too much burden to a family. Moreover in the present scenario care has been shifting from Hospitals to community-based care and Home care. Elder people should have the self-confidence to take care of their own health care management to avoid mental and physical depression. By the recent survey Information and Communication Technology is used in Health care facilities such as smartphones, Tablets. However, Elder people are not familiar with smartphones and Tablet Applications [1]. Due to the Above Instances, It will be too much use if the devices are Human friendly.A robot partner is the best choice for the Problem

2 Introduction

Socially Assistive Robots (SAR) provide a viable solution. Several papers examined different scenarios and the use of SAR [2] as well as their application for cognitive training [3]. Moreover, research has shown that SAR yields better results than other technologies, like tablets.Social Assistive Robots enter our homes and workspaces to assist and sustain independent living. Thus, robots operate in environments specifically designed for humans.Distinguish social robots from conventional robots by emphasizing that social interaction with users plays a key role. Social robots are envisioned to autonomously interact with humans in a socially meaningful way [4]. To work with humans in environments designed for humans, robots should be designed optimally for such conditions, in form, behavior, and personality.For robots to communicate and collaborate with people in a natural way, it is important that they possess the capacity for social inter-action. Emotion is a vital ingredient in social interaction,and is conveyed in part through visual cues (facial expression, pose, move-ment, etc.) and tone of voice [5-6].

Natural human-human interaction is multimodal, and therefore social robots should also be capable of multimodal.interaction [7]. Multiple modalities enable robots to perform the same action in different ways. However, in certain situations the choice of modalities is unambiguous. If the user is not looking at the robot, speech could be used to obtain the attention of the user. However, often the same action can be performed in multiple manners. If the user is looking at the robot, it could do a pointing gesture to direct user's attention. Moreover, it could combine speech and gesture and use both modalities simultaneously. The exact choice of modalities should depend on the user profile, which we describe with modality preferences.But in Our case we will mainly focus on Dialogue based Interaction.

Dialog-based interaction models are excellent for information exchange and resolving differences to achieve common ground.In human-robot interaction,People tend to attribute intent to technology. This is reinforced in the case of robots, since they have physical "bodies" with which to interact with the world. Young [8] argues that this "embeddedness in the physical world and the socially situated context of interaction creates a unique and affect-laden sense of active agency similar to that of living beings. In a sense, therefore, for many people, interacting with a robot is more akin to interacting with an animal or another human being than with a technology." This agency leads people to believe that the robot makes autonomous, intelligent decisions [9].Since the thesis is a work which concerns different topics,the literature review will deal with below given subtopics Dialogue-based interactions,proactive agents, assistive robotics, sentiment analysis, cognitive exercises, statistical tools to evaluate human-centered experiments.proactive agents, assistive robotics

3 Relevant Literature

In the literature, most work deals with different types of dialogue-based interactions between humans and robots.one of the biggest challenges in robotics is actually understanding human interaction. Unlike games, where the number of available actions is always finite (final) and therefore can be known in advance, in human-robot interaction it is often impossible to create a universal evasion mechanism that an AI system could use, even if the scenario is quite limited. This is supported by the analysis of Riek [10], who found that WoZ is most often used in place of natural language processing and nonverbal behaviour.Cloud robotics allows the robot to benefit from the far superior computing and storage resources of data centres - while saving its own energy (which is critical for mobile robots). This is useful for a variety of tasks, including natural language processing, object recognition, and navigation. As always, there are some drawbacks, notably potential difficulties in controlling the robot due to latency and limited information, and if the robot is too dependent on the cloud, it will become useless if the network fails.

During the Interaction with Robots.A person's emotions play an important role, which are usually measured directly by self-report or indirectly by observation-based methods in HRI experiments. The majority of HRI experiments that have addressed the following issues.Most HRI experiments that have addressed emotion have used categorical models of emotion representation; these models are typically derived from Ekman's "Basic Emotions' and typically include 6-8 states.

The statistical tools to measure the emotion in Human Robot Interaction geneva Emotion Wheel.Although the robotics literature rarely mentions a third type of Affect classification, which is based on valuation theory [11]. This approach defines emotions as "a process in which multiple components, such as physiological responses, cognitive representations (of both the triggering events and self-perceived response patterns) are synchronized over a limited period of time" [12]. For example, emotion components include physical symptoms (such as endocrine levels) and the user's subjective feelings. Proponents of the componential approach believe that observing the "response "chronization" measured components provides an empirical measure of Emotion in humans can be [13].

Human-robot interaction in terms of the framing effect. The framing effect is an example of a cognitive bias that affects a person's decision depending on whether it is presented as positive or negative. Here, we implemented verbal communication content based on positive or negative framing. In addition, we conducted a demonstration experiment to examine the effects of each expression on the motivation of the elderly.

To enhance good output in HRI. Robots should have an ability independently plan their actions and interact with humans, recognizing human emotions is critical. For most people, nonverbal cues such as pitch, volume, range, and speech rate are efficient carriers of emotion. The properties of the sound of a spoken voice likely contain crucial information about the speaker's emotional state. In this framework, a machine could use such properties of sound to detect emotions.It can be performed through "Affective Computing" [14], according to which the main goal is to build machines that recognize, express, model, communicate, and respond to users' emotion indicators.Humans cannot always hope that robots will be able to respond in a timely and rational manner, especially if they are not able to capture all affective information through their sensors.

The emotional features that the robot needs to capture are not always provided by different sources of the human body at the same time. Perhaps the information collected is not robust enough or the robot lacks a specific sensor to extract the emotional feature. Towards this goal, this research is based on the possible effects of some crucial language features on emotion detection in communication with humans. It is well known that emotions cause mental and physiological changes that are also reflected in the uttered speech [15] . It is possible to find connections between emotional cues in speech, and these can be used to learn about human emotions. Once such connections are learned, it is theoretically possible to compute the features and then automatically detect emotions in human speech utterances, keeping in mind that the emotional content of speech does not depend on the speaker or the lexical content. Decoding emotions in speech through various features is a challenging research topic that has become increasingly important.n robotics, because of the emotional factors that the robot can process and learn in social situations. A variety of different features have been used in classifying emotions based on language, and no rule has yet been established to follow.

Major Research theories on shows results The study of Speech Emotion Recognition focused on measuring speech quality using two features, pitch and Intensity. These features were chosen because they reflect even very small deviations from the more familiar SER features such as pitch and intensity of an utterance. Moreover, unlike certain observable prosodic features that can be influenced, pitch and Intensity are assumed to reflect only genuine emotions. Pitch has been used extensively in clinical diagnosis, as in [16], more than speech recognition applications. pitch and Intensity have been used sparingly along with many other features for emotion recognition, as indicated in [17]. The based on pitch, Intensity and their combination. These studies focused on the recognition of neutral and six basic emotions, namely happiness, surprise, anger, sadness, fear, and disgust,. The first two are emotions with positive valence, and the last four are emotions with negative valence.

4 Conclusion

In this Literature Review . I have explained mainly on Human Robot Interaction based on Cognitive Exercise and emotional Recognition based up on voice features , statistical tools to measure the output of the system .The state of art of both has been investigated to establish a context for the future work. Even though these topics have been widely treated in literature, the thesis aims at exploring the best solution which allows the Emotional Recognition of user measured through Oliver API which Improve Dialog management system with proactive Virtual assistant for Elder people.

5 References

[1] Y. Chen, S. Y. M. Lee, and C.-R. Huang. Are emotions enumerable or decomposable? and its implications for emotion processing. In Proceed- ings of the 23rd Pacific Asia Conference on Language, Information and Computation, pages 92–100. Association for computation Linguistics, 2009.

[2] S. M. Rabbitt, A. E. Kazdin, and B. Scassellati, "Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use," Clinical Psychology Review, vol. 35, pp. 35 – 46, 2015.

[3] A. Tapus, C. Tapus, and M. J. Mataric, "The use of socially assistive robots in the design of intelligent cognitive therapies for people with dementia," in 2009 IEEE International Conference on Rehabilitation Robotics, June 2009, pp. 924–929.

[4] K. Lee, W. Peng, S. Jin, and C. Yan, "Can Robots Manifest Personality?: An Empirical Test of Personality Recognition, Social Responses, and Social Presence in Human-Robot Interaction", Journal of Communication, vol. 56, no. 4, pp. 754-772, 2006.

[5] Mark Coulson. 2004. Attributing emotion to static body postures: Recognition,accuracy, confusions, and viewpoint dependence. Journal of nonverbal behavior 28, 2 (2004), 117–139. [6] Elwin Marg. 1995. Descartes' error: emotion, reason, and the human brain. Optometry and Vision Science 72, 11 (1995), 847–848.

[7] T. Kollar, A. Vedantham, C. Sobel, C. Chang, V. Perera, and M. Veloso, "A multi-modal approach for natural human-robot interaction," in Social Robotics. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 458–467.

[8] J. E. Young, J. Sung, A. Voida, E. Sharlin, T. Igarashi, H. I. Christensen, and R. E. Grinter, "Evaluating human-robot interaction," International Journal of Social Robotics, vol. 3, no. 1, pp. 53–67, 2011.

[9] C. Bartneck, M. Verbunt, O. Mubin, and A. Mahmud, "To kill a mockingbird robot," 01 2007, pp. 81–87.

[10] L. D. Riek, "Wizard of oz studies in hri: a systematic review and new reporting guidelines," Journal of Human-Robot Interaction, vol. 1, no. 1, pp. 119–136, 2012.

[11] Didier Grandjean, David Sander, and Klaus R Scherer. 2008. Conscious emo- tional experience emerges as a function of multilevel, appraisal-driven response synchronization.Consciousness and cognition 17, 2 (2008), 484–495.

[12] Klaus R Scherer, Vera Shuman, Johnny RJ Fontaine, and Cristina Soriano. 2013. The GRID meets the Wheel: Assessing emotional feeling via self-report. Compo- nents of emotional meaning: A sourcebook (2013), 281–298

[13] Klaus R Scherer. 2005. What are emotions? And how can they be measured? Social science information 44, 4 (2005), 695–729

[14] R. Picard, Affective Computing. The MIT Press, United States, 1998.

[15] J.Q. Wang, N. Trent, E. Skoe, M. Sams and N. Kraus, "Emotion and the auditory brainstem response to speech," Neuroscience Letters, vol. 469, no. 3, pp. 319–323, 2010

[16] P. Lieberman, "Some acoustic measures of the fundamental periodicity of normal and pathological larynges", Journal of the Acoustic Society of America. 1963, vol.35, no. 3

[17] Kwang-Dong Jang and Oh-Wook Kwon, "Speech Emotion Recognition for affective human–robot interaction," SPECOM Conference 2006, St. Petersburg, 25-29 June 2006, pp. 419-422