International Journal of Enhanced Research in Science Technology & Engineering, ISSN: 2319-7463
Vol. 3 Issue 1, January-2014, pp: (362-365), Impact Factor: 1.252, Available online at: www.erpublications.com

Cognitive Learning for Social Robot through Facial Expression from Video Input

Neeraj Rai[1], Deepak Rai[2]
[1]Department of Automation & Robotics, [2]Department of Computer Science & Engg.
[1,2]Ajay Kumar Garg Engineering College, Ghaziabad, UP, India

Abstract: This paper presents research on Cognitive Learning through Facial Expression for making Artificial Intelligent Social Robot. In this paper we have made a system for Cognitive Learning through facial expression from continuous video input. The goal of this research paper is to train a machine with the help of facial expression of any person. In this we have we have used the new technology namely, MATLAB. Use of MATLAB have made the implementation task easier and also increased the performance like speed and accuracy. The past work is very complex and has slow working. Our model works on our own proposed algorithm which has made it easier to implement. In this paper we have reviewed the previous work, discussed current concepts, tools and purpose of further work with initial finding on new Robotic environment and concepts through simulation in MATLAB. The Model was created on the basis of Cognitive Learning through Facial Expression as a part of this research paper to support customizable sets of concepts and theory in order to easily interface with diverse sets of educational material. Developing new methods to improve recognition performance is a major concern of this research work. The results are discussed in terms of previous research and theory. Implications of the study for practice and suggestions for future research are also discussed.

Keywords: Cognitive Learning, Facial Expression, MATLAB, Social Robot.

1. INTRODUCTION

There are so many types of robot like nano robot, industrial robot, mobile robot, programmable robot and many more, but, one of them is Artificial Intelligent Robot. In this two type of robot comes one is Humanoid and another is Cog Robots. The name Cog was given by MIT on the basis of word Cognitive because all things like Understanding, Learning, Speech Recognition and Perception comes under Cognitive Science. The COG Robots have capacity to understand and express the emotions and also have feelings of touch ability as like human. The implementation will be based on models of the development of Cognitive Learning through Facial Expression of human being which account for behaviors observed expression on human face and also learn "what changes has to be done by human face within mean time during analysis of our Robot.

The Cog Robots are also known as Social Robots because they are able to detect and understand natural human and cues the low-level social conventions that people understand and use every day, such as head nods or eye contact so that anyone can interact with them without special training or instruction. They also employ these conventions to perform an interactive exchange. The necessity of these abilities influences the robots control-system design and physical embodiment.

Imagine the year 2019. Mr. Smith, 90 years old, is still able to live autonomously thanks to his social robot Suzy. Suzy cleans his house, monitors his health and plays card games with him. Mr. Smith states: *"He is my best friend and I cannot live without him anymore."* Will this be reality ten years from now [1]? It is possible in upcoming days; the social robots will be able to build relationship like friendship. Fong et.al in 2003 [2] stated that "If technology adheres to human social expectations, people will find the interaction enjoyable, feeling empowered and competent." Therefore, it is important to study hedonic factors as well, to get a more complete view of which factors play an important role in the acceptance and usage of social robots

Cognitive Learning is a part of Cognitive Learning through Facial Expression. Cognitive Learning through Facial Expression (ToM) is a cognitive capability that allows us to understand another's internal states (intention, goal, and belief) and predict future behaviors of others [3] [4]. From the observation of other's behavior, facial expression, and speech, we can infer the person's internal state (emotions, thought, decision making, and plans). It was known that this function is supported by widely distributed areas of human brain [5] [6]. According to this we develop a model on Cognitive Learning through Facial expression. Scassellati in 2005 built "finding faces and eyes and distinguishing animate from inanimate stimuli" functions for humanoid robots [7].

The problem of finding and analyzing faces is a foundational task in computer vision. Though great strides have been made in face detection, it is still challenging to obtain reliable estimates of head pose and facial landmarks, particularly in unconstrained "in the wild" images. Ambiguities due to the latter are known to be confounding factors for face recognition [8]. Eye detection is required in many applications like eye-gaze tracking, iris detection, video conferencing, auto-stereoscopic displays, face detection and face recognition. R.Tanmay et.al in 2009 modeled every facial landmark as a part and use global mixtures to capture topological changes due to viewpoint [9]. We have shown that tree-structured models are surprisingly effective at capturing global elastic deformation, while being easy to optimize unlike dense graph structures [10].

2. RELATED WORK

2.1. Existing System

Kismet is an active vision head augmented with expressive facial features. Kismet was designed to receive and send human-like social cues to a caregiver, who can regulate its environment and shape its experiences as a parent would for a child visual perception (omitting the eyelids!). The eyes can turn independently along the horizontal (pan), but turn together along the vertical (tilt). The neck can turn the whole head horizontally and vertically, and can also crane forward. Two cameras with narrow "foveal" fields of view rotate with the eyes. Two central cameras with wide fields of view rotate with the neck. These cameras are unaffected by the orientation of the eyes. Kismet has three degrees of freedom to control gaze direction, three degrees of freedom to control its neck, and fifteen degrees of freedom in other expressive components of the face, including eyebrows (each with two degrees of freedom: lift and arch), ears (each with two degrees of freedom: lift and rotate), eyelids (each with one degree of freedom: open/close), a mouth (with one degree of freedom: open/close), and lips which can curl at each of the four corners. The robot is able to show expressions analogous to anger, fatigue, fear, disgust, excitement, happiness, interest, sadness, and surprise as shown in fig.1 & 2 [11], which are easily interpreted by an untrained human observer.

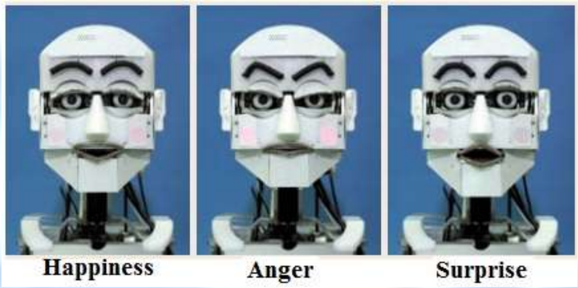

Fig.1. Basic facial expressions of kismet

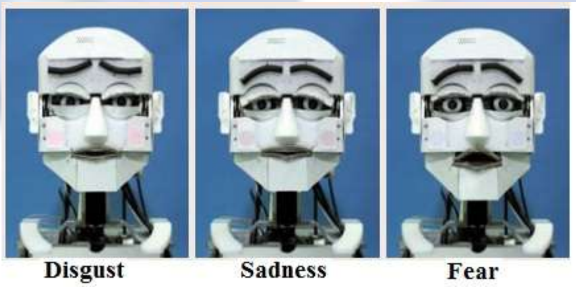

Fig.2. Basic facial expressions of kismet

To perceive its caregiver, Kismet uses a microphone, worn by the caregiver, and four color CCD cameras. The visual system on Kismet differs from Cog's in the configuration and type of cameras. Kismet has two single-board CCD cameras, one behind each eye, that have a narrow field of view slightly larger than the foveal cameras on Cog. Between the eyes, there are two unobtrusive central cameras fixed with respect to the head, each with a wider field of view but correspondingly lower acuity. This configuration leads to a less anthropomorphic visual sensing system, but has the benefit that certain visual tasks become simpler to implement. For example, smooth-pursuit tracking of a visual stimulus (that is, moving the eyes to maintain the object within the center of the field of view) becomes simpler when the two cameras along the robot's midline are used. Because these cameras do not move with the eyes, the visual processing required for tracking need not compensate for camera motion, leading to a simpler tracking algorithm.
The computational system for Kismet is considerably more heterogeneous than Cog's, although the components used for vision are nearly identical. Kismet's vision system is implemented on a network of nine 400 MHz commercial PCs running the QNX real-time operating system. Kismet also has a motivational system which runs on a collection of four Motorola 68332 processors. Machines running Windows NT and Linux are also networked for speech generation and recognition respectively.

3. PROPOSED SYSTEM

In our research work we have divided our task into following segment by which our task becomes easier to maintain the work and make an algorithm to complete our task in following step:

1. First of all, we divide the video which is perceived by the system through video camera into number of frames of images.
2. After that we detect the human face and part of human face like eye, nose, mouth and eyebrow within the images.
3. In this step, we find the changes sequentially in our image which is divided one by one frame set and we detect these changes on the basis of change in pixel value on comparison between two images.
4. At Last, we train these set of images one by one with the help of neural network tool nnstart and due to this our model start learning.

In whole scenario we only develop learning not the representation of facial expression. Our model only does the Pattern recognition (face detection and face parts detection) and power of itself learning. This is very critical to embed all features within a single robot to makes it multipurpose robots. The overall proposed system is shown in fig.3.

In real world all type of things are develop but they don't integrate within a single unit. If we merge all things then model goes to failure condition because different model work on different platform that's why they don't integrate and it is necessary to build whole model within a single platform by using the design of the algorithm for the whole scenario after that it implemented systematically according to their algorithm.

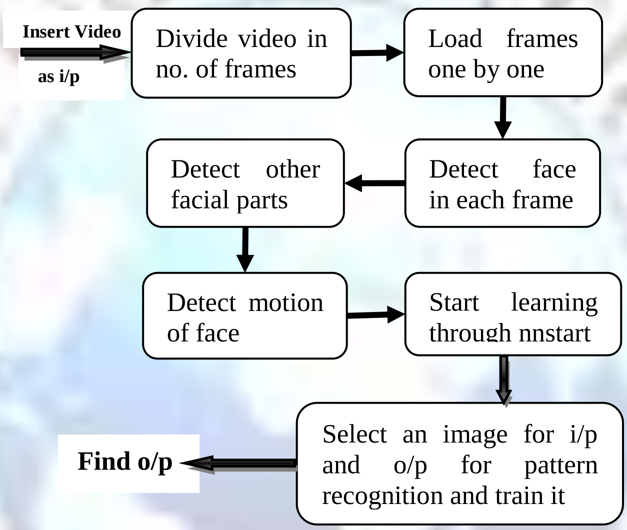

Fig. 3: Proposed System

4. CONCLUSION

Existing work of cognitive learning through facial expression by Scassellati et.al worked perfectly in determining social cues anthropology and sociology but lacks in some edges which can be improved by our above proposed functions. By utilizing various function of MATLAB we can enhance the performance of machine expert in recognizing the facial expression. Current model has been in developing phase and this is based on new technology. This model may provide assurance of more intelligent and fast learning methodology for detection of facial expression. This can be further improved and enhanced according to the need of time.

REFERENCES

[1]. Klamer, Tineke, and S. Ben Allouch. "Acceptance and use of a social robot by elderly users in a domestic environment." Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2010 4th International Conference on-NO PERMISSIONS. IEEE, 2010.
[2]. Fong, Terrence, Illah Nourbakhsh, and Kerstin Dautenhahn. "A survey of socially interactive robots." Robotics and autonomous systems 42.3 (2003): 143-166.
[3]. D. G. Premack, and G. Woodruff, "Does the chimpanzee have a theory of mind?," Behavioral and Brain Sciences, vol. 1, pp. 515-526, 1978.
[4]. C. Zimmer, "How the mind reads other minds," Science, vol. 300, pp. 1079--1080, 16 May 2003.

[5]. Siegal, Michael, and Rosemary Varley. "Neural systems involved in'theory of mind'." Nature Reviews Neuroscience 3.6 (2002): 463-471.
[6]. J. Call, and M. Tomasello, "Does the chimpanzee have a theory of mind? 30 years later," Trends in Cognitive Sciences, vol. 12, no. 5, pp. 187-192, 2008.
[7]. B. Scassellati, Foundations for a Theory of Mind for a Humanoid Robot, Ph.D. Thesis, Massachusetts Institute of Technology, 2001.
[8]. Zhao, Wenyi, Rama Chellappa, P. Jonathon Phillips, and Azriel Rosenfeld. "Face recognition: A literature survey." Acm Computing Surveys (CSUR) 35, no. 4 (2003): 399-458.
[9]. Rajpathaka, T., R. Kumarb, and E. Schwartzb. "Eye detection using morphological and color image processing." Proceeding of Florida Conference on Recent Advances in Robotics. 2009.
[10]. Zhu, Xiangxin, and Deva Ramanan. "Face detection, pose estimation, and landmark localization in the wild." Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. IEEE, 2012.
[11]. http://www.eat.lth.se/fileadmin/eat/Litteraturkompendium_MAMN10.pdf.