Human Interaction with Binocular Vision Robots in Ubiquitous Environment

Janaka Chaminda Balasuriya Keigo Watanabe Kiyotaka Izumi
Department of Advanced Systems Control Engineering
Graduate School of Science and Engineering
Saga University, 1-Honjomachi, Saga 840-8502, Japan

Abstract

Human interactive robotics is a fast developing area today. There are many aspects to be fulfilled for a robot to be more similar to human behavior. For example, a robot with a binocular vision head needs to recognize various gestures of the humans in order to respond to them accordingly. It is also necessary to have a sense of environmental changes, especially if interaction is being made while in motion. This is achieved with the help of an overhead monitoring system. Making the robot be "ubiquitous" will be possible by manipulating Bluetooth Wireless Technology. In order to apprehend the demand of more human like intelligent system, this research project is planned to search for new solutions for better achievement.

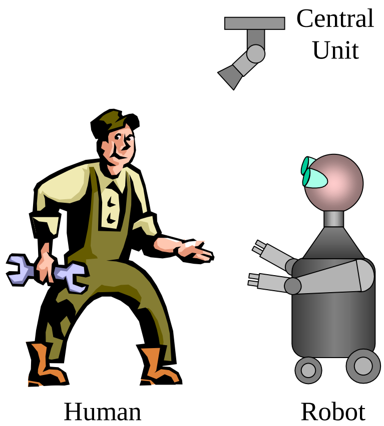

Figure 1: Human interaction

1 Introduction

For few years, research is going on to assist people by finding new robotic solutions. In the near future there may be robots in everywhere as our partners. These "ubiquitous" or present everywhere robots are human interactive not only to voice commands but also to human gestures as well (see Figure 1). In conversation, humans are capable of getting sense of other activities going around themselves. These include static / dynamic environment, environmental conditions, other sounds and noises in the vicinity, etc. But unfortunately there are human limitations as well. For an example, they are incapable of presenting themselves in anytime, anywhere or anyway. These may be due to human limitations or due to some natural disasters like fire, earthquakes, etc. However, it is useful to have a companion, who can represent a human. This may include basic things like; watching, listening, understanding, communicating as well as many other things. An attempt to build such a human interactive intelligent system is shown in Figure 1.

2 Binocular Vision

Humans are capable of getting a vision in breadth, width and depth, i.e. three-dimensional view. This has gained humans lot of features of characterizing the object in concern. In humans, the brain gets images from both (bi) eyes (ocular) at the same time and combines those two images into one, to make vision. The images that the brain gets from the eyes are however slightly different from each other and this small difference is used to work out how far away an object is. This is called the depth perception. It also helps to work out how quickly an object is moving towards or away from a person. This is the movement perception [1]. In order to interact effectively with humans, Ubiquitous robots require such information. Some examples are, to recognize a particular object or person (as a requested by third party interactor), to clearly identify face impressions / body gestures (while in conversation), to guidance in motion, etc.

2.1 Object Tracking

In order to do any of the above tasks, the primary objective is to obtain a clear view of the object. This requires continuous gazing at the right place at the right time. In other words, it is required to focus the lenses of the vision system according to the position changes of an object, especially if it is moving. As these trajectories of motion are not always simple as straight lines, the tracking mechanism becomes complicated and advanced. For a binocular vision system, both eyes (or cameras) should be focused separately over the object under consideration continuously.

2.2 Fuzzy Logic System

Matlab simulation results of one such tracking program written using fuzzy reasoning, is given in Figure 2. Once received the position of the object, relative distance to it from the current focal point is calculated. This is taken as 'Distance Error (DE)'. In addition to above, considering the motion of the object and the robot; rate of change of distance error (RateChangDE) is calculated. Obtaining these two values, will serve as the inputs to the Fuzzy Logic system, resulting in controlling the focal length of the lens (FocusLens). Rules such as; If DistanceError is positive big and RateChangDE is positive big then FocusLens is positive big, If DistanceError is negative small and RateChangDE is negative small then FocusLens is negative small, etc. are used. Here, for simplicity, a straight-line motion has been considered. But in realistic situation, the trajectories of motion will be much more complex. An adaptive neuro-fuzzy system will be much useful.

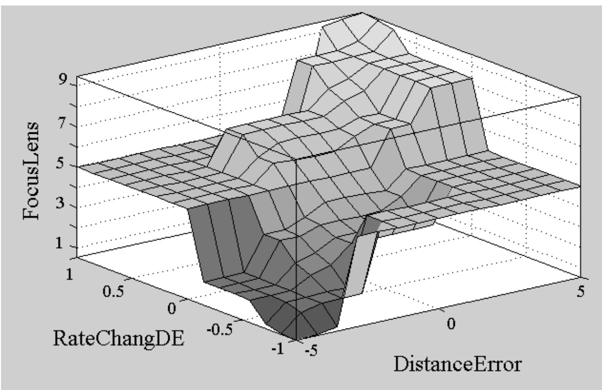

Figure 2: Fuzzy object tracking

2.3 Vision Function

On the other hand, if we consider the trends in modern communication techniques, video teleconferencing over the Internet provides an arguably more realistic interface into a remote space; but it is more of an enhancement to existing telephone communication technology rather than a new form of communication. With video conferencing, the participants find themselves fixed, staring almost all the time through the gaze of an immovable camera atop of someone's computer monitor. As actions and people pass across the camera's field of view, the observers are helpless to pan and track them or follow them into another room. In essence, here, it still lacks mobility and autonomy. The observers cannot control what they see or hear. Even if there are cameras in every room and the ability to switch between them is provided, the experience would still lack the spatial continuity of a walk around a building.

It will always be better to deliver a more realistic perception of physical embodiment of the user within the remote space being explored. Such system must immerse the user in the remote world by providing continuity of motion and control. These would provide the user the visual cues necessary to stitch together the entire visual experience into a coherent picture of a building and its occupants. It will be much advantageous and also be thoroughly appreciated if the user is provided with the necessary means of communication and interaction with the remote world and its real inhabitants using this new system. Furthermore, as Paulos et al. [2, 3] and Schulz et al. [4] have planned in their research, such a system can be further developed to give access to any user in the Internet with standard software running on currently existing computer architecture.

3 Localization and Target Tracking

Ubiquitous robots, finding its own path by means of visuals, sounds or any other method, have many problems due to the limitations of the sensory elements. Robots with binocular vision system have many advantages in this regard; but there are some limitations such as reduction in visible area [5] once the cameras are focused to obtain a closer view of an object as shown in Figure 3. This gets worse when there are moving objects in the surrounding area. There may be possible collisions due to blank angles and hence necessity of a continuous environmental awareness system arises.

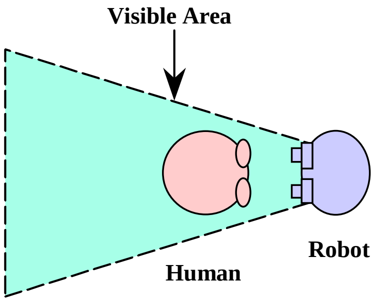

Figure 3: Limitation of visibility

There are platforms that estimate relevant quantities in the vicinity formed by combining information from multiple distributed sensors. For an example, robots in a team estimate their relative configuration by combining the angular measurements obtained from all of the ominidirectional images and performing triangulation operation as described by Spletzer et al. [6]. There are other variants such as the system proposed by Gurinaldo et al. [7] for controlling the perceptual process of two cooperative non holonomic mobile robots by formalism called perceptual anchoring. Their system enhances the awareness of the system by employing an anchor based active gaze control strategy to control the perceptual effort according to what is important at a given time. But such a system is of little use or not adequate for the robots whose main intention is to interact with humans. The situation gets critical, when such robots require to interact with a human while in motion. To place other robot units in the vicinity just to obtain an idea of the surrounding will be redundant and expensive. According to Splitzer et al. [6], questions of the quality and of how informative the gathered data are also arise, because they are obtained from individual robot units (sensors). In addition, there may be other issues like how the robot units (sensors) should be deployed in order to maximize the quality of the estimates returned by the team, because the robot platforms are mobile.

On the other hand, a different set of questions arises when one considers the problem of integrating information from a number of fixed distributed sensors such as cameras. Cost associated with transmitting and processing data, sensors that should be used to form an estimation for a given time, coordination among the sensors, automatically relating events among each other, sensor geometry, effects due to characteristic differences, etc. are some of the problems to be solved. In multiple video streams generated by multiple distributed cameras, finding correspondence is the key issue as observed in Lee et al. [8]. Hence a newer, simpler yet versatile localization and tracking system is required.

3.1 Networking

It should be possible to recruit such ubiquitous robots covering a large area. Hence to combine information gathered by one robot as well as to seek information / instruction, effective and convenient communication medium is required. In other words, these robots, while in motion, should be able to interconnect easily to a computer network as shown in Figure 4. Here the central unit with overhead camera monitors the environment continuously. Shapes of the various objects, positions at time to time, direction of motion, velocities, etc. are observed. These data is distributed among the ubiquitous robots in the surrounding area. In addition to the above, the central unit gathers the information send by these robots and sends to the respective recipients.

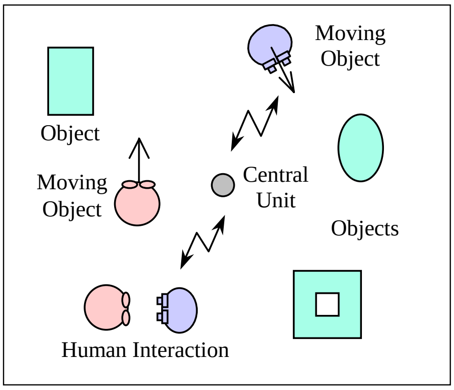

Figure 4: Interconnection of robots

3.2 Bluetooth Technology

In order to achieve the above mentioned communication, Bluetooth wireless technology, which facilitates an electronic equipment to make its own connections without wires, cables or any direct human intervention, will be very useful [9]. This type of a system will facilitate connection and synchronization; it will control the other electronic devices such as computers, printers, cell phones, PDA, televisions, alarm systems and telephone systems. These are the equipment developed by more than 1200 companies worldwide who represent the Bluetooth Special Interest Group including the giant leaders in the Electronic / Computer Industry like Ericson, IBM, Intel, Nokia, Toshiba, 3COM, Microsoft, Motorola, etc. [10]. All of these communications can take place in an ad hoc manner, without being aware and totally automatically. This gives the robots the freedom of appearing anywhere, any time and in any way very effectively.

4 Summary

It is difficult to design a human like robot in a ubiquitous environment that caters for interaction of any kind at the first instance. This research project tries to tackle such difficulties by manipulating some of the effective systems in a more realistic way. The approach is focused on the designing of a more human like robot while allowing services of many other systems at the same time. Here it tries to manipulate the services, capabilities and advantages of other existing systems as they are without any changes or with minimum alterations if required. But more complex interaction, will alternatively complicate the interface in return. This may make obtaining the so called 'Ubiquitous' more difficult. In the first instance, interaction may not provide the full facilities and functionalities such as two humans in conversation or in manipulating a task. But more general approach now will be available for a greater diversity of interactions to be managed in a later stage.

References

[1] T. Brown, *Stereoscopic phenomena of light and sight: A Guide to the Practice of Stereoscopic Photography and its Relations to Binocular Vision*, Reel Three-D Enterprises Publications, 1994.

[2] E. Paulos and J. Canny, "Space browsers: A tool for ubiquitous tele-embodiment," in *Proc. of International Conference on Computer Graphics and Interactive Techniques (SIGGRAPH)*, Digital Bayou 1996.

http://www.siggraph.org/conferences/siggraph96/core/conference/bayou/26.html

[3] E. Paulos and J. Canny, "Delivering real reality to the world wide web via telerobotics," in *Proc. of IEEE International Conference on Robotics and Automation*, vol. 2, 22–28 April 1996, pp. 1694 – 1699, 1996.

[4] D. Schulz, W. Burgard, A. B. Cremers, D. Fox, and S. Thrun, "Web interfaces for mobile robots in public places," *IEEE Robotics and Automation Magazine*, vol. 7, no. 1, pp. 48–56, March 2000.

[5] D. J. Coomb, "Real time gaze holding in binocular robot vision," *Technical Report, University of Rochester*, Dept. of Computer Science, 1992.

[6] J. Spletzer and C. J. Taylor, "Sensor planning and control in a dynamic environment," in *Proc. of IEEE International Conference on Robotics and Automation (ICRA '02)*, vol. 1, pp. 676–681, May 2002.

[7] S. A. Gurinaldo, K. Watanabe, and K. Izumi, "Enhancing awareness in cooperative robots through perceptual anchoring," in *Proc. of the 9th International Symposium on Artificial Life and Robotics (AROB)*, vol. 2, pp. 523–526, January 2004.

[8] L. Lee, R. Romano, and G. Stein, "Monitoring activities from multiple video streams: Establishing a common coordinate frame," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 758–767, Aug. 2000.

[9] M. Miller, *Discovering Bluetooth*, Sybex Inc. Publications, 2001.

[10] B. R. Miller and C. Bisdikian, *Bluetooth Revealed: The Insider's Guide to an Open Specification for Global Wireless Communication*, 2nd Edition, Prentice Hall, 2001.

[11] C. Stauffer and W. E. L. Grimson, "Learning patterns of activity using real-time tracking," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 747–757, Aug. 2000.

[12] T. B. Sheridan, *Telerobotics, Automation and Human Supervisory Control*, MIT Press, 1992.

[13] H. Monson, "Bluetooth Technology and Implications," *System Optimization Information Article*, December 14, 1999.

(http://www.sysopt.com/articles/bluetooth/)

[14] http://www.bluetooth.com/news/dex.asp?A=2&PID=1192

[15] http://www.open.ac.uk/oubs-future/Millennium/ 636.htm