Design and Implementation of Human-Robot Interactive Demonstration System Based on Kinect

Liying Cheng[1, 2, 3], Qi Sun[1], Han Su[1], Yang Cong[3], Shuying Zhao[1, 3]

1. College of Information Science and Engineering, Northeastern University, Shenyang Liaoning, 110819, China

2. Shenyang Normal University, Shenyang 110034, China

3. State Key Laboratory of Robotics, Shenyang Institute of Automation, Chinese Academy of Science, Shenyang 110016, China

E-mail: clypzb@163.com

Abstract: With the development of technology, humanoid robots gradually enter our life, not only for education, but helping people with housework and many other tasks, some of which are inevitable for human. But only few people know how to control and interact with a humanoid robot, which hinders the development of humanoid robot. So a human-robot interactive demonstration system is designed to help non-expert users to control the humanoid robot, Aldebaran humanoid robot Nao in this case. Users just need to use the most natural body gestures to interact with the robot. Microsoft Kinect is applied in this system to recognize different body gestures and generate visual Human-Robot interaction interface, then the controlling signals of different body gesture modules are sent to Nao through wifi, which can stimulate Nao to complete tasks. This kind of system aims to enrich the interactive way between human and robots and help non-expert users to control the robot freely, making Human-Robot interaction much easier.

Keyword: Humanoid robot, gesture recognition, Kinect, HRI

1 INTRODUCTION

Currently, the robot system has been developed to a very advanced level, and robots have moved away from industrial settings and closer into our lives, the question arises of how to interact with these robots [1][2]. Up to now, the interaction between human beings and robots is limited by keyboards and mice [3]. Some of the advanced robots can identify pictures and voice commands, but few can understand human body gestures. That limits robots to learn about the environment and interact with people [4]. Therefore, increasing amount of attention has been paid to the study of interaction between human and robots. With the steady development of interactive technology, a method of interaction by recognizing human body gestures has been presented. Through this method, robots can recognize different body gestures, and simplify the process of interaction. Thus, we can control them easily, precisely and naturally. This Human-Robot Interaction allows amateur users or those with disabilities to control robotic systems [3][4], and help them accomplish missions which seems impossible to them.

Most previous researches, for example, interacting with robots through putting a large amount of motion sensors on bodies, seem limited in the field of Human-Robot interactive system. This way not only needs numerous sensors connected to the computer, but also influences the effect of Human-Robot interaction. After that, interaction between humans and robots through voice commands seems a little more natural, but resulting from the differences between different people's intonation, it requests an extremely accurate voice recognition system to separate the voice command from noisy circumstance. In a word, this interactive way is limited, too.

For above-mentioned cases, a new Human-Robot interactive demonstration system is designed and implemented. Microsoft Kinect is used in this system as the body gesture recognizer. After collecting and identifying different body gestures, it can provide the

This work is supported by the Fundamental Research Funds for the Central Universities under Grant N110804005

visual interpretation of gestures and send it to robots, hence achieving the interaction between human and robots. This system can make robots able to accomplish different actions and enrich the way of Human-Robot interaction.

This paper is organized as follows: Section 2 describes the structure and the operational principles of this demonstration system. In Section 3 the key technology is discussed, including the recognition of body gestures, the communication between Nao and Kinect as well as the behavior management of Nao. Experiment results are given in Section 4. Section 5 gives the conclusion and the plan of our future work.

2 SYSTEM DESIGN

2.1 System structure

There is only one humanoid robot Nao and one Microsoft Kinect in this system, without any other complex connection devices, wires or sensors. Humanoid robot Nao is designed and developed by Aldebaran, France. It has 25 degrees of freedom, several kinds of sensors and smooth actions [5]. With these advantages Nao has been regarded as the leader of humanoid robot system and chosen as the international standard platform of humanoid football match. Kinect is developed by Microsoft, which contains a low cost RGB camera, a standard CMOS camera and a depth camera which can collect the depth information [6]. The structure of this demonstration system is shown in Figure1.

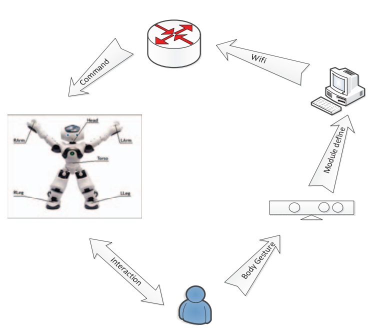

Fig 1. System Structure

2.2 Principles of system operation

In this interactive demonstration system, humanoid robot Nao is controlled by body gestures. Kinect is used as the vision part to collect the controller's body gestures. By means of processing the collected images, the body gesture modules are defined to be recognized by computer. When controllers interact with the robot using their bodies, the body gestures will match with the defined action modules, and send the matched signal to the LAN through wifi. As humanoid robot Nao is equipped with a wireless transmission module, it can receive the wireless signal. This interactive demonstration system is realized as the following steps:

(1) Collect body gestures in 3D space by Kinect.

(2) Process the images from Kinect and define the gesture module.

(3) Transmit instructions and data.

(4) Execute the behaviors of humanoid robot Nao.

3 KEY TECHNOLOGY

3.1 Recognition of body gestures

In this interactive demonstration system, Kinect is used to recognize body gestures and generate the skeleton model of the controller. The recognized skeleton figure is shown in Figure2.

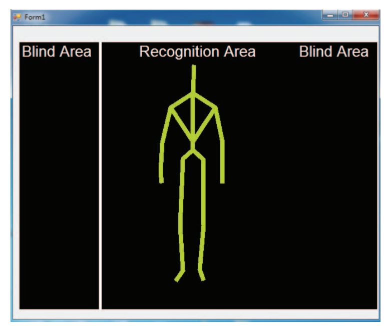

Fig 2. Recognized Skeleton Figure

The Microsoft Kinect has a horizontal range of 57 degrees, a vertical range of 43 degrees and a depth range of 1.2-3.5 meters given officially [6].

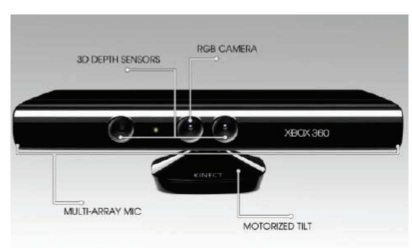

Fig 3. Kinect Sensor

Kinect is able to capture the surroundings in 3D by combining the information collected from the depth sensors and the standard RGB camera [7]. The result of this combination is an RGBD image with 640x480 resolutions, where each pixel is assigned certain color information as well as depth information. In ideal conditions the resolution of the depth information can be 3 mm in height, using 11 bits resolution. Kinect works with the frequency of 30 Hz for both RGB and depth camera. On the left side of the Kinect is a laser infrared light source which generates electromagnetic waves with the wavelength of 830 nm. Information is encoded in light patterns which are deformed as the light reflecting from objects in front of the Kinect [8].

In this system, the ideal interactive distance is given as 2 meters after a number of experiments. Kinect can not only offer natural interaction interface, generate the 3D model of objects, but also track the skeleton in 3D space. At the same time it can also separate the specific character from complex circumstance. We can get the x-y-z coordinates of the interested pointed through Kinect which are used to defined gesture modules.

Skeleton tracing is the core technology of Kinect and because of this the Human-Robot interactive demonstration system can be achieved. Kinect can track 20 skeleton points of human body at most [6], and these specific points are shown in Figure4.

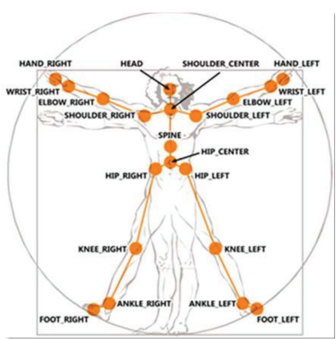

Fig 4. Distribution of Skeleton Points

In this system, coordinates of some key points are chosen to define the interactive gesture modules, such as head, hands, knees, spine, shoulders, elbows and hip. Here the dot product of vectors is used to calculate the angle of each joint. Because in some cases, three points describing the joint cannot create a triangle in a single plane. The solution to this problem is the usage of the dot product of vectors. For instance, for the elbow roll, the vector pointing from the shoulder to the hip and the vector pointing from the elbow to the hand are used. Once the dot product of the two vectors is calculated, the angle can easily be determined [9] (where θ is the angle between the two vectors):

$$\vec{a} \cdot \vec{b} = \|\vec{a}\| \|\vec{b}\| \cos\theta \tag{1}$$

Different body gestures modules are shown in Figure5.

Fig 5. Different Body Gestures

In the early phase of the designing of this system, there are two cases making the users' skeleton image distorted, which are shown in Figure6.

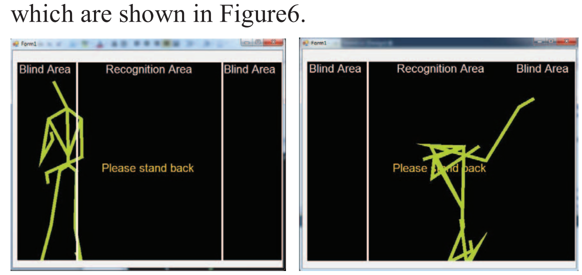

Fig 6. Two Case of Distorted Image

One occurs when user stands close to the border of horizontal range, the other occurs when the distance between user and Kinect is less than 0.8 meters. These two cases cause the wrong recognition and make robots execute wrong behaviors. To solve this problem, Kinect is fixed within a certain range. Because the resolution of Kinect is 640*480 and it convert this resolution to the range of -1 to +1, the horizontal view is fixed from -0.75 to +0.75, and depth distance from 1.5 to 2.5 meters in order to eliminate the distortion. It proves efficient in avoiding this kind of wrong control.

3.2 Communication between Nao and Kinect

As Nao is programmed under the architecture of .Net 3.5, and Kinect is .Net 4.0, this brings along a big challenge. Because the wifi module of Nao and this whole system is implemented in a local area, the UDP communication protocol is used, which proves to be stable in a local area to transmit the signal of different gesture modules. The whole process of communication is shown in Figure7.

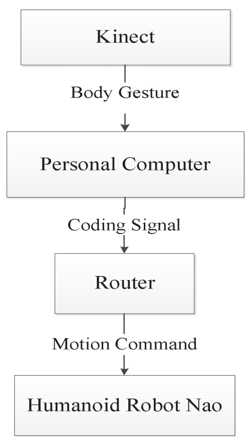

Fig 7. Flow Chart of Communication

By this way, the Human-Robot interactive demonstration system containing humanoid robot Nao, Kinect and personal computer is successfully built, and every signal generated by Kinect is transmitted to robot Nao lossless.

3.3 Behavior management of Nao

Since the users of this system are amateurs, the behaviors stimulated by different gesture modules have been loaded into the behavior management of Nao. Hence the process of controlling can be totally simplified. Nao's resource can only be occupied by one behavior. If two behaviors are designed to use the same resource at one time, there will be a source conflict. Based on this case, a flag is inserted into Nao's memory to avoid such conflict. When one behavior is running, the flag will be set. Before the running of any behavior, the program will check the flag to see whether there is another behavior running. By this way, the source conflict is avoided. The controlling process is shown in Figure8.

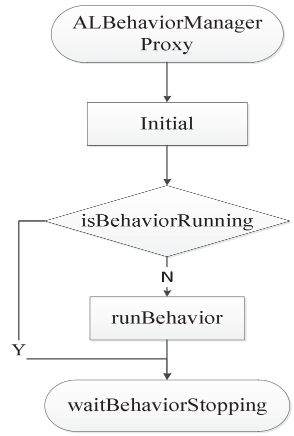

Fig 8. Flow Chart of Behaviors Management

4 EXPERIMENT RESULT

Some behaviors are loaded into Nao's behavior manager and linked to related gesture modules. The correspondence is shown in Table1.

Table1. Table of Gestures-Behaviors

Body gesture	Robot's behaviors
Raise right arm forward	Stand up
Raise right arm towards the right	Introduce itself
Raise left arm forward	Play Taiji
Raise left arm towards the left	Dance
Wave left hand	Wave hand
Squat	Sit down
Raise right knee, then left knee	Walk

In the process of experiments, Nao needs a few seconds to execute the behaviors after users send the signal of body gestures. If another signal is sent in the same period of time, it will cause the collision and make Nao shut down. Based on this case, the program is set to sleep a short period after one signal is sent. It proves efficient in solving this problem.

5 CONCLUSION AND FUTURE WORK

In this Human-Robot interactive demonstration system, the recognition and identification of body gestures as well as the interaction with humanoid robot Nao have been achieved, which makes Human-Robot interaction more natural. This system can pave the way of allowing robots work for humans. Meanwhile, as a new way of interaction, it can be applied in technology museums. This system makes a breakthrough for Human-Robot interaction. Our future work is to add the function of recognizing dynamic body gestures to this demonstration system.

REFERENCES

[1] Michael Van den Bergh, Daniel Carton, Roderick De Nijs, Nikos Mitsou, Christian Landsiedel, Kolja Kuehnlenz, Dirk Wollherr, Luc Van Gool, Martin Buss, Real-time 3D Hand Gesture Interaction with a Robot for Understanding Directions from Humans, IEEE RO-MAN, 357-362, 2011.

[2] Michal Tolgyessy, Peter Hubinsk, The Kinect Sensor in Robotics Education.

[3] Hee-Deok Yang, A-Yeon Park, Seong-Whan Lee, Human-Robot Interaction by Whole Body Gesture Spotting and Recognition, 18th International Conference on Pattern Recognition, Vol.4, 774-777, 2006.

[4] Radu Muresan, Stefano Gregori, Natural User Interface for Robot Armature Control, February 7, 2011

[5] Aldebaran Robotics, Nao Academics Data Sheet, Available:http://www.aldebaran-robotics.com/en/node/1166, France, 2011, [Aug, 20, 2011].

[6] Microsoft Corp, Kinect for Xbox 360.

[7] Michal Tolgyessy, Peter Hubinsk, The Kinect Sensor in Robotics Education.

[8] The Bilibot website, Available: http://ww.bilibot.com.

[9] Faculty of Science, Universiteit van Amsterdam, Dutch Nao Team Technical Report 2011, Netherland, September 28,2011