Journal of Basic and Applied Engineering Research
p-ISSN: 2350-0077; e-ISSN: 2350-0255; Volume 3, Issue 10; July-September, 2016, pp. 823-827

http://www.krishisanskriti.org/Publication.html

A Speaking 3R Robot

Aakansha Bansal[1] and Manoj Soni[2]

[1,2]IGDTUW
E-mail: [1]aakanshabansal10@gmail.com, [2]manoj_soni2002@yahoo.com

Abstract—*In this research work, a speaking 3R robot with voice customization capability has been developed which converts position of 3R robotic arm into speech. Speech produced can be converted into any target speech voice. This approach consists of both hardware module and software application. In hardware module - The position of the 3R robotic arm is determined with the help of accelerometer and atmega 328 microcontroller. The design of 3R robot and the concept of decoding its position by considering axis orientation is discussed .In software part - an arduino program is written that converts position of 3R robotic arm into speech. Further ,speech can be converted to any target speech. The entire process of speech synthesis is tested and test results displaying robotic arm position have been shown.*

Keywords- *accelerometer, arduino application, axis orientation, microcontroller ,voice conversion, LP analysis*

1. INTRODUCTION

One of the most important criteria for human robot interaction is need for effective communication between them. Talking robots may be used as shopping assistants ,tutors, domestic chores helper, elderly people companion, complex environment navigator, physical therapist and what not. It can also be used in application aiding physically challenged community like blind and dumb people [1]. This research will help to empower the people who cannot speak or see and enable them to work in manufacturing industries. We have developed a 3 R robotic arm with the end effector attached to the fixed surface. It can be used for welding, painting, assembly, pick and place with accuracy and precision. Accuracy can vary with speed and position within the working envelope and with payload . A typical robot can, of course make a positional error exceeding that and that could be a problem for the process. . The main challenge in human robot interaction is haring of workplace so that robot do not cause any harm to the human beings. Thus to avoid the problems caused by inaccuracy in robotic arm position, a communication system is developed which converts its position into speech and alerts the workers if error occurs in robotic arm position. Position recognition of robotic arm can be done by two major techniques namely vision based approach and haptic based approach[2]-[4]. Because of large data processing , low speed response in vision based approach , the haptic approach is considered in this paper[4].For position recognition of robotic arm ,several recognition algorithm like neural network approach [7]-[9], hidden markov model(HMM)[10]-[13],and feature extraction analysis are used. All these techniques and vision based approaches require complex data processing. So for portability of the system decreases and cannot be carried everywhere. In this thesis, this is overcome by using arduino based program and a haptic system. The positions are recognized by implementation of simple template matching [14]-[15] techniques considering the maximum and minimum voltage values according to accelerometer towards gravity. The major factor in this system is enhancement of portability and the conversion of ADC values back to voltage for reducing the range and for improved coding and decoding of positions. Also, a voice conversion module[5,6,20] is interfaced with the arduino board to convert given speech into any target speech. This speaking robot can be used as a communication aid for the dumb people who are unable to speak by using a switch array in which different user defined commands are programmed like emergency stop, fire alert, system failure etc. Whenever a dumb person presses the corresponding switch ,same information will be delivered out as voice using voice conversion module.

2. SYSTEM DESIGN AND DESCRIPTION

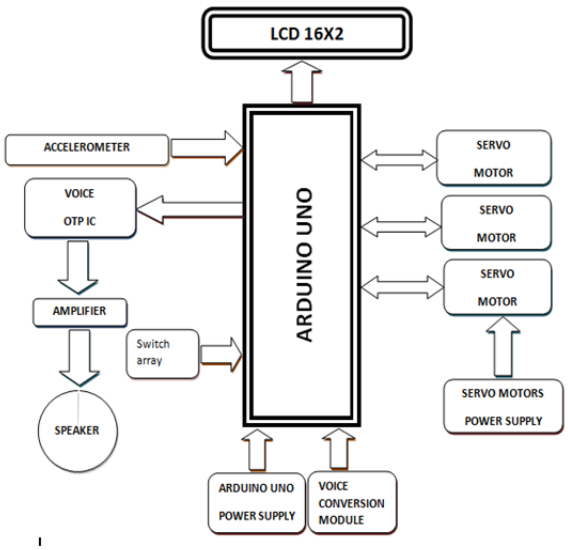


Fig. 1.

2.1. Hardware Module

The overview of the system is shown in Fig.1. The main purpose of this model is to recognize position of end effecter of 3R robot. This is done using 3 axis accelerometer(ADXL 335)[16] and microcontroller. The 3 servo motors of robotic arm are interfaced to the microcontroller i.e. ATMEGA 2560 .The accelerometer is attached to end effector of the robotic arm and its position is input to the microcontroller. Arduino uno converts this position into speech using voice integrated chip , amplifier and speaker. Power supply is developed for servomotors so that each motor gets 5v for proper functioning. Also , the power supply is developed for arduino board so that it can carry out the required operation.

Hardware Details:

- Motors:

In this setup 4 motors are used out of which three are Tower pro servos of torque rating 10kg-cm for the joint control and one dc motor for the end effector. Servo Motor used are MG995 High Speed Metal Gear Dual Ball Bearing Servos.

- Torque Calculations

Free body diagram of the setup is shown in following figure.

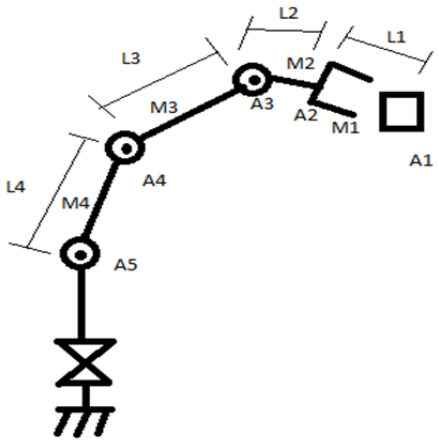


Fig. 3: Free Body Diagram Of Experimental Setup

The torque required is given by the equation:

$$T = F * L$$

Where F= Force on the link and L= Length of the link

Now F= Mass* Acceleration due to gravity= weight W of the link

Therefore, $T = W * L$

Here

M_1, M_2, M_3, M_4 = weights of the links

A_2, A_3, A_4 = Weights of the servo motors

A_1= Weight of the object

L_1, L_2, L_3, L_4= Respective link lengths

T_1, T_2, T_3, T_4 =Respective torques required. The torque equations for the setup are given

$$\boldsymbol{T_1} = L_1 * A_1 + \left(\frac{L_1}{2}\right) * M_1$$

$$\boldsymbol{T_2} = (L_1 + L_2) * A_1 + \left(\frac{L_1}{2} + L_2\right) * M_1 + L_2 * A_2 + \frac{L_2}{2} * M_2$$

$$\boldsymbol{T_3} = (L_1 + L_2 + L_3) * A_1 + (L_2 + L_3) * A_2 + L_3 * A_3 + \left(\frac{L_1}{2} + L_2 + L_3\right) * M_1 + \left(\frac{L_2}{2} + L_3\right) * M_2 + \frac{L_3}{2} * M_3$$

$$\boldsymbol{T_4} = (L_1 + L_2 + L_3 + L_4) * A_1 + (L_2 + L_3 + L_4) * A_2 + (L_4 + L_3) * A_3 + \left(\frac{L_1}{2} + L_2 + L_3 + L_4\right) * M_1 + \left(\frac{L_2}{2} + L_3 + L_4\right) * M_2 + \left(\frac{L_3}{2} + L_4\right) * M_3 + \frac{L_4}{2} * M_4$$

Assuming the weight of links negligible as compared to the motor weight, i.e

$M_1 = M_2 = M_3 = M_4 = 0$

A_1= 0.5kg

$A_2 = A_3 = A_4$ =55 gms = 0.055 Kg

L_1=6cm, L_2= 2cm, L_3=5cm, L_4=5cm.

Solving the above equations we get

T_1= 3.03 kg-cm

T_2 = 4.16 Kg-cm

T_3 = 7.26kg-cm

T_4 = 9.96kg-cm

Due to the availability constraints keeping in mind the torque requirements all servo motors of rating 10 kg-cm have been selected.

- **Power supply**:

A 12 V adaptor is used to provide power supply to the board externally as the 5V supply provided from the computer is not enough to supply power to all the motor**s.**

- **ATMEGA 328:**

utput of accelerometer is fed to the ATMEGA 328(operating voltage 2.7-5.5v) on port A that is used as analog input for built in ADC. By default ADC work on half the crystal frequency (7.372 MHZ).It has 10 bit built in ADC but we have used its 3 bits only which convert it into digital form.3 bit code for particular phoneme is sent through USART(universal synchronous and asynchronous receiver and transmitter) at baud rate i.e. 9600 to speak jet.

- **Speakjet**

The output of ATMEGA which sends serial data at baud rate of 9600 to speakjet that has mathematical sound architecture(MSA) predefined allophones that are spoken depending upon the data received from the microcontroller. It generates the voice audible to human beings[18] .

- **Signal Amplification**

The synthesized voice output from speakjet is not clearly audible to user, therefore it is fed to the amplifier (LM386) that increases its volume .It provides a gain of 200 that makes it quite natural to human ears. An 8 ohm speaker is then used to get final output .It has wide supply voltage range of 4v-12v ,low quiescent current drain of 4 mA, voltage gains from 20 to200 and ground referenced input

- **Electrical specifications of this project are:**

Power Consumption : 250 mV

Circuit Current : 67.5ma

Operating Voltage: 3.7V-12V

2.2. Software Module:

Arduino is an open-source prototyping platform based on easy-to-use hardware and software. Arduino boards can read inputs and convert them into desired output. In our program arduino board reads analog signal from accelerometer and convert it into speech output. Also ,it can convert input from switch button pressed in switch array into desired speech output. Signals are send to the microcontroller using a step of instructions written using arduino programming language and IDE. Arduino is an open-source prototyping platform based on easy-to-use hardware and software. As voice conversion involves complex equation solving is required , the arduino board is interfaced with MATLAB using hardware support package for arduino.

3. POSITION RECOGNITION

3.1. Analysis of position of 3 R robotic arm

Different positions of the robotic arm is found by analyzing change in position and orientation of robotic hand. An accelerometer is used to detect the position of the robotic arm in terms of its X,Y &Z coordinates. We have considered four position in our working model.

- Position 1: Home position
- Position 2: Position at which robotic arm picks the object
- Position 3: An intermediate position
- Position 4: Position at which robotic arm drops the object

Our program converts these position into speech.

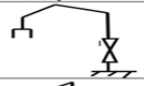
3.2. Position Recognition Technique

For different position ,the accelerometer readings vary with respect to gravity in each plane and produces maximum value in any one of the planes. Considering this, all the position of the robotic arm can be recognized. The value is either maximum or minimum in any one of the x, y or z planes for each position ,which is used for identifying the position. These set of values will be unique for each positions and each position can be differentiated easily.

3.3. Formation of sequence table

As observed from the Fig.2,orietation analysis for position 1,2,3,&4 depends upon accelerometer readings which varies with gravity.

Table 1.

S.no.	Figure	Position	X-axis	Y-axis	Z-axis
1.		Position 1 (Home position)	245-250	355-360	350-355
2.		Position 2 (Pick up position)	255-260	430-435	320-325
3.		Position 3 (Intermediate position)	365-370	370-375	275-280
4.		Position 4 (Drop down position)	260-265	380-385	310-315

3.4. Sequence Matching and Coding Positions

The ADC values obtained for different orientations are first collected and tabulated. The ADC values vary randomly at various heights for same position. Hence, it is difficult to code the positions with these values. Fixation of range is also laborious job and prone to errors as the range is wide for 10 bit ADC. Moreover determining range for all the position is difficult. Thus, range is obtained using the following formula

$V_{dc} = V_{adc}$ * Range Of Supply Voltage / $2^{bits\ of\ adc}$

Hence coding of positions is made easy by fixing the ranges either to a maximum value or to a minimum value. For e.g. the range(P_{1y} >650 && P_{1y} <950) is coded as P_{1y} <1.5. The maximum value will be above 2 in the accelerometer when axis is oriented towards gravity and minimum values is below 1.5 for axis oriented away from gravity.

3.5. Decoding Of Positions

When the controller gets the input sequence ,it makes computation and compares them with coded templates of values. When the condition gets satisfied it sends the position name to voice integrated chip which further sends it to speaker via amplifier.

4. VOICE CONVERSION MODULE

Voice conversion is the process by which source voice is converted to any target voice using speech parameters. Prior to analysis speech signal is passed through pre emphasis filter which reduces the dynamic range of speech spectra. Then the pre-emphasized signal is segmented in frames for short term analysis using hamming window. Here, frame size of 30 ms is used because human pitch cannot go below 50 Hz i.e. 20mS .After pre- emphasis and framing , next phase is LP analysis which allows us to predict future samples based on few previous samples. The no. of samples required to predict the next sample is called prediction order of the filter p. After analysis phase LP coefficients of both source and target speech have been extracted. These parameters are used to model the vocal tract filter of target speech and an inverse filter to extract excitation component of source speech which is applied to all pole vocal tract filter which shapes the spectrum of source excitation. Modification in excitation component refers to pitch modification which is done by speech synthesizer approach. . For voiced component pitch value is determined and pulse train of determined pitch period is generated, whereas for unvoiced components white Gaussian noise is generated. The amplitude of the generated excitation signal is scaled by gain value and then passed through a filter characterized by LP coefficients of target speech and hence the target speech is generated.

5. AID FOR DUMB

Dumb (muteness) is a situation in which a person is unable to speak .There is a need in our society to empower physically challenged people so that they can also work in different fields like manufacturing industry , power plants, hospitals, house hold purpose, schools, colleges, offices, public gatherings etc. In this prototype system , a communication aid for the dumb people is also developed which is found to serve the purpose effectively. A switch array is used in which messages are entered such that if the dumb person wants to communicate with normal people he/she can press the specified switches in the switch array. The same information will be delivered out as a voice signals by using a text to speech converter module which is again given to a amplifier unit to get the exact information without any noise, and from there to a speaker. This module describes the design of a simple, low-cost controller based Text-to-Speech conversion system. We can get both the voice announcements as well as a 16X2 LCD display is used to display the text messages on it.

6. CONCLUSION

A portable 3R speaking robot is developed which converts position of robotic arm into speech using the accelerometer, microcontroller, arduino uno , voice integrated chip, amplifier and speaker. Such a system will enable the blind people to come to know which operation is being performed. Speaker will speak the position only when the robotic arm is at the desired position. If the actual position of the robot will deviate from the commanded position ,speaker will not generate output and alert the factory worker that positional error has occurred. Whenever positional inaccuracy occur in robotic arm due high speed or payload, even dumb people can alert the others .We have interfaced a switch array in the above system and messages such as 'system failure' can be entered into the system such that if the dumb person wants to communicate with normal people he/she can press the specified switches in the switch array. The same information will be delivered out as a voice signals by using a text to speech converter module which is again given to a amplifier unit to get the exact information without any noise, and from there to a speaker. Source voice can also be converted to any target voice.

REFERENCES

[1] S. F. Ahmed, et al., "Electronic speaking glove for speechless patients," in the IEEE Conference on Sustainable Utilization and Development in Engineering and Technology, Petaling Jaya, Malaysia, 2010, pp. 56-60..

[2] L.Bretznar & T.Linderberg, “Relative orientation from extended sequence of sparse point and line correspondences using the affine trifocal sensor,” in Proc. 5th Eur. Conf. Computer Vision, Berlin, Germany, June 1998, Vol.1406, Lecture Notes in Computer Science, pp. 141- 157, Springer Verlag.

[3] D. Xu, “A neural network approach for hand gesture recognition in virtual reality driving training systern of SPG,” presented at the 18th Inc .Conf. Pattern recognition, 2006.

[4] M.Delliraj, et al., "Design Of Smart E-Tongue For Physically Challenged People" in International Conference on recent trends in information technology,2013

[5]Wei Chao Xie ,et al .,"research on voice conversion based codebook and GMM "12th IEEE (ICCT),2010,1403-1406

[6] Yannis Stylianou ," Continuous Probabilistic Transform for Voice Conversion” in IEEE TRANSACTIONS ON SPEECH AND AUDIO PROCESSING, VOL. 6, NO. 2, MARCH 1998,131-142

[7] Russell, Stuart, and Peter Norvig. Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs, NJ, 1995.

[8] Krose, Ben J. A., and P. Patrick van der Smagt. An Introduction to Neural Networks University of Amsterdam, 1995

[9] Auer, A. Pinz, and M. Gervautz. Tracking in a Multi- User Augmented Reality System. In Proceedings of the IASTED International Conference on Computer Graphics and Imaging,249-253, 1998
[10] Rabiner, L. R. “A Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition.” Proceedings of the IEEE 77(2):267-296, 1989
[11] Huang, X. D., Y. Ariki, and M. A. Jack.,” Hidden Markov Models for Speech Recognition,” Edinburgh University Press, Edinburgh, 1990.
[12] Rabiner, L. R., and B.H. Juang,” An Introduction to Hidden Markov Models,” IEEE ASSP Magazine, 4-16, January, 1986.
[13] Charniak, Eugene “Statistical Language Learning,” MIT Press, Cambridge, 1993.
[14] Sturman, David J. “Whole-hand Input. Ph.D dissertation, Massachusetts,” Institute of Technology, 1992
[15] Rubine, Dean. “Specifying Gestures by Example,” In Proceedings of SIGGRAPH’91, ACM Press, 329-337, 1991.
[16] Ruize Xu, Shengli Zhou, Wen J. Li, “MEMS Accelerometer based Non-Specific User Hand Gesture Recognition”, in IEEE Sensors Journal, Vol. 12, no. 5, pp.1166-1173.
[17] Watson, Richard. “A Survey of Gesture Recognition Techniques,” Technical Report TCD-CS-93-11, Department of Computer Science, Trinity College Dublin,1993
[18] Syed Faiz Ahmed, Syed Muhammad Baber Ali,Sh. Qureshi," Electronic speaking love for speechless patient" in IEEE conference 2010
[19]Yan -You-Chen ,"voice customizable text to speech for home care system" ,IEEE 2013
[20] Yannis Stylianou ," Continuous Probabilistic Transform for Voice Conversion” in IEEE Transactions On Speech And Audio Processing, Vol. 6, No. 2, March 1998,131-142