

Design and Fabrication of Speech Controlled Robot Scribe

¹Dr PHV Sesha Talpa Sai, ²Dr. Amiya Bhaumik, ³Abhishek Tripathi, ⁴Arun Prasad K V, ⁵Pradeep Shivappa
Asundi
^{1,2} Professor, ^{3,4,5} UG Student
¹Malla Reddy College of Engineering & Technology, (Autonomous)
^{2,3,4,5} Lincoln University College, Malaysia

Article Info

Volume 83

Page Number: 16631 – 16635

Publication Issue:

May - June 2020

Article History

Article Received: 1May 2020

Revised: 11 May 2020

Accepted: 20 May 2020

Publication: 24May 2020

Abstract

One of the most challenging jobs in robot control is handwriting, is of particular interest due to the fine control needed to write alphabets to express words and sentences. Hence, the proposed work is aimed to accomplish accurate robot handwriting of speech in English using a robotic arm system that accurately writes the words spoken by the user. In this approach, every word spoken by the user is converted into the form of text by using Google cloud speech Application Program Interface by using powerful Artificial Neural Network models. Google cloud speech Application Program Interface (API) can convert speech to text in noisy environments and it is accent insensitive, hence making it the most viable option to convert the speech to text. Two dimensional inverse kinematics equations are used to generate the control angles to drive the motors of the 6 Degrees of Freedom robotic arm to replicate the words spoken by the user.

Index Terms; *Robotic arm, Text to Speech, Inverse Kinematics, DOF, API*

I. INTRODUCTION

Development of assistive robots is a very important research field but a robot needs accurate control and manipulation to assist people in their daily chores. Among the most challenging robot control tasks, the writing task is of particular interest as very fine control is needed to replicate alphabets and words in real-time. Hand writing is one of the most complex tasks that are of a very great importance because of the applications it has in real world [1].

1.1 The NAO System

The Nao robot presented in this work replicates the alphabet or shape that is drawn on the tablet that is connected to the nao on a paper using a marker. The robot uses inverse kinematics to replicate alphabets real time, but it does not use speech processing. This robot was mainly engaged in teaching school students from ages 6-8 such that they learn from the robotic arm, how to write. [2]

The problems associated with optimization of cursive letter dynamics is solved by using a velocity linearized arm dynamics model. A custom robotic arm is fabricated and tested and results were considered as further feedback of the arm system. This work presents a KUKA robotic arm that could write Chinese characters in terms of multiple strokes. Visual feedback is provided by a camera node. A data base of multiple alphabets is created and a learning node is trained on this system to improvise the alphabets after every iteration. Even though the system presented is very accurate and precise, the robotic arm by KUKA is very expensive, which prevents many people with limited financial resources [3].

The robot demonstrates replication of signatures made by the user on a tablet by using a linear 3D robot. The pen used is considered to be a combination of a spring and a dashpot. Even though this system is very accurate, it uses a Cartesian coordinate control, which is not the case in human

beings. In this paper presents motion analysis of a human like 6DOF Robotic arm for hand writing task [4]. Modeling separates the required movement into smooth global and fast local motion. It uses distributed positioning and Inverse Kinematics for positioning the pen tip. In comparison with the systems that are presented in the introduction, a robotisdynamixel system is relatively cheaper (200\$ as of April 2019). The system exhibits very precise output with an accurate feedback. [5]

II. HARDWARE DETAILS

Generally HX servos are used in do it robotics and low-cost projects, but there are many backlogs of them such as absence of any feedback, and they require individual wiring and different power connections. A better option would be to use a smart robot actuator such as Dynamixel AX-12A by Robotis, represented in Figure.4. Each dynamixel has an 8 bit microcontroller in itself which holds a unique ID for each of the servo. The motors can be connected in a daisy chain. The speed, torque, angle of rotation, mode of operation and voltage limits. [6] Even though they offer many features and high precision, they are not very expensive. Few drawbacks include the usage of contact encoders, which may wear out through the course of time, comparatively low resolution. This results in lack of extreme precision and causes disparities between the planned path and the path followed by the arm. The proposed speech controlled Robot scribe model is depicted in the Figure 1. Using clamp 1 and clamp 2 as represented in Figure 2, Figure 3 respectively and Dynamixel AX-12A given in Figure 4, the robotic arm is assembled as shown in Figure 5. A better alternative is a smart servo such as the Dynamixel AX-12A by Robotis, represented in Figure.4 [7]

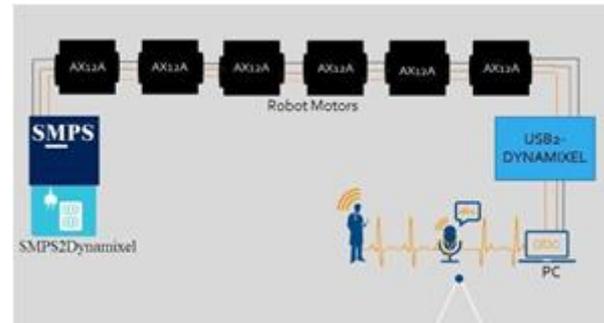


Figure 1: Block Diagram

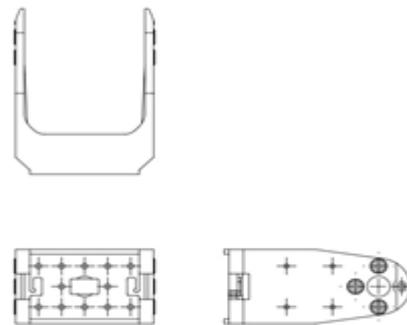


Figure 2: Clamp1

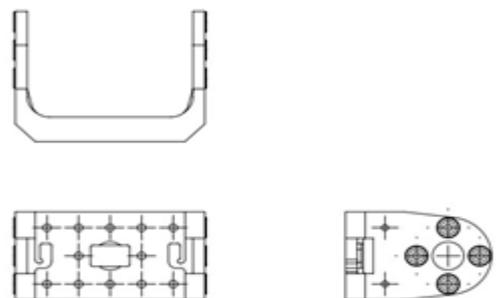


Figure 3: Clamp2



Figure 4: Dynamixel AX-12A [Source: www.robotis.com]



Figure 5: Assembled hardware Robotic arm by proposed method



Figure 6: Laptop mic used for speech to text

III. SOFTWARE ASPECTS

3.1 Speech to text

For the robot writing to be accurate, an accurate speech to text conversion system is needed. Even though a system of that sort can be created, it might not work with the users with multiple accents. The system is supposed to be trained for various accents and that would involve a very high amount of effort. So a readymade API that can convert speech to text can be used to reduce the inaccuracy of the speech to text system. There are many APIs that use neural networks and machine learning to convert text to speech, Google Cloud speech API works the best due to the extremely large training set that is created for machine learning by users from different countries with different accents. Google cloud speech API becomes the best option when it comes to conversion of text to speech because:

- It uses a very advanced deep learning based Artificial Neural Network models to recognize speech with a very high accuracy.
- It can return partial results without waiting for the user to complete speech and after the speech ends, the sentence will be tailored based on the context.
- Digital noise reduction is used to eliminate the ambient noise and speech. Figure 6 represents the speech to text setup used in the system.

3.2 Motion Generation

The human hand has 19 DOF; but recreating those in a robot is very hard. The writing task can be achieved by using a 2 DOF Robotic Arm. The manipulation of the arm can be controlled by using standard inverse kinematics equations.

3.3 Acquiring coordinates of alphabets

To acquire the joint space coordinates from the Cartesian space, inverse kinematics is used. In order to create Cartesian space values for various alphabets, equal sized images of the different characters of the alphabet are considered and processed. A threshold is applied on the image that is being considered to convert it into a grey scale image. Each pixel of the image is scanned through and the pixels that are black, i.e. value 0 are considered to be the black pixels. Black pixels are where the robot is supposed to move to write that particular alphabet. The pixel coordinate of all the black pixels are written to a .dat file with the respective letter name for further use.

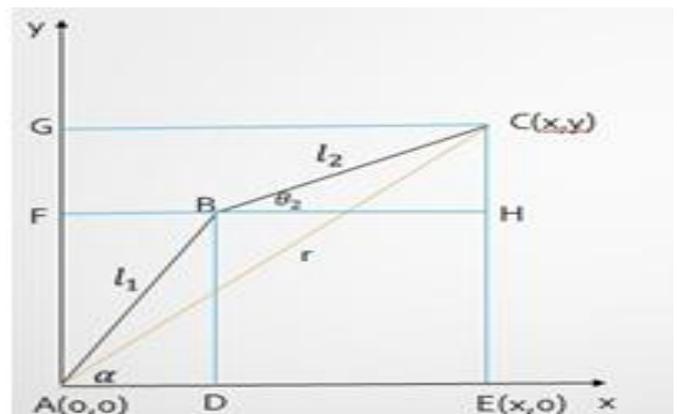


Figure 7: To find theta 1

3.4 Mapping

The coordinates file that was created holds the coordinates of the pixels, so a suitable mapping algorithm is needed to convert it into Cartesian coordinates so that it can be used by the inverse kinematics equations. For this purpose, a linear mapping method is used where the x and y coordinates of the alphabet are multiplied by a constant and scaled proportionally.

3.5 Replication of the alphabet on paper

The coordinates that are acquired after scaling up the data is parsed through 2 DOF inverse kinematics equations in order to convert them into joint angles. The joint angles thus acquired are written to the AX12A Dynamixel motors which move the mechanical structure and a pen attached to it, recreating the alphabet.

- Define the current (c) and goal (g) x-y points: $P_c = [x_c, y_c]$, $P_g = [x_g, y_g]$.
- Define the value X with NI equally-spaced points between x_c and x_g .
- Define the value Y with NI equally-spaced points between y_c and y_g .
- For each [X, Y], compute the inverse kinematics to obtain the associated angles as shown in the Figure 7:

$$\theta_1 = \cos^{-1} \frac{x}{\sqrt{x^2 + y^2}}$$

$$= \cos^{-1} \frac{[x^2 + y^2 + l_1^2 - l_2^2]}{2l_1 r}$$

$$\theta_2 = \cos^{-1} \frac{[x^2 + y^2 - l_1^2 - l_2^2]}{2l_1 l_2}$$

θ_1 is the angle of rotation of the motor 1

θ_2 is the angle of rotation of the motor 2

l_1 is the length of the link 1

l_2 is the length of the link 2 x_c is the current point

x_g is the goal point P_c is current position

P_g is the goal position

NI is the control resolution

X and Y are X and Y axis respectively

For the robotic arm used here, the length of the link 1 is 0.1075 m and the length of the link 2 is 0.365 m.

IV. RESULTS

The robot produced impressive replica of human hand writing of Alphabet "B" as shown in Figure.8



Figure 8. Robot Arm writing the Alphabet "B"

V. CONCLUSION

A 6 DOF Robotic arm with 5 yaw joints and 1 pitch joint is created. A program to acquire pixel coordinates from the laptop screen and scale it down to smaller size is created. Inverse Kinematics is applied to convert Cartesian coordinates into spherical coordinates. Thus acquired spherical coordinates are fed to the motor control program, which enabled the robot to replicate alphabets real-time. The alphabet or the word uttered by the user is converted into text and then used to control motors and replicate specific alphabets. As of now, the robot can write four letter words precisely.

VI. REFERENCES

- Hood, D.; Lemaignan, S.; Dillenbourg, P. When Children Teach a Robot to Write: An Autonomous Teachable Humanoid Which Uses Simulated

- Handwriting. In Proceedings of the 10th Annual ACM/IEEE International Conference on Human-Robot Interaction, Portland, OR, USA, 2–5 March 2015; pp. 1–8.
- [2]. H. Fujioka and H. Kano, "Design of cursive characters using robotic arm dynamics as generation mechanism," Proceedings 2006 IEEE International Conference on Robotics and Automation, 2006. ICRA 2006., Orlando, FL, 2006, pp. 3195-3200.
- [3]. N. Huebel, E. Mueggler, M. Waibel and R. D'Andrea, "Towards robotic calligraphy," 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems, Vilamoura, 2012, pp. 5165-5166.
- [4]. K. Franke, L. Schomaker and M. Koppen, "Pen force emulating robotic writing device and its application," IEEE Workshop on Advanced Robotics and its Social Impacts, 2005, pp. 36-46.
- [5]. V. Potkonjak, M. Popovic, M. Lazarevic and J. Sinanovic, "Redundancy problem in writing: from human to anthropomorphic robot arm," in IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 28, no. 6, pp. 790-805, Dec 1998.
- [6]. Tresset, P.; Leymarie, F. Portrait drawing by Paul the robot. *Comput. Graph.* 2013, 37, 348–363.
- [7]. www.robotis.com.nt)