

Implementation of a Hybrid Voice guiding System for Robot Arm

Mohamed FEZARI, Hamza Attoui, N.E. Debbache and R. Lakel

Faculty of Engineering, Department of Electronics, Badji Mokhtar University, Annaba
Laboratory of Automatic and Signals, Annaba, BP.12, Annaba, 23000, ALGERIA
Mohamed.fezari@uve.ac.uk, hattoui@yahoo.fr, ndebbache@yahoo.fr

Abstract- In this paper, a voice command system for a robot arm (manipulator) is designed as a part of a research project. The methodology adopted is based on a hybrid technique used in automatic speech recognition system. To implement the approach on a real-time application, a Personal Computer interface was designed to control the movements of a five degree of freedom robot arm by transmitting the orders via radio frequency circuits. The voice command system for the robot arm is designed. The main parts of the robot are a microcontroller from Microchip PIC16F84 and a radio frequency receiver module. The possibility to control other parts of the automatic arm are investigated.

Keywords: Human-machine interaction, Hybrid technique, DTW, Voice command, Robot Arm.

I. INTRODUCTION

Human-robot voice interface has a key role in many application fields. Robotics has achieved its greatest success to date in the world of industrial manufacturing. Robot arms, or *manipulators*, comprise a 2 billion dollar industry. Bolted at its shoulder to a specific position in the assembly line, the robot arm can move with great speed and accuracy to perform repetitive tasks such as spot welding and painting. In the electronics industry, manipulators place surface-mounted components with superhuman precision, making the portable telephone and laptop computer possible [1].

Yet, for all of their successes, these commercial robots suffer from a fundamental disadvantage:

lack of human voice control. A fixed manipulator has a limited range of commands provided by a manipulator. Mainly using a keyboard, joystick or a mouse.

This paper proposes a new approach to the problem of the robot arm command. Based on the recognition of isolated words, using a set of traditional pattern recognition approaches and a discrimination approach based on test results of classical methods [2][5] and [7] in order to increase the rate of recognition. The increase in complexity as compared to the use of only traditional approach is negligible, but the system achieves considerable improvement in the matching phase, thus facilitating the final decision and reducing the number of errors in decision taken by the voice command guided system.

Moreover, speech recognition constitutes the focus of a large research effort in Artificial Intelligence (AI), which has

led to a large number of new theories and new techniques. However, it is only recently that the field of robot arm control and AGV navigation have started to import some of the existing techniques developed in AI for dealing with uncertain information.

Hybrid method is a simple, robust technique developed to allow the grouping of some basic techniques advantages. It therefore increases the rate of recognition. The selected methods are : Zero Crossing and Extremes (CZEXM), linear Dynamic Time Warping (DTW), Linear Predictive Coefficient (LPC) parameters , Energy Segments (ES), and cepstral coefficients. This study is part of a specific application concerning robot control by simple voice commands. The application uses a set of commands in Arabic words in order to control the directions of five degree of freedom robot arm. It has to be implemented on a DSP [8] and has to be robust to any background noise confronted by the system.

The aim of this paper is therefore the recognition of isolated words from a limited vocabulary in the presence of background noise. The application is speaker-dependent. Therefore, it needs a training phase. It should, however, be pointed out that this limit does not depend on the overall approach but only on the method with which the reference patterns were chosen. So by leaving the approach unaltered and choosing the reference patterns appropriately, this application can be made speaker-independent [9].

As application, a vocal command for a five degree of freedom robot arm is chosen. The robot arm is "TERGANE - T45". There have been many research projects dealing with robot control, among these projects, there are some projects that build intelligent systems [10-12]. Since we have seen human-like robots in science fiction movies such as in "I ROBOT", making intelligent robots or intelligent systems became an obsession within the research group. Voice command needs the recognition of isolated words from a limited vocabulary used in Automatic Robot Arm Control system (RACS)[13] and [14].

II. DESIGNED APPLICATION DESCRIPTION

The application is based on the voice command for a set of motors for a robot arm T45 [20]. It therefore involves the recognition of isolated words from a limited vocabulary used to control the movement of selected parts of the arm.

The vocabulary is limited to nine words to select the arm part (upper limb, limb, hand and forceps or grip) and to command

the selected part (up, down, left right and stop). These commands are necessary to control the movement of the T45, Up movement, Down movement, stop, turn left and turn right. The number of words in the vocabulary was kept to a minimum both to make the application simpler and easier for the user.

The user selects the robot arm part by its name then gives the movement order on a microphone, connected to sound card of the PC. A speech recognition agent based on hybrid technique recognises the words and send to the parallel port of the PC an appropriate binary code. This code is then transmitted to the robot T45 via a radio frequency emitter.

The application is first simulated on PC. It includes two phases: the training phase, where a reference pattern file is created based on pre-recorded data base, and the recognition phase where the decision to generate an accurate action is taken. The action is shown in real-time on parallel port interface card that includes a set of LED's, showing what command is taken, and a radio Frequency emitter.

III. THE SPEECH RECOGNITION AGENT

The speech recognition agent is based on a traditional pattern recognition approach. The main elements are shown in the block diagram of Figure 1. The pre-processing block is used to adapt the characteristics of the input signal to the recognition system. It is essentially a set of filters, whose task is to enhance the characteristics of the speech signal and minimize the effects of the background noise produced by the external conditions and the motor.

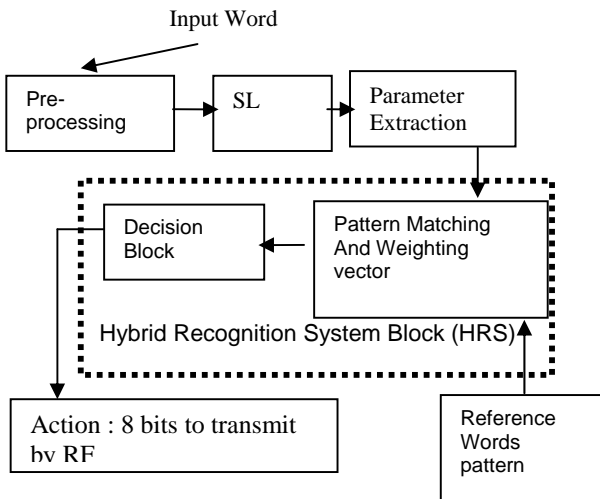


Figure 1: Block Diagram

The Speech locator (SL) block detects the beginning and end of the word pronounced by the user, thus eliminating silence. It processes the samples of the filtered input waveform, comprising useful information (the word pronounced) and any noise surrounding the PC. Its output is a vector of samples of the word (i.e.: those included between the endpoints detected).

The SL implemented is based on analysis of crossing zero points and energy of the signal, the linear prediction mean square error computation helps in limiting the beginning and the end of a word; this makes it computationally quite simple.

The parameter extraction block analyses the signal, extracting a set of parameters with which to perform the recognition process. First, the signal is analysed as a block, the signal is analysed over 20-mili seconds frames, at 256 samples per frame. Five types of parameters are extracted: Normalized Extremes Rate with Normalized Zero Crossing Rate (CZEXM), linear DTW with Euclidian distance (DTWE) as in figure 2, LPC coefficients (Ai), Energy Segments (ES) and cepstral parameters (Ci) [14].

These parameters were chosen for computational simplicity reasons (CZEXM, ES), robustness to background noise (12 Cepstral parameters) and robustness to speaker rhythm variation (DTWE) [19].

The reference pattern block is created during the training phase of the application, where the user is asked to enter ten times each command word. For each word and based on the ten repetition, ten vectors of parameters are extracted from each segment and stored.

The matching block compares the reference patterns and those extracted from the input signal. The matching decision integrate: a hybrid recognition block based on five methods, and a weighting vector.

Tests were made using each method separately. From the results obtained, a weighting vector is extracted based on the rate of recognition for each method. Figure 1 shows the elements making up the main blocks for the hybrid recognition system (HRS).

A parallel port interface was designed to show the real-time commands. It is based on two TTL 74HCT573 Latches and 10 light emitting diodes (LED), 4 green LED to indicate each recognized word for robot arm parts ("Dirrar", "Saad", "Meassam", "mekbadh") and 5 yellow LED to indicate each recognised command word ("fawk", "tahta", "Yamine", "Yassar", "Kif"), respectively and a red LED to indicate wrong or no recognized word. Other LED's can be added for future insertion of new command word in the vocabulary example the words: "Iftah", "Ighlak" for the grip as shown in Figure 2.

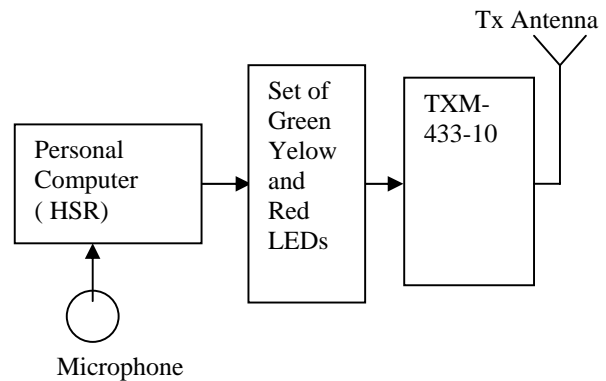


Figure 2 : Block Diagram of Parrallel Port Interface.

A voice command system and an interface to the robot arm are implemented. The voice commands of T45 via a radio frequency transmitter system are selected from commands used to control a Robot arm and their meanings are listed as in Table 1,

Table 1. The meaning of voice commands

1) Diraa	Upper limb motor (M1)
2) Saad	Limb motor (M2)
3) Meassam	Wrist (hand) motor (M3)
4) Mikbath	Grip motor (M4)
5) Fawk	Up movement M1, M2 and M3
6) Tahta	Down movement M1, M2 and M3
7) Iftah	Open Grip, action on M4
8) Ighlak	Close grip, action on M4
9) Kif	Stop the movement, stops M1, M2 M3r or M4

IV. ROBOT ARM CONTROL SYSTEM

As in Figure 3, the structures of the mechanical hardware and the computer board of the five degree of freedom robot arm in this paper are similar to those in [5-6]. However, since the robot arm in this paper needs to perform simpler tasks than those in [5-6] do, robot arm electronic control system can be designed simply. The computer board consists of a PIC16F84, with 1K-instruction EEPROM (Electrically Programmable Read Only Memory) [15], four H bridges drivers using BD134 and BD133 transistors for DC motors, a RF receiver module from RADIOMETRIX which is the SILRX-433-10 (modulation frequency 433MHz and transmission rate is 10 Kbs) [16]. In order to protect the microcontroller from power feedback signals, a set of opto-transistors were added. Each motor within the robot arm T45 is designed by a name and performs the corresponding task to a received command as in Table 1. Commands and their corresponding tasks affected to robot arm motors may be changed in order to enhance or addapt the application.

In the recognition phase, the application gets the word to be processed, treats the word, then takes a decision by setting the corresponding bit on the parallel port data register and hence the corresponding LED is on. The code is also transmitted in serial mode to the TXM-433-10.

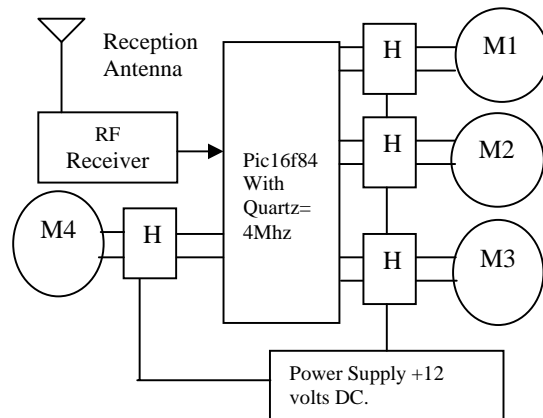


Figure 3. Robot Arm Control System block diagram

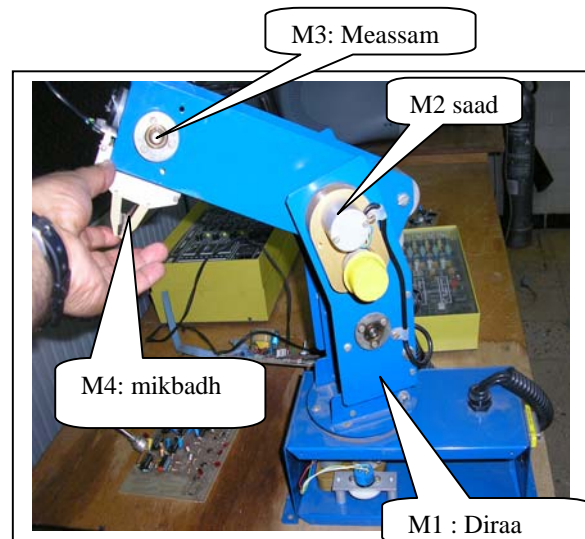


Figure 4. Overview of the Robot Arm T45 and Parallel interface

V. BENEFITS OF THE DESIGNED SYSTEM

In this work, a voice command system is designed and is constructed to manipulate a robot arm T45.

This voice command system can be applied to other systems as well as robot systems. The followings are the advantages of the proposed system:

1) The proposed system, to command and control a robot arm by human voices, is totally based on microcontroller and a PC central command as shown in figure 5.

2) Compared with previous projects that build intelligent systems as in [5-6], the cost of the proposed voice command system is much lower.

3) A Manipulator robot arm controlled by human voices is one of the projects that can be assigned to a heterogenous research group and therefore require the cooperation of each member. Depending on the research field of group members, this robot arm control can be divided into several modules and each module can be assigned to one individual researcher. For example, one person designs a voice command system and the other person the architecture for the robot arm, while a third person may work on behaviour of arm tracking objects.

4) Several interesting competitions of voice-controlled robot arm will be possible. In the competition, robots will perform various tasks by human voice commands. Each team may use its own voice commands different from those of other teams. One example of the competitions is a robot arm objects detection and classification by human voice commands.

This is different from a robot arm surgery by vision as in [17]. Another example is to get the robot arm detecting a plate take the food from a plate and feed a handicapped person by human voice commands.

5) While previous intelligent systems as in [5][6][17] are under a full automatic control, voice-controlled robot arm is under a supervisory control. Therefore, it can be used to solve some problems in the supervisory control. One of problems in supervisory control is due to the time delay. The time delay mainly caused by the recognition time of voices and the time of reception of an RF signal then reaction of the robot, the effect of time delays in controlling robot arm manipulator can be observed.

6) Other systems besides robot arm can be combined with the proposed voice command system. For example, a voice-controlled remote switch can be built by using this voice command system, an infrared transmitter/receiver pair, and some relays. A voice-controlled remote controller of consumer electronic products can also be built by using this voice command system and an infrared transmitter/receiver pair if the codes of the remote controller in consumer electronic products are known, example the RC5 code for Philips products[18].

VI. EXPERIMENTS ON THE SYSTEM

The developed system has been tested within the laboratory of L.A.S.A, there were two different conditions to be tested: The distance of the microphone from the speaker, and the rate of recognition in periodic noise and nonstationary noise (NSN) environment. The system, first, had been tested in the laboratory and outside in order to detect the environment effect on the recognition rate, then. After testing the recognition of each word 25 times in the following conditions: a) outside the Laboratory (LASA) with NSN, b) outside the LASA with periodic noise (PN), c) inside the LASA with NSN, and d) inside the LASA with PN. The results are shown

in figure 5 where the numbers in abscess axe corresponds to the order of voice command word as they appear in Table 1.

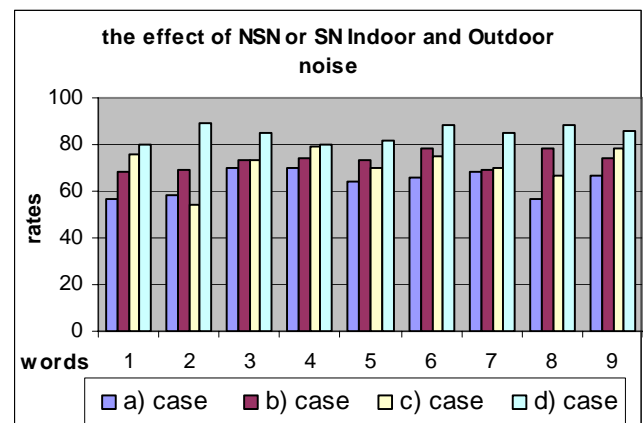


Figure 5. The effect of SN or NSN in and out the laboratory

VII. CONCLUSION AND FUTURE DIRECTIONS

A voice command system for robot arm manipulator is designed and is implemented based on a HRS for isolated words. Since the designed robot consists of a microcontroller and other low-cost components namely RF transmitters, the hardware design can easily be carried out. The results of the tests shows that a better recognition rate can be achieved inside the laboratory and especially if the phonemes of the selected word for voice command are quite different. However, a good position of the microphone and additional filtering may enhance the recognition rate. Several interesting applications of the proposed system different from previous ones are possible as mentioned in section 5.

Beside the designed application, a hybrid approach to the implementation of an isolated word recognition agent HSR was used. This approach can be implemented easily within a DSP or a CMOS DSP microcontroller.

The use of hybrid technique based on classical recognition methods makes it easier to separate the class represented by the various words, thus simplifying the task of the final decision block. Tests carried out have shown an improvement in performance, in terms of misclassification of the words pronounced by the user. The increase in computational complexity as compared with a traditional approach is, however, negligible. Segmentation of the word in three principal frames for the Zero Crossing and Extremes method gives better results in recognition rate. The idea can be implemented easily within a hybrid design using a DSP with a microcontroller since it does not need too much memory capacity. Finally we notice that by simply changing the set of command words, we can use this system to control other objects by voice command such as an electric wheelchair movements or a set of autonomous robots. One line of research is under way is the implementation of feedback force sensors for the grip and adapted algorithms to get smooth objects.

