SPOTTED WORDS TO COMMAND A SET OF MOBILE ROBOTS

MOHAMED FEZARI, Hamza Attoui
University of Annaba,Faculty of Engineering, Department of Electronics
Laboratory of Automatic and Signal, Annaba, BP.12, Annaba, 23200,ALGERIA
Mohamed.fezari@uwe.ac.uk, hattoui@yahoo.fr

Abstract: *An approach based on spotted word recognition system is applied to control a set of autonomous robots. The methodology adopted is based on hybrid techniques used in speech recognition which are zero crossing and extremes with dynamic time warping followed by a decision system based on independent methods test results. To test the approach on a real application, a PC interface was designed to control the movement of five mobile robots using a radio frequency transmission. Also, it was checked whether the basic commands used would be enough to control the robots. Some experiments are tested by actually running a set of powered-toy vehicle and each robot is secured from frontal collision by some sensors.*

Keywords: Voice command, hybrid technique, spotted words, Robots**.**

1. INTRODUCTION

Speech command has a key role in many application fields. Various studies made in the last few years have given good results in both research and commercial applications using automatic speech recognition systems. However, the performances are not equal to those of humans, especially in real conditions: spontaneous speech, in presence of noise, continuous talk, mixed signals [1-3]. This paper proposes a new approach to the problem of the recognition of spotted words, using a set of traditional pattern recognition approaches and a decision system based on test results of classical methods in order to increase the rate of recognition [4-5]. 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.

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 and Automatic Guided Vehicle (AGV) navigation have started to import some of the existing techniques developed in AI for dealing with uncertain information.

The recent sophisticated methods for intelligent signal processing have led to the development of various applications of hybrid techniques in the field of telecommunications as well as automatic speech recognition. 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 (ZCEXM), Dynamic Time Warping with Cepstral coefficients (DTWC), and Linear Predictive Coefficient (LPC) parameters.

Since the apparition of "I ROBOT" movie on the screen, making intelligent robots or intelligent systems became an obsession within the research group in Laboratory of Automatic and Signal, Annaba (LASA).

As application, a vocal command for a set of autonomous robots is chosen. Voice command needs the recognition of isolated words from a limited vocabulary used in AGV (Automatic Guided Vehicle) system [6-9]. A micromouse in [6] is a computer controlled autonomous mobile robot that can find a predetermined destination when placed in an unknown maze. The micromouse competition has earned high marks on publicity and technology. But the micromouse competition has to be conducted on a maze of unreasonable specification [7]. A vision-guided autonomous vehicle is proposed as an alternative of the micromouse competition in [7]. Thus, an other approach to control a set of robots is voice command system based on spotted words instead of isolated vocabulary.

This study is part of a specific application concerning system control by simple voice commands. The application uses five Arabic words as robots names and six other Arabic words as commands. The operator can use, within a phrase of many words, a selected robot (robot name) and a task affected to that robot.. However, some spotted words should be used in order to generate a task for a special robot, mainly: the robot name and the task affected to it. The system will recognise the spotted words which are the robot name and the task then generates on the parallel port of the Personal Computer (PC) an 8 bit command, where the four first bits design the command code and the following four bits design the robot name code. The 8 bits will be transmitted by a radio frequency module.

2. DESCRIPTION OF THE DESIGNED APPLICATION

The application is a design of a voice command system for a set of mobile robots. It therefore involves the recognition of spotted words from a limited vocabulary used to select a robot and to control the movement of that robot.

The application is designed in two steps which are the speech recognition system implemented on a PC and the mobile robot design. The speech recognition system gets a sentence from the operator; it isolates the words, and then processes each word in order to recognise the spotted words. The recognised words within this sentence should have a correct syntax phrase that means: a name of a robot and a command word. Example: if the users says: " robot red go forward" then the system will recognise the robot name: "Red" and the command word "forward".

Once the robot name and the command were recognised, the system will generate on the parallel port of the PC a binary code. The three most significant bits define the robot name and the 5 other bits define the task to be performed by that mobile robot.

The mobile robots task specifications are limited to six commands which are necessary to control the movement of a robot: switching on and off the engine 'motor', forward movement, backward movement, stop, turn left and turn right. More commands can be added. However, the number of words in the vocabulary was kept to a minimum both to make the application simpler and easier for the user.

The six words in the Arabic vocabulary are the following:

- "Mouharek": This switches the motor on or off at average speed.
- "Ameme": This makes the movement upward.
- "Wara": This makes the movement backward.
- "Kif": This command stops the movement.
- "Yamine": This makes the turn right;
- "Yassar": This makes the turn left.

The mobile robots names were chosen as colours in Arabic language for these robots: "Ahmar", "Azrak", "Akhdar", "Asuad", and "asfar" witch mean respectively : "Red", "Blue", "Green", "Black" and "yellow".

The application is first simulated on PC. It includes two phases: the training phase, where a reference pattern file is created based on the 11 spotted words 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 and a Radio frequency transmitter module.

3. SPEECH RECOGNITION S*YSTEM MAIN PARTS*

The speech recognition system 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.

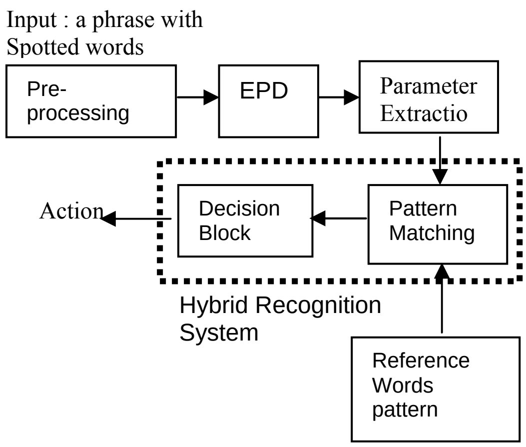

Figure 1: Block diagram of speech recognition system (spotted words).

The End Point Detector (EPD) block detects the beginning and end of each word pronounced by the user, thus eliminating silence and separating words. It processes the samples of the filtered input waveform, comprising useful information (the spotted words). Its output is a vector of samples for each word (i.e.: those included between the endpoints detected).

The EPD is based on analysis of crossing zero points and energy of the signal [3], 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 10-mili seconds frames, at 100 samples per frame. Three types of parameters are extracted: ZCEXM, DTWC and LPC.

These parameters were chosen for computational simplicity reasons (ZCEXM), robustness to background noise (10 Cepstral and LPC parameters) and robustness to speaker rhythm variation (DTWC). The parameter extraction and ordering tool made the task simpler and more efficient. In addition, calculation of the Cepstral parameters does not create an additional computational load

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

The matching block compares the reference patterns and those extracted from the input signal. The matching and decision integrate: a hybrid recognition block based on the three 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 2 shows the elements making up the main blocks for the hybrid recognition system (HRS). The inputs of HRS block are three vectors representing the parameters of the input word obtained from the three methods. The Hybrid Recognition block compares the input parameters with the references parameters of the six command words and five robots names. It then generates a vector of five values; the elements are the recognized word number. If the input word is “Amame” and the ZCEXM method recognizes the word then it generates the number 2 (which means the second command word).

However it may recognize another word from the used vocabulary or none of them, in this case it generates either the number of the recognized word (1 to 5 for commands and 6 to 11 for robot names) or a Zero.

The last block in the HRS is the weighting block. It is made of a threshold system which discriminates which of the six words was pronounced by the user on the basis of the information provided by the HRS. The resulting vector is multiplied by the weighting block that contains three rates. These rates were fixed based on results obtained by tests of each method. The best word to fit the input word is chosen.

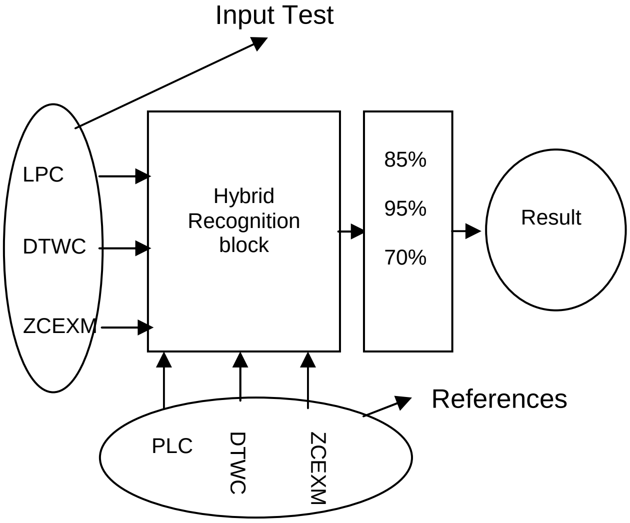

Figure 2**:** HRS Block Structure.

4. INTERFACE CARD FOR IMULATION

A parallel port interface was designed to show the real-time spotted words command. It is based on two TTL 74HCT245 buffers and 16 light emitting diodes (LED), six green LED to indicate each recognized command, five yellow LED to indicate each recognised robot name and a red LED to indicate wrong phrase or no recognized command. The other LED's were added for future insertion of new command word in the vocabulary example the words: "Faster", "slower" and "light" and a radio frequency transmission module is used to send task to mobile robots as shown in Figure 3. A graphic user interface (GUI) was developed with MATLAB language under windows XP as operating system In GUI, the training and recognition phases are presented in the right, the six LED for the six words to recognize are presented by squares, the red square is the recognized word and the pronounced command is plotted in blue as shown in Figure 5.b

The training phase is the first step in which the parameters for the vocabulary were created. In this phase, the speaker repeats five times each word, the parameters are produced and saved in a file.

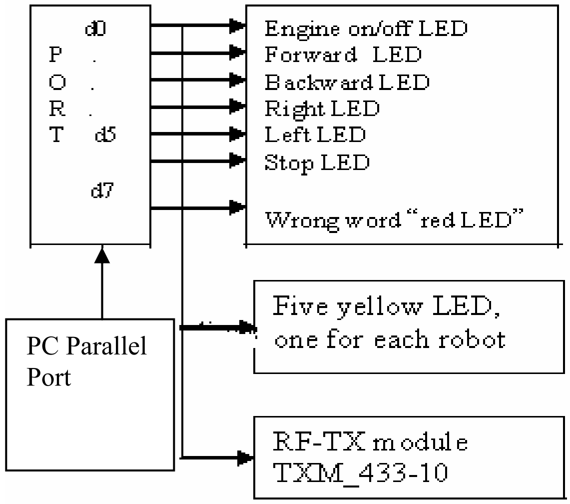

Figure 3 : Block diagram of parallel port interface.

In the recognition phase, the application gets a phrase containing the spotted words to be processed, treats the words, then takes a decision by setting the corresponding bit on the parallel port data register and hence the corresponding LED is on. The corresponding box on the GUI is changed to red colour. The corresponding LED's on parallel port are set to on and an 8-bit binary code is sent to the RF transmitter.

5. AUTONOMOUS ROBOT MAIN PARTS

As shown in Figure 4, the structures of the mechanical hardware and the computer board of autonomous robots in this paper are quite similar to those described in [6]. However, since the autonomous robots in this paper need to perform simpler tasks than those in [6] and [12] do, these autonomous robots can more easily be constructed. The computer board of each autonomous robot consists of a PIC16F84 microcontroller , with 1K-instruction EEPROM (Electrically Programmable Read Only Memory) [10], two motor controller circuits, it consists of an L293 which is an H bridges drivers (600 mA) for DC motors, a RF receiver module from RADIOMETRIX which is SILRX-433-10 (frequency 433MHz and transmission speed 10kbs) that has been used in the remote controller of consumer electronic products

[11], and a four bit micro-switch to fix the address of each robot. In addition, to avoid frontal obstacles, we integrated on each robot an Ultrasonic module type MSU05 from LEXTRONIC [14]. Each autonomous robot performs the corresponding task to a received command. However, if there is any conflict to execute the ordered task (an obstacle in front of the robot) the robot will not execute this task.

Commands and their corresponding tasks in autonomous robots may be changed in order to enhance or change the application. For example, adding horn and light or adding other robots.

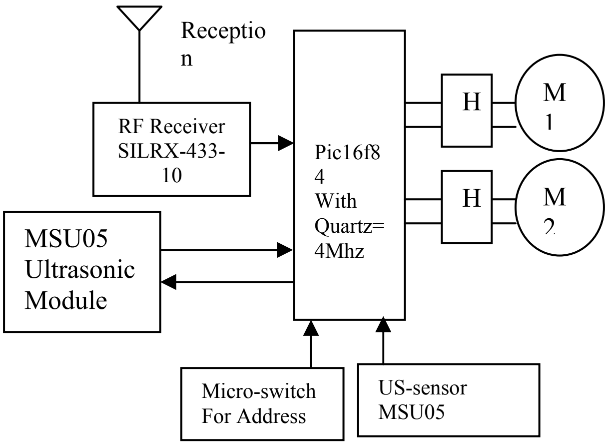

Figure 4. Autonomous robot block diagram

6. RESULTS OF SIMULATION

First, some tests were done on each method and the rate of recognition was registered In the ZCEXM method, the rate is lower; however the rate is speaker dependant. DTWC gives better results specially if there is any distortion in locution rhythm. LPC and Cepstral coefficients give better rates than those cited earlier. And based on the tests for each method separately, a weighting vector is produced. The second tests were made using the hybrid method, the recognition rate is improved as it is shown if figure 5.a.

The Ultrasonic module MSU05 was tested and gave acceptable results. The error in computing the distance from the obstacle is 1 to 2 cm for a distance of 10 cm to 200 cm [13].

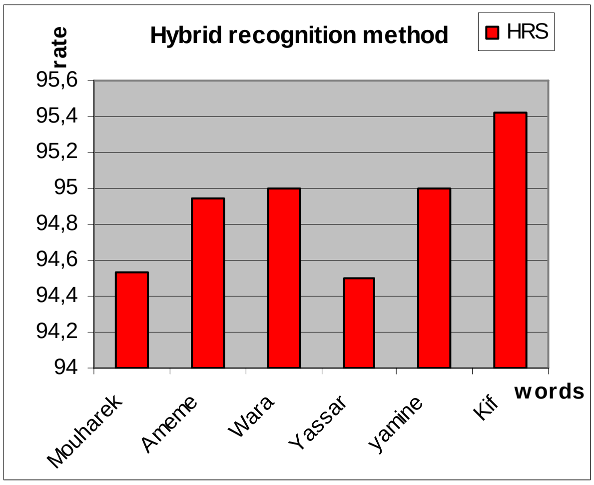

Figure 5.a: Recognition rate for the six command words using hybrid method.

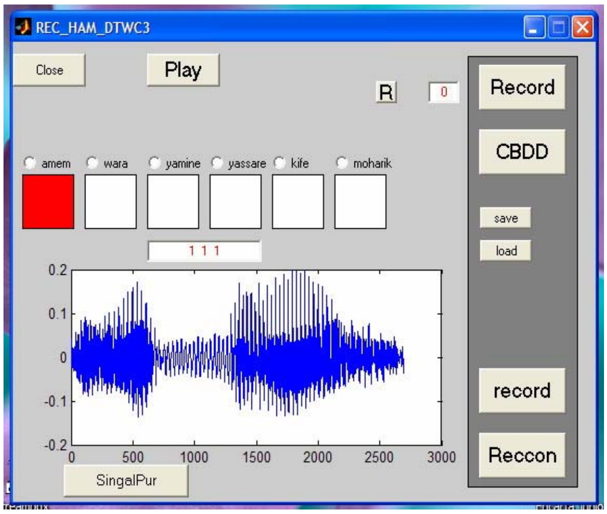

Figure 5.b. GUI Human-application Interface

7. CONCLUSION AND FUTUR WORK

This article presented a new methodological approach to the implementation of spotted words recognition system based on a mixed techniques to control a set of autonomous robots.

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 frames for the Zero Crossing and Extremes method gives better results in recognition rate. The method can be implemented into a hybrid design using a DSP and a microcontroller since it does not need too much memory space.

Moreover, an autonomous robot controlled by human voice is one of the projects that can be assigned to a heterogynous research group and therefore require the cooperation of each member. Depending on the research field of group members, this autonomous robot 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 designs an autonomous robot with sensors, while a third person may work on behaviour of several robots.

The reinforcement learning algorithms will be investigated as a future work. Results from research and development efforts in this area should have application to a broad range of assistive technology systems. 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 intelligent-robot colony or robot-hands control.

REFERENCES

[1] Gu L. and Rose K., “Perceptual Harmonic Cepstral Coefficients for Speech Recognition in Noisy Environment,”, *In Proc ICASSP 2001, Salt Lake City,* Utah, May 2001.

[2] Wang T. and V Cuperman., "Robust Voicing Estimation with Dynamic Time Warping," *Proceedings IEEE ICASSP'98*, , pp. 533-536, 1998.

[3] Niles L. Silverman H., “Combining Hidden Markov Models and Neural Network classifiers”,. In: CASSP'90, 1990, pp. 417–420, 1990.

[4] Glotin H., and Berthommier F., “Test of several external posterior weighting functions fur multiband full combination ASR”, *In proceedings of ICSLP 2000*., pp. 333–336, 2000.

[5] YOSHIZAWA S., HAYASAKA N., WADA N. and MIYANAGA Y. ,“ Cepstral Gain Normalization for Noise Robust Speech Recognition”, *in proceedings of International Conference on speech and signal processing ICASSP 2004*, Montreal Canada, pp: 209-212, 2004.

[6] Gulec N., Unel M. and Sabanovic A., “Coordinated Task Manipulation by Nonholonomic Mobile Robots”, I*n proc. AMC the 9th International workshop on Advanced Motion Control, Turkey,* pp. 255-260, *2006.*.

[7] Chen N., “A Vision-Guided autonomous vehicle: An alternative micromouse competition”, *IEEE Transactions on Education*, vol. 40, no. 4, pp. 253-258, 1997.

[8] Kim J. H., et al, “Cooperative Multi-Agent robotic systems: From the Robot-Soccer perspective”, 1997 *Micro-Robot World Cup Soccer Tournament Proceedings*, Taejon, Korea, 1997, pp. 3-14, 1997.

[9] Durali M., Amini Javid G. and Kasaiezadeh A., ” Collision Avoidance Maneuver for an Autonomous Vehicle “, In proc. *AMC the 9th International workshop on Advanced Motion Control, Turkey*, 2006, pp. 49-54, 2006.

[10] Data sheet PIC16F84 from Microchip inc. User’s Manual, 2001, http://www.microchip.com.

[11] Data sheet , TSM-433-10 and SILRX-433-10, User’s manual, 2005, http://www.radiometrix.com

[12] J. H. Kim, et al, “Cooperative Multi-Agent robotic systems: From the Robot-Soccer perspective”, 1997 Micro-Robot World Cup Soccer Tournament Proceedings, Taejon, Korea, 1997, pp. 3-14, 1997.

[13] Data sheet and application using SRF05 (MSU05) Ultrasonic sensor, 2005.
http://www.robot-electronics-co-uk-htm-sfr05.tech.htm.

[14] M. Fezari and M. Bousbia-Salah, “ Voice and Sensors to Improve the Control of Wheelchair by Disabled Persons”, *In Proceedings of International Conference on Human Machine iNteration*, HuMaN’07, Timimoun Algeria, march 2007,pp.295-298, 2007.