# Voice-Driven Monitoring Robots Navigation

Rahul R Reddy
Student, Dept. of E & C,
BTL Institute of Technology and Management,
Bangalore 560 099,   India

Poornima H L
Student,  Dept. of E & C,
BTL Institute of Technology and Management,
Bangalore 560 099,   India

Kavya M
Student, Dept. of E & C,
BTL Institute of Technology and Management,
Bangalore 560 099,   India

Kavana J
Student, Dept. of E & C,
BTL Institute of Technology and Management,
Bangalore 560 099,   India

*Abstract:* **The multiprocessor system is composed of pipelining a DSP processor for speech enhancement, a new voice recognition module for isolated word recognition and a microcontroller for transforming the voice order into  coded byte to be transmitted to the robots via wireless transmission module. The resulting design is used to control via master-slave protocol a set of small mobile robots with sensors. The speech processor recognises the  words then generates a code. A microcontroller produces a coded byte and transmit in to via a RF module the to robots. Since the system is an embedded device developed in order to be portable, it should be easy to carry and use, with low power consumption, thus the choice of power less consumption processors.**

*Keywords: Speech recognition; Embedded systems; Mobil Robot; VR-Stamp processor and Bluetooth.*

## I. INTRODUCTION

The easiest way of making a robot go to a goal location is simply to guide it to this location. This guidance can be done in different ways: burying an inductive loop or magnets in the floor, painting lines on the floor, by placing beacons, markers, bar codes, by vision, rare are paper dealing with voice command to guide mobile robots[1] The main requirement for a service robot in human robot communication is to provide easy humanlike interaction, which on the one hand does not load the user too much and on the other hand is  effective in the sense that the robot can be kept in useful work as much as possible. Note that learning of new tasks is not counted as useful work! The interface should be natural for human cognition and based on speech and gestures in communication. Because the robot cognition and learning capabilities are still very limited the interface should be optimized between these limits by dividing the cognitive tasks between the human brains and robot "intelligence" in an appropriate way[1-3]. Mobile robots are also found in industry, military and security environments. They also appear as consumer products, for entertainment or to perform certain tasks like vacuum cleaner. Techniques and materials are introduced to improve the command of a mobile Robot such as IR-telecontrol module or program sequencer or a computer and speech recognition techniques namely: DTW (Dynamic Time Warping), Crossing Zero, HMM (Hidden Markov Model) and GMM ( Gaussian Mixture Model) [4-5]. And a way to increase the rate of recognition is to process the input signal in order to eliminate any distortion and noises that can affect the speech signal during recording phase [3-4]. Because, special components for speech processing have been emerged in the last decade and implemented in various application fields [5]. Also thanks to the fact that increasingly faster computers have become accessible for simulation and emulation of the new components to a growing number of users, special processor in signal processing implemented in DK (system Development Kit) such as SDK TMS320C6711 or C6713 made pre-processing steps more independent and can be presented as modules, in this work the DSP TMS is used to eliminate echoes and background noises based on implemented Kalman filter [6]-[10].

This paper proposes a new approach to the problem of the recognition of spotted words, using a speech recognition development kit from 'sensory' and implements it for voice command of a colony of small robots from Pob-Tech, a set of robots type POB-BOT with the minimum of sensors in laboratory LASA [8]- [14]. The study is part of a specific application concerning system control by simple voice commands. The objective of this design is therefore the recognition of spotted words from a limited vocabulary in the presence of stationary background noise. This application is speaker-dependent. However, it should 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. To enhance the designed, a pre-processing step is added using a DSP TMS320c6711. A wireless transmission of the commands is provided by Bluetooth modules.

The application to be integrated in this embedded system is first simulated using MPLAB, then implemented in a RISC architecture microcontroller adapted to a speech recognition development kit 'Easy-VRStamp produced by MikroElectroniKa'. Experimental tests showed the validity of the new hardware adaptation and Test Results, within the laboratory experience area, are acceptable.

**Special Issue - 2016**

**International Journal of Engineering Research & Technology (IJERT)**
**ISSN: 2278-0181**
**ICRET - 2016 Conference Proceedings**

## II. GENERAL DESCRIPTION OF THE DESIGNED EMBEDDAD SYSTEM

The designed System as shown in Figure 1.a and 1.b is developed as a client-server system, the client system is composed of the following components:

- The VR-Stamp based on RSC4128 special processor, which is the heart of the vocal command system [6].
- A DSP TMS320C6711, to eliminate echo and background noises based on Kalman-filter.
- A microcontroller PIC18F252 as a main processor.
- A special designed keyboard with eight switches, four interrupters for robot selection and four push buttons for direction control.
- And a Bluetooth module Rok101007 from Ericsson Microelectronics.

These components are controlled by a CMOS- RISC microcontroller from Microchip, a new generation of powerful computation, low-cost, low- power microcontrollers. The client-system is fed by a rechargeable Li+ battery as a power supply.

The server-system is composed of the following main components:

- A set of Four mobile robots from of type POB- BOT From POB-Technology (www.pob- technology.com)
- A Bluetooth (POB-Tooth) as a wireless communication interface.

Unsatisfactory functionality is still the major reason for the poor customer's acceptance of speech control. One could improve the acceptance with high level speech recognizers and sophisticated user interfaces. On the other hand, speech command for robots is expected to provide an accuracy, to be robust against noise and environment variations, but also to be very cost efficient.
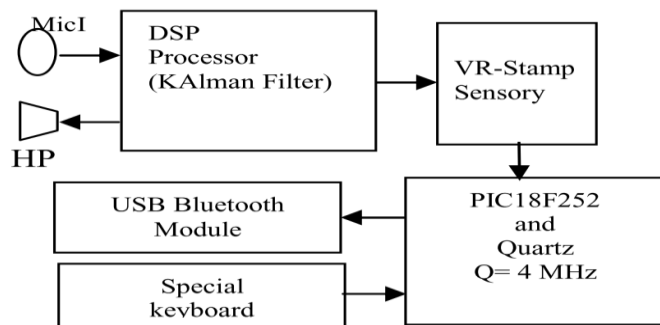


Figure.1.a Block diagram of the master system.



Figure 1.b Block diagram of
Slave - system Bluetooth module integrated onto one of set of four mobile robots Pob-Bot

## III. DETAILS OF MASTER PART

The master part is developed around the VD364 and a DSP processor controlled by the PIC18f252. For best performance, the system gives better results in a quiet environment with the speaker's mouth in close proximity to the microphone, approximately 5 to 10 cm.

### 3.1. *DSP processor*

A TMS320C6711 DSP processor were used to do two jobs, enhancing the speech signal by reducing the environment noise using Kalman Filter and reducing the effect of echo. Moreover this unit presents words of the sentence as a set of isolated and filtered words to the speech processor VR-stamp. The TMS320C6711 DSK module was chosen as it provides low cost gateway into real-time implementation of DSP algorithms. This module has the following features: A 150MHz TMS320C6711 DSP capable of executing 1200 Million Instruction Per Second (MIPS), 4M-bytes of 100MHz SDRAM, 128K-bytes of flash memory, a 16-bit audio codec, a parallel port interface to standard parallel port on a host PC. The TMS320C6711 DSK module is accompanied by the Code Composer Studio IDE software, developed by Texas Instruments.

### 3.2 *VR-Stamp ( Voice recognition Stamp)*

Voice Recognition Stamp is a new component from Sensory inc. It has more capabilities designed for embedded systems. It was designed for consumer telephony products and cost-sensitive consumer electronic applications such as home electronics, personal security, and personal communication because of its performances: - Noise-robust Speaker Independent (SI) and Speaker Dependent (SD) recognition. - Many language models now available for international use. - High quality, 2.4-7.8 kbps speech synthesis & Speaker Verification Word Spot (SVWS) -Noise robust voice biometric security. The module VR-Stamp is based on the following components: a special microcontroller RSC4128, a reference word storage 24C65 of EEPROM type that holds the parameters of referenced word produced during the training phase, a Flash program memory of 4 Mega-byte that holds the main program of word recognition, and a parallel interface of 24 lines ( divided into 3 by 8-bit ports) to generate the results

of recognition or to introduce commands, and audio communication lines for microphone and speakers .

In training phase, the module gets features of the 10 spotted words used in the vocabulary and presented in table 1, among these words, the starting keyword "Lasa" which is the name the laboratory and finishes with the Keyword "Tabek" which means 'execute' the command, so whenever the user wants to submit a voice command the sentence should start with the word "Lasa" and finishes with the word "Tabek". In recognition phase the VR-Stamp should detect some spotted words in the sentence and then submit the code of recognized words to the microcontroller, example: " Lasa Aswad Yassar tabek", in this sentence the module will submit the codes: 3 for the robot name and 7 for the action to be taken by that robot, the codes are presented in table 1 with the corresponding words.

### 3.3 The microcontroller PIC18F252

As an interface between the wireless transmission circuit Bluetooth and the vocal module VD364, a microcontroller with at least 16 input/output lines and minimum of 4 kilo instructions is needed. Therefore a better choice was the PIC18F252 from Microchip [17]. The main function of the microcontroller is to get the information from the VR-STAMP and based on the order of the codes it will submit this command to the Bluetooth or signal to the user (client) that an error in recognition or in comprehension: as an error recognition every sentence with no starting word, or non recognised word. As comprehension error, a sentence containing correct spotted words however it does not have a meaning: yade Mikbath Fawk". The microcontroller gets also high priority command from special keyboard.

### 3.4 Bluetooth wireless communication system

Initially Bluetooth wireless technology was created to solve a problem of replacing cables used for communication between such devices as: laptops, palmtops, personal digital assistant (PDA), cellular phones and other mobile devices [13][14] and [15]. Now Bluetooth enables users to connect to a wide range of computing and telecommunications devices without any need of connecting cables to the devices.

| Word | Meaning |
|---|---|
| Ahmar | Name of the robot 1 |
| Azrac | Name of the robot 2 |
| Akhdar | Name of the robot 3 |
| Aswad | Name of the robot 4 |
| Ammam | Go forward |
| Wara | Go Backward |
| Yamine | Left turn |
| Yassar | Right turn |
| Kif | Stop the movement, |
| lasa | Starting keyword |

Table 1. The Meaning Of The Vocabulary Voice Commands, Assigned Code.

Bluetooth is a radio frequency specification for short range, point-to-multipoint voice and data transfer. It operates in the 2.4 GHz ISM (Industrial- Scientific-Medical) band. This band is free for use, so it is not necessary to have special license for communicating in this frequency range. Of course this range is full of other signals from different devices, so it should have special methods of preventing interference with other signal. Bluetooth uses the frequency hopping (FH) technology and it avoids from the interference. It is also the reason, why Bluetooth is very secure protocol. Bluetooth is based on a low-cost, short-range radio link and enables communication via ad hoc networks. Main features of the Bluetooth communication protocol are [13]:

- nominal link range is 10 m, it can be extended to 100 m by increasing transmit power;
- connection is created every time, it is needed by using the ad hoc networks;
- basic unit in Bluetooth networks is a piconet, it supports up to 8 devices (1 master device and up to 7 slave devices);
- one Bluetooth device can be a part of different piconets, they can exist simultaneously;
- possible transmission through solid, non-metal objects;
- built-in methods of security and preventing interferences;

Bluetooth allows easy integration of TCP/IP for networking; Despite of the short range, Bluetooth protocol has also a lot of advantages that are very important in designed type of system [14]-[15].

### 3.5 Special keyboard

The special keyboard is composed of 8 switches , four interrupters for robot selection and four push buttons for direction of the selected robot, more over the navigation can be directed by a joystick once the robot is selected.

### IV. PRACTICAL DETAILS OF SLAVE PART

The slave agent based on a Bluetooth module from Pob-Technology connected to the main part of the robot POB-eye The vocabulary to be recognized by the system and their meanings are listed as in Table 1. It is obvious that within these words, some are object names and other are command names. The code to be received is composed of 1 bytes, four most significant bits are used to code the robot name and the four least significant bits are used to code the command to be executed by the selected robot. Example: "Lasa ahmar ammem tabek" which means "Lasa Red robot Go Forward execute"

A Graphic User Interface was used to display the real-time commands and the movements of Robots in the space work.
However, a simulation card was designed to control the set of four robots directly by MATLAB software. It is based on a buffer 74LS245, a 74LS138 3 to 8 decoder. The figure 3.a shows one of the robots POB-BOT in the work space avoiding obstacle.

Figure. 3 Mobile robot POB-BOT in space work navigation   and avoiding obstacle  via voice command

### Bluetooth

A Bluetooth  module is placed on Serial port of the POB-eye module to get wireless commands from Master agent. This Pob-Tooth with low consumption of energy is connected to serial port of the robot and can be controlled by any device with Bluetooth or JAVA program.

### POB-BOT Presentation

As illustrated in the following figure 3.b , the POB-BOT is composed of three main parts, POB-EYE based on an ARM 07 Microprocessor with a Monochrome Camera, A POB-LCD in which the user can get messages from the robot and POB-Proto a card integrating a PIC16F877 microcontroller to process inputs from sensors and outputs to navigation motors. This robot is programmed using C language and it is provided with en a development tools software called Pob-Tools [ 18 ]

## V. DETAILS ON SPEECH RECOGNITION MODULE

The development of the application for the VR-Stamp needs to follow some steps:

First the user creates the vocabulary words using  any voice record software, however we recommend the use of windows media play.

The recorded words should be compressed using quick synthesiser 4 (QS4) from sensory and built, the qs4 will produce an adapted file 'qs4/HPWC_voice.h' to be included in the C program of the application.

Some libraries from Fluent Chip software of Sensory have to be included also namely techlib.h.
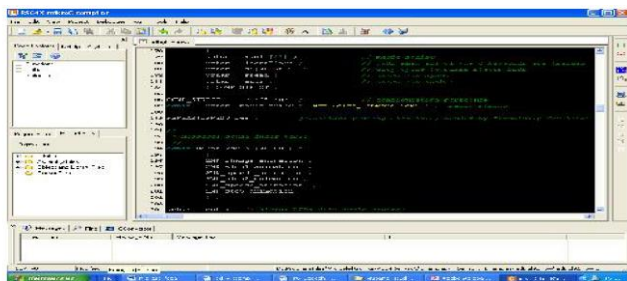


Figure. 4.bMain windows of RSC4 mikroC Development tool.

The main program is developed in C language on RSC4 mikroC Compiler. And then the produced Hex file is programmed on the VR-Stamp. As shown if the following figures 4.

## VI.  DESCRIPTION OF THE APPLICATION

The application involves the recognition of spotted words from a limited vocabulary ( 10 words) divided into two groups as shown in table 1. a set of word for robots name( Ahmar, Azrak, Akhdar, Aswad) and a set of actions to be donne by these robots( forward, backward, left right, stop) and two keywords at the beginning and end of the sentence ( Lasa, Tabek). The vocabulary specifications are 5 commands that are necessary to control the direction of the robots: Left, right, forward and backward and stop. The number of words in the vocabulary was kept to a minimum both to make the application simpler and to make it easier for the user to use. However, this number can be increased if any improvement is necessary such as adding words to control the robot displacement in the environment such as rotation in some degrees or full rotation (1800) and to add more robots in the arena.

In order to control the movements robots safely within the traced workspace and comfortably by voice commands, a set of sensors were integrated on robots in order to avoid obstacles and avoid collision with other robots in movement. For security reasons, the client can control the robots by the special keyboard .

External noise affects the system since it is by nature in movement. In designing the application, account was taken to reduce the affecting noise on the system at various movements. To do so, the external noise was recorded and spectral analysis was performed to study how to limit its effects in the recognition phase. However this is just done within the experience area and implemented on the DSP TMS320C6711.

The vocal command system works in two phases: The training phase and the recognition phase or verification phase. In the training phase, the operator will be asked to pronounce 'say' command words one by one. During this phase, the operator might be asked to repeat a word many times, especially if the word pronunciation is quite different from time to time. Once the 13 words have been used for training the system, the operator can start the second phase. The recognition phase represents the use of the system. In this phase, the system will be in a waiting state, whenever a sentence is detected.

The acquisition step will be activated, the DSP processor will eliminate echoes and filters the sentence then present the words of the sentence to the VR- Stamp, and then the parameters of each word are extracted and compared to those of reference words. If there is any matching between a reference word and the user word, the likelihood rate is high, and then the appropriate code will be generated. The PIC microcontroller will get the set of codes representing the command sentence and starting with 0, it processes the sequence of codes if it is right then it is sent to the server via the wireless communication system Bluetooth as shown in the flowchart Figure5.

The joystick or special keyboard is used to avoid any misleading in vocal command, hens this input device has higher priority than voice command.

## VII. RESULTS OF THE SIMULATION

The PIC 18F252 program was simulated by MPLAB which is a Windows-based Integrated Development Environment (IDE) for the Microchip, under windows XP. In the training phase, the speaker repeats two times each word to construct the database of referenced words. In the recognition phase, the application gets the words to be processed, treats them, then takes a decision if the sentence begins with the word "Lasa" followed by a robot name then a corresponding movement for that robot and finally finishes with the word "Tabek".
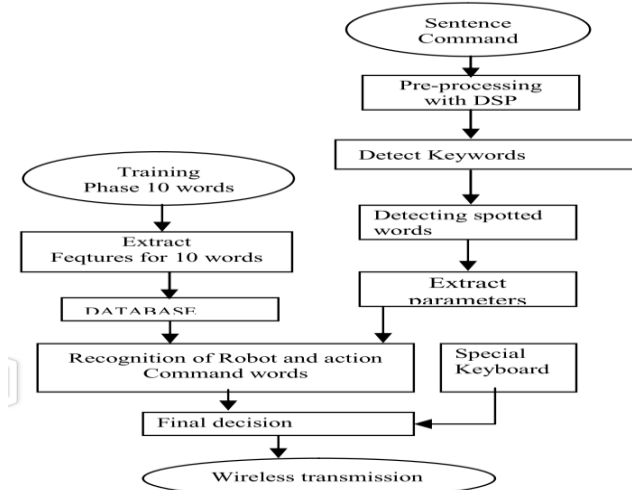


Figure.5. System phases flowchart.

Then the two input devices ( voice command and special keyboard) were tested and results are presented in figure 6. the column indicates the overage recognition rate for the robot names and actions taken by the robots within the space work, it is clear that special keyboard is more accurate than joystick or voice command words. The special keyboard has a robust base thus it is well controlled by the operator. Joystick on the other hand, is subject to vibration of operator arm, witch is a source of errors in guiding mobile robots. The word recognition module gives acceptable results in direct control of robots, however, in tele- operation some errors are produced due to word confusion and transmission line perturbation. Since special keyboard is connected with position control and joystick with rate control. Similarity between them proves that hand movement are more efficient than voice commands for tele-operation. However, when the hand is occupied by other control, the voice command can replace it.
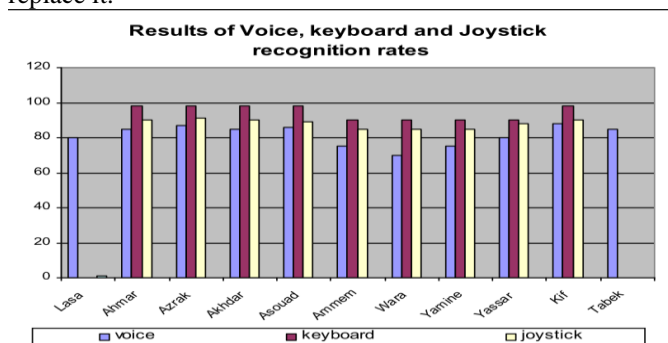


Figure.6. Results of speech recognition system and keyboard control system

## CONCLUSION

In this article, a software design of a master-slave system based on voice command for a colony of robots and a hardware conception of a special portable voice command system for a tele operation the colony of robots are presented. The bulky and complex designs have, however, been overcome by exploring new speech recognition kit. Interfacing this special vocal microprocessor to the robots was controlled by the PIC18F252 and a wireless transmission system ( Bluetooth). Thus the program memory capacity is improved in order to design more complex controls, and no need to an AD and DA Converters, since they are already integrated within the VR-Stamp. The application might be used to enhance AGV in robotics or other type of vocal command system. However, in order to increase the recognition rate, the training and recognition phase should be done in the same area of tests, which means that stationary noise has no effect on the recognition rate. In addition, More software work for this module are in development stage such as image feedback from the work space and speaker identification in the master part[17].

## REFERENCES

[1] J. Leonard and H. Durrant-Whyte, "Mobile robot localization by tracking geometric beacons," IEEE Trans. Robot. Automat., vol. 7, pp. 376–382, June 1991.

[2] I. Cox, "Blanche—An experiment in guidance and navigation of an autonomous robot vehicle," IEEE Trans. Robot. Automat., vol. 7, pp. 193–204, Apr. 1991.

[3] Mohamed FEZARI and Mounir Bousbia-Salah," Speech and Sensors in Guiding an Electric Wheelchair", AVT Journal( Automatic Control and Computer Sciences Journal), SpringerLink Publication, Vol . 41 no. 1, pp.39-43, Mai 2007.

[4] Ng, T.C., Ibañez-Guzmán, J., Shen, J., Gong, Z., Wang, H., Cheng, C., "Vehicle Following with Obstacle Avoidance Capabilities in Natural Environments," Proc. IEEE Int. Conf. on Robotics & Automation, (2004).

[5] Mohamed Fezari , Bousbia-Salah Mounir, " voice command system based on HMM fro a set of mobile robots", in rocedings of ACIT2007 , Amman Jordan.

[6] Sensory Data sheet for RSC-4128, Sensory Inc., 2005. Available at : www.sensory.com

[7] Texas Instruments, "TMS320C6000 CPU and Instruction Set", 2004. available at WWW.TI.com

[8] Pob-Bot robot datasheet, available at www.pob- technology.com 2011.

[9] M. D. Galanis, A. Papazacharias, E. Zigouris, "A DSP Course For Real Time Systems Design and Implementation Based On the TMS320C6211 DSK", Proc. of 14th IEEE Intl. Conf. on Digital Signal Processing (DSP2002), Santorini, Greece, pp. 853-856, vol. 2, 2002.

[10] Zheke Zhu; Shengh Zhu; Yuesong Lin; "Implementation of Pseudo-Linear Kalman Filter on DSP for two stations bearings-only target tracking", in Control and Decision Conference, 2008. CCDC 2008. china

[11] Gautier, M.; Khalil, W. & Restrepo, « Identification of the dynamic parameters of a closed loop robot", Proc. IEEE on Int. Conf. on Robotics and Automation, pp. 3045-3050, Nagoya, May 1995

[12] larson Mikael, (1999). "Speech Control for Robotic arm within rehabilitation". Master thesis, Division of Robotics, Dept of mechanical engineering Lund Unversity, 1999.