Embedded Multiprocessors Vocal Command system: a way to improve the Tele-Control of a didactic Manipulator Arm

Mohamed Fezari, Hamza Atoui , Al-Dahoud Ali *

University of Annaba, Laboratory of Automatic and Signals Annaba

** Al-Zaytoonah University Amman Jordan,Faculty of science and technology*

Mohamed.fezari@uwe.ac.uk, hamza.atoui@gmail.com, aldahoud@alzaytoonah.edu.jo

***Abstract*—A robot arm control project based on the implementation of New specialised processors is described in this work. The adopted design is based on grouping a microcontroller and a DSP processor for speech enhancement with a new voice recognition module for isolated word and speaker dependent. The resulting design is used to control a six degree of freedom manipulator arm that can be used on wheelchair of a severely handicapped individuals based on a vocal command. A way to gain in time design, experiments have shown that using Kits , is the best. More over, A DSP processor is integrated in order to enhance the quality of speech signal by reducing noise and echoes. The input of the system is a sentence of some spotted Arabic words used to control the objects and movement of a manipulator arm. The output is a corresponding command byte sent by Bluetooth module to the PC server that command the Robot arm..**

The system is developed in order to be portable. Therefore 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, Robot Arm, VR-Stamp processor and wireless transmission.

I. INTRODUCTION

The main requirement for a service robotin 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].

For these reasons, robots are very useful supporting tool in many fields such as chemical plants or nuclear power plants. In recent works, different techniques were introduced to improve the command of a Robot Arm such as joystick 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 or noise 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]. This is due to various studies made in the last few years which led to good results in both research and commercial applications. 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 vocal command of a robot arm TR-45, a didactic manipulator arm with five degree of freedom in laboratory LASA [11]-[14]. The study is part of a specific application concerning system control by simple vocal commands. It has to be implemented on a portable system and has to be robust to any background noise confronted by the system. 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.

II. General Description of The Designed Embedded 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 around 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 echos and background noises based on Kalman-filter.
- A microcpontroller PIC18F252 as a main processor.
- A special designed keyboard with eleven switches (two for each motor: one for rotation in one direction and the other for the other direction) and one for the stop command.
- 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.

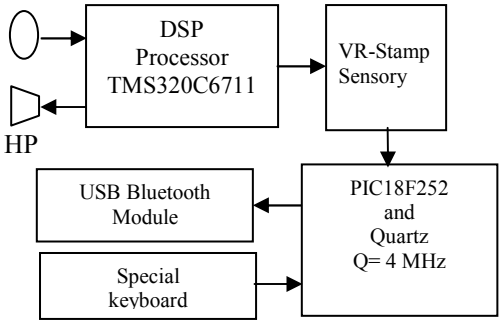

Fig.1.a Block diagram of the client system.

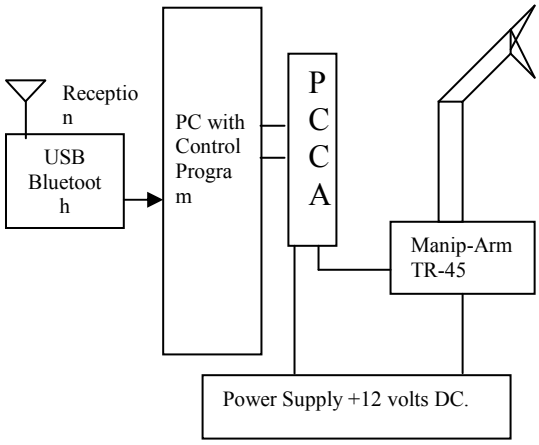

Fig 1.b Block diagram of server-system
PCCA: Power Circuit for Conrol ARM

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

A Personal Computer Pentium 2 GHz with 2 Gb DDRAM, with the program server that control the movements of the robot arm through a power circuit interface.

A power circuit interface that control the movements of the TR-45 five motors (base, upper limb , limp, wrist and gripper).

A Bluetooth 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 a accuracy, to be robust against noise and environment variations, but also to be very cost efficient.

III. Practical Details of client Part

The Client 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 echos. 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.

Fig. 2 Development Starter Kit of Texas Instrument

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 Flash program memory of 4 Mega-byte that holds the main program of word recognition, a reference word storage 24C65 of EEPROM type that holds the parameters of referenced word produced during the training phase, and a parallel interface of 24 lines (divided into 3 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 13 spotted words used in the vocabulary and presented in table 1, among these words, the starting words "yade" which is the name we gave to the manipulator arm TR-45, so whenever the user wants to submit a voice command the sentence should start with the word "yade" . 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, exaple: " yade Saad fawk please", in this sentence the module will submit the codes: 0, 3 and 8.

TABLE 1.

THE MEANING OF THE VOCABULARY VOICE COMMANDS, ASSIGNED CODE AND MOTOR CONTROLLED.

0) Yade (0)	Name of the manipulator
1) Assass	Base motor (M0)
2) Diraa (2)	Upper limb motor (M1)
3) Saad (3)	Limb motor (M2)
4) Meassam(4	Wrist (hand) motor (M3)
5) Mikbath(5)	Gripper motor (M4)
6) Yamine (1)	Left turn (M0)
7) Yassar (2)	Right turn (M0)
8) Fawk (3)	Up movement M1, M2 and M3
9) Tahta (4)	Down movement M1, M2 and M3
10) Iftah (5)	Open Grip, action on M4
11) Ighlak (6)	Close grip, action on M4
12) Kif (7)	Stop the movement, stops M0,M1 M2, M3 or M4

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. 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].

IV. Practical Details of Server Part

4.1. PC Interface

The client agent based on multiprocessor system will detects words within a phrase, and process each word. Depending on the probability of recognition of the object name and the command word a code will be transmitted to the PC via the wireless transmission system Bluetooth wireless module integrated in USB Port. 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 object name and the four least significant bits are used to code the command to be executed by the selected object. Example: "yade diraa fawk tabek" which means "Arm upper limb up execute".

A Graphic User Interface was used to display the real-time commands and the movements of Robot Arm. A hardware interface based on the following TTL IC (integrated circuits): a 74LS245 buffer, and a set of LEDs to indicate object and command transmitted.

However, a simulation card was designed to control the set of four stepper motor directly by MATLAB software. It is based on a buffer 74LS245, a 74LS138 3 to 8 decoder, and four power circuits for the motors as shown in figure 3.

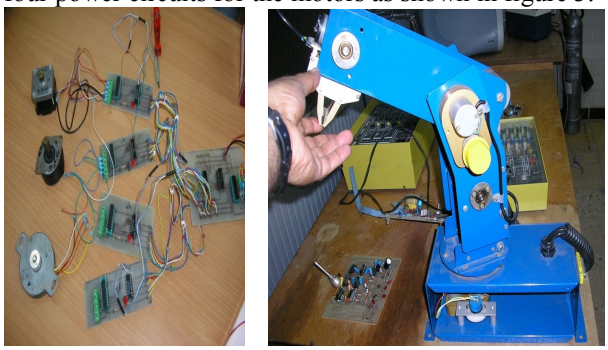

Fig. 3 Four stepper motors controlled by parallel port

and an overview for the robot-arm TR-45

4.2 Bluetooth

A Bluetooth module is placed on USB port to get wireless commands from client agent.

4.3 Power system module

A Power system module is installed to provide the necessary power for the manipulator arm motors.

V. Details on Speech Recognition Application

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 FluentChip software of Sensory have to be included also namely techlib.h.

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.a 4.b

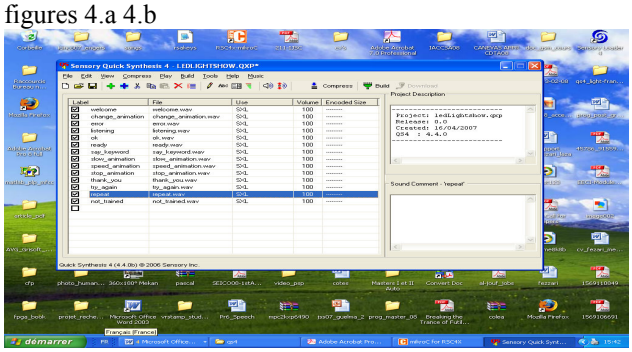

Fig. 4.a QS4 for vocabulary compression

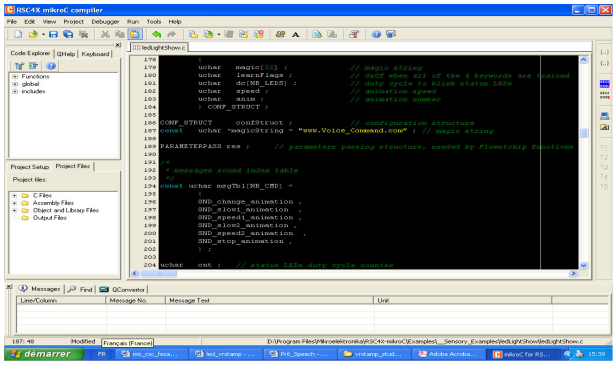

Fig. 4.b Main windows of RSC4 mikroC Development tool.

VI. Descrpton of The Application And Operation

The application involves the recognition of spotted words from a limited vocabulary (13 words) divided into two groups as shown in table 1. a set of object name(base, upper-limb, limb, wrist and gripper) and a set of object movements(up, down, left right, stop, open close) and a starting word (yade). The vocabulary specifications are 6 commands that are necessary to control the main parts of the robot arm: Left and right for the base and wrist , up down for the upper-limb and limb, up down left right for the wrist, and finally open close for the gripper. The words stop will stop al the movements. 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 arm displacement in the environment.

In order to control the movements of motors of robot arm safely and comfortably by vocal commands, a set of sensors were integrated on motors in order to cut power when attending

limits of movement. Also, the client can control the manipulator 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 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 words begin with "yade" followed by an object name then a corresponding action for that object name.

The filtering Kalman filter algorithm was developed using CCS Development tools provided by TI (Texas Instruments DSP unit) and implemented on TMS320C6711 SDK. It was fist simulated by Matlab tools and some results of filtering and classification for the words "kife" "yamine" and "Yassare" are presented in the following figures 6.a 6.b and 6.c

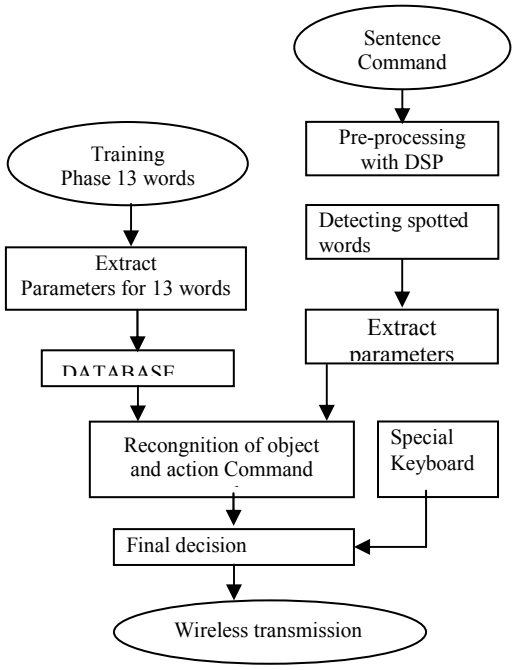

Fig.5. System phases flowchart.

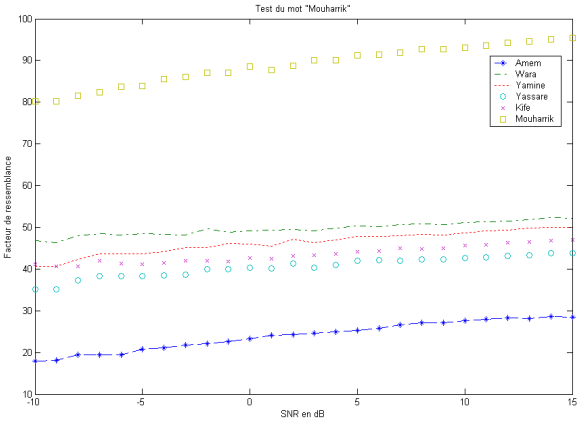

Fig 6.a the word "mouharek"

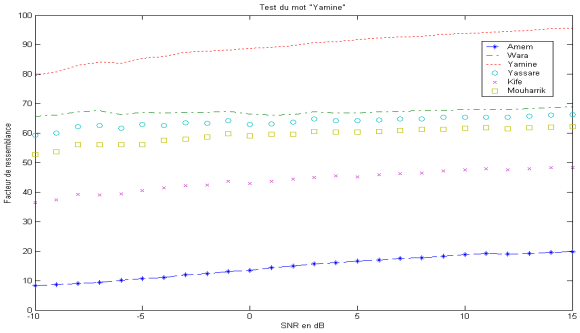

Fig 6.b the word "yamine"

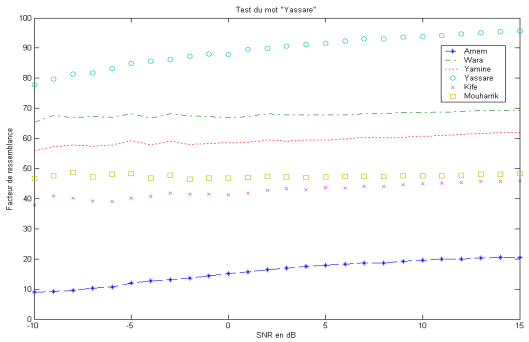

Fig 6.c the word "yassare"

It is obvious that the Kalman filter distinguish the noisy spotted word among the others and the system will present the clean corresponding word to the recognition system witch is the VR-Stamp processor. Then the two input devices (voice command and special keyboard) were tested and results are presented in figure 6.c the column indicates the overage recognition rate for the seven actions that can be taken by the objects of the robot arm TR-45, it is clear that special keyboard is more accurate than joystick or voce 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 the robot arm. The word recognition module gives acceptable results in direct control of robot arm, however, in teleoperation 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 teleoperation.

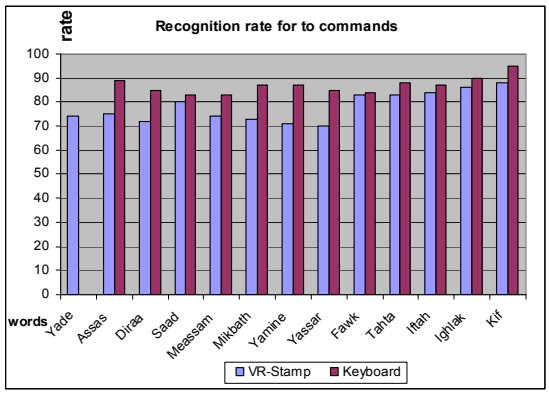

Fig.6.c Results of speech recognition system and keyboard control system

VIII. CONCLUSION

In this paper, a hardware design of a special portable vocal command system for a teleoperation of a manipulator arm is presented. The bulky and complex designs have, however, been overcome by exploring new speech recognition kit. Interfacing this special vocal microprocessor to the robot arm TR-45 was controlled by the PIC18F252 and a wireless transmission system. 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 server and speaker identification in the client agent.

Bitmap design to model the two dimensional plane where the robot arm is being used to facilitate the movements [17].

REFERENCES

[1] Fong T., Conti F., Grange S. and Baur C (2000). "Novel Interfaces for Remote Driving: Gesture, Haptic and PDA", SPIE Telemanipulator and Telepresence Technologies VII, Boston, MA, November 2000

[2] 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.

[3] Halme, Aarne; Leppänen, Ilkka; Suomela, Jussi; Ylönen, Sami; Kettunen, Ilkka, "WorkPartner: Interactive Human-like Service Robot for Outdoor Applications", The International Journal of Robotics Research , 2003. Vol. 22, nro 7-8, pp. 627-640

[4] Djemili R. , Bedda M. , Bourouba H. Recognition Of Spoken Arabic Digits Using Neural Predictive Hidden Markov Models. International Arab Journal on Information Technology, IAJIT, 2004, Vol.1, N°2, pp. 226-233.

[5] Gu L. and Rose K.,(2001). Perceptual Harmonic Cepstral Coefficients for Speech Recognition in Noisy Environment. Proc ICASSP 2001, 2001.

[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] 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.

[9] 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

[10] Gautier, M.; Khalil, W. & Restrepo, P. P. (1995). 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

[11] Halme A., Sievilä J., Kauppi I., Ylönen S., "Performing skilled work with an interactively operated service robot", In S.Yuta et all (Eds), Field and Service Robotics – Recent, Advances in Research and Applications, Springer, 2006 .

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

[13] "Bluetooth protocol architecture", White Paper, Bluetooth SIG, 1999.

[14] "Bluetooth technology", White Paper, Compaq Computer Corporation, Nov. 2000.

[15] B. A. Miller and Ch. Bisdikian, Bluetooth Revealed: The Insider's Guide to an Open Specification for Global Wireless Communications. Prentice Hall PTR, 2002.

[16] Kim Won S., T. Franck, E. Stephen and S. Lawrence," A Comprson of Position and Rate Control for Telemanipulator Syetems Dynamics", IEEE Journal of Robotics an Automation Vol.3, N. 5, 1987.

[17] Microchip manual, PIC18F252 Application notes, Microcontrollers, Microchip Technology Inc., 2003. Available at : www.microchip.com .

[18]. A. A. Alshennawy and Ayman A. Aly, " Edge Detection in Digital Images Using Fuzzy logic technique", in proc. WASET, march 2009. vol. 39 pp. 185-193, 2009