Journal of Engineering and Applied Sciences 2 (1): 30-35 2007

Voice and Sensors Fusion for Guiding a Wheelchair

Mohamed Fezari
Department of Electronics, Faculty of Engineering, Laboratory of Automatic and Signal,
Badji Mokhtar University, Annaba Annaba, BP.12, Annaba, 23200, Algeria

Abstarct: A new study based on voice command and sensors is being developed to reduce the cognitive and physical requirements of operating a power wheelchair for people with ranging impairment that limit their access to power mobility. The adopted system is based on hybrid method that combines a set of classical speech recognition algorithms and set of four peer of ultrasonic sensors modules. The hybrid method is obtained by mixing zero crossing, extremes, dynamic time warping, energy and fundamental frequency. A weight vector, computed on test results, is used in decision section. To test the study on a real application, a PC interface was designed to control the movement of a wheelchair for a handicapped person by simple vocal messages. On the wheelchair, a set of ultrasonic sensors were installed. The main role of these sensors is to detect frontal obstacles and to enable or disable the left or right rotation of the wheelchair when it is in is Tests showed that the system needs some outputs follow a logic in line with the words uttered.

Key words: Speech recognition, hybrid methods, DTW, voice command

INTRODUCTION

It is important nowadays, to assist elderly people or people with special needs. Creating a sense of independence for challenged individuals consists in controlling equipments with voice. Autonomy in controlling a power wheelchair is an important aspect of self-esteem, confidence and self worth. Combining speech and sensors for security in driving a wheelchair by handicapped persons is the main work presented in this study. First, Speech recognition system is designed to increase the autonomy of a physically disabled individual. Researches made in the last few years on automatic speech recognition systems have given good results in laboratory applications[1-4]. However, this paper proposes the hybrid approach to the problem of the recognition for a limited vocabulary, using a set of traditional pattern recognition method followed by a balance vector computed from test results of classical. In this novel approach ASR (Automatic Speech Recognition), the best elements in each classical method are taken into account in order to increase the rate of recognition as outlined in[5-8]. 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.

Second, the improvement expected to reach is an intelligent system on a wheelchair. In this study a set of ultrasonic sensors were added in order to prevent and avoid frontal and lateral obstacles.

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 AGV navigation have started to import some of the existing techniques developed in AI for dealing with uncertain information.

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 recently. 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: Crossing Zero and Extremes (CZEXM), Linear Dynamic Time Warping (DTW), Linear Predictive Coefficient (LPC) parameters, Energy Segments (ES) and cepstral coefficients. The application uses ten Arabic words used in AVG command. To be easy to implement on a DSP (Digital Signal Processor)[9], a simple approach is a better choice[5-8]. However, it has to be robust to any background noise confronted by the system.

The application is speaker-dependent. 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.

As application, a vocal command for a Handicapped Person WheelChair (HPWC) is chosen. A wheelchair is an important vehicle for physically handicapped persons. However, for the injuries who suffer from spasms and paralysis of extremities, the joystick is a useless device as a manipulating tool. It therefore involves the use of voice or head to control the movement of the wheelchair. Voice command needs the recognition of isolated words from a limited vocabulary used in AGV (Automatic Guided Vehicle) system[10,11].

Moreover, for driving the wheelchair with security by handicapped persons some Ultrasonic sensors[12,14] were added to the designed system, four peer of ultrasonic sensors are used and controlled by a microcontroller. The microcontroller gives environment information to the PC which takes the decision to execute the vocal ordered movement or not.

APPLICATION DESCRIPTION

The application is based on the voice command of wheelchair for a handicapped person and the implementation of a set of sensors around the wheelchair.

The Handicapped Person Wheel Chair (HPWC) specifications are fixed to ten command words which are necessary to control the movement of the wheelchair: switching on and off the engine motor forward movement, backward movement, stop, turn left, turn right, light on or off and horn on or off. The number of words in the vocabulary was kept to a minimum both to make the application simpler and easier for the user.

The ten words in the Arabic vocabulary are the following:

- Mouharek: This switches the motor on or off at average speed.
- Amame: This makes the movement upward.
- Wara: This makes the movement backward.
- Kif: This command stops the movement.
- Yamine: This makes the turn right.
- Yessar: This makes the turn left.
- Sarian: Speed up
- Batian: Speed down
- Dhoo: Turn on or off the light
- Zammara: Turno on or off the horn.

The system is by nature in movement on the wheelchair. Thus it is affected by environment noise. Some conditions were 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[7]. This is just done within the experience area.

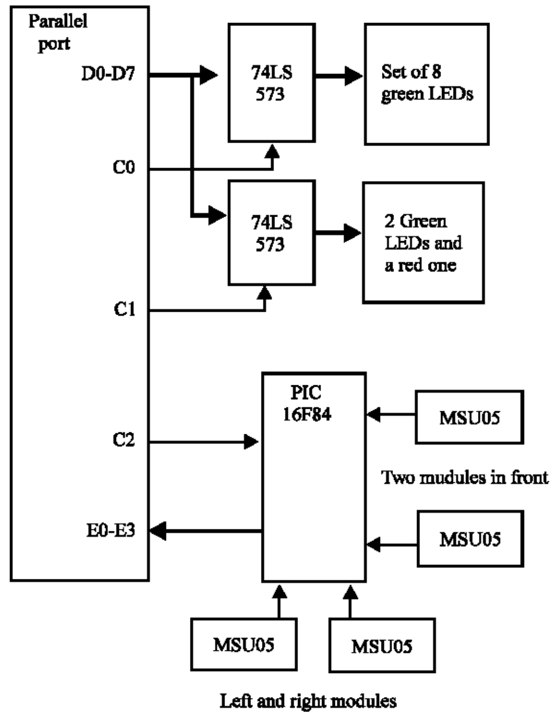

Fig. 1: Block diagram of hardware interface

Whenever, A voice command is given to the system, It recognise the command using the hybrid method then it tests the set of sensors. If the command word amème is chosen and if there is no obstacle in front then the command is executed. If the command word is yamine or yassare then the system computes the distance between the wall or obstacle and the wheelchair on left or right side. If the distance is greater than the limits then the action is allowed as shown in Fig. 1.

The application is first simulated on PC. It includes two phases: the training phase, where a reference pattern file is created 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 to show what command is taken. Moreover, on this parallel port are connected the set of four ultrasonic sensor modules type MSU05 from Lextronique[15].

THE HYBRID SPEECH RECOGNITION AGENT

The speech recognition system is based on a traditional pattern recognition approach. The main blocks are shown in Fig. 2.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

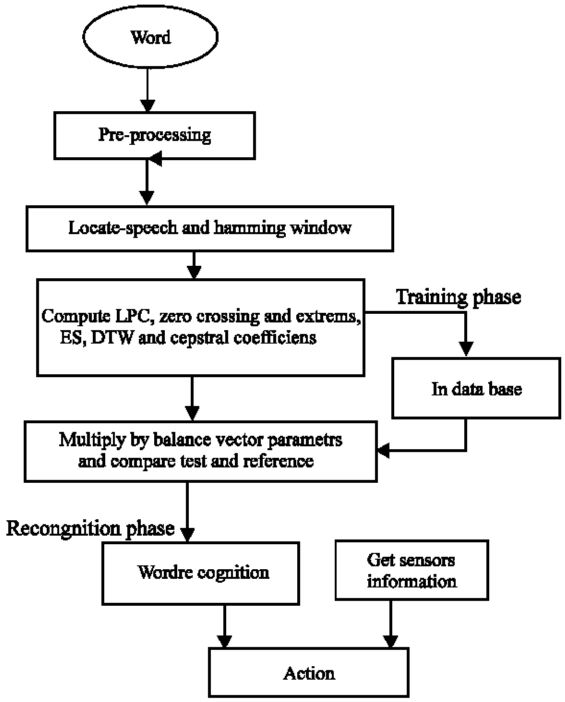

Fig. 2. Designed system main blocks

and minimize the effects of the background noise produced by the external conditions and the engine of the wheelchair.

The sound of the engine is recorded and then it is suppressed from the composite signal which includes the command word and motor sound. This makes it possible to increase the signal-to-noise ratio, with obvious benefits for the application.

The speech-locate and Hamming window block detects the beginning and end of the word pronounced by the user, thus eliminating silence and divide the word into segments. It processes the samples of the filtered input waveform, selecting useful information and filtering any noise generated by engine. Its output is a vector of samples of the word (i.e., those included between the endpoints detected).

The speech-locate procedure is based on detection and 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 10-mili seconds frames, at 100 samples per frame. Five types of parameters are extracted: Normalized Extremes Rate with Normalized Zero Crossing Rate (CZEXM), [linear DTW with Euclidian distance (DTWE), LPC coefficients [Ai), Energy segments (ES) and cepstral parameters [Ci)[5].

These parameters were chosen for computational simplicity reasons (CZEXM), robustness to background noise (10 Cepstral and Parcor parameters) and robustness to speaker rhythm variation (DTWE). 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, because they are obtained from the autocorrelation coefficients previously calculated by the Parcor parameters.

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 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 five methods and a weighting vector.

HYBRID RECOGNITION SYSTEM

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 3 shows the elements making up the main blocks for the Hybrid Recognition System (HRS). The input of HRS block is a set of five values representing the parameters of the input word obtained from the five methods. The Hybrid Recognition block compares the input parameters with the references parameters of the six words. It then generates a vector of five values; the elements are the recognized word number. If the input word is amame and the CZEXM method recognizes the word then it generates the number 2 (which means de second 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 or a Zero.

The last block in the HRS is the weighting Vector block. It is made of a threshold system that 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 five rates. These rates have been fixed based on results given by each method. The best word to fit the input word is chosen.

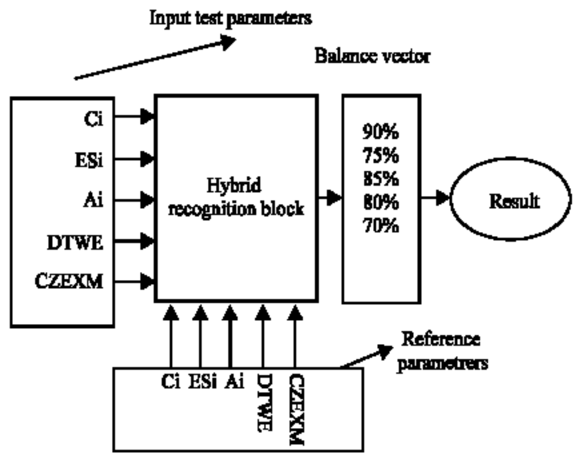

Fig. 3: HRS (Hybrid recognition system) block structure

INTERFACE CARD FOR IMULATION

A parallel port interface was designed to show the real-time commands. It is based on two TTL 74HCT573 Latches and 16 Light Emitting Diodes (LED), ten green LED to indicate each recognized command (Engine ON/OFF, Forward, Backward, Left, Right, Stop, Faster, Slower, light and horn) respectively and a red LED to indicate wrong or no recognized word. The other LED's were added for future insertion of new command word in the vocabulary. The set of four MSU05 modules are controlled by a microcontroller PIF16F84[17]. The information collected from the sensors by this microcontroller is forwarded to the PC via parallel port status register as shown in Fig. 3. The application goes through tree phases:

Training phase: The training phase is the first step in which the database is created. In this phase, the speaker repeats ten times each word, the parameters are extracted by some algorithms. For each vocabulary word, a vector of parameters is produced and saved in a file representing the Data base.

Recognition phase: 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 corresponding box on the GUI is changed to red colour.

Action phase: In this phase, the system gathers the information about the environment from the set of sensors and recognised command word and then it sends the appropriate commands to the actuators.

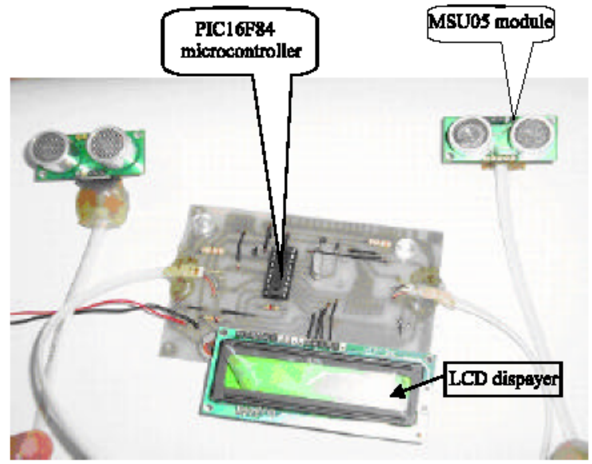

Fig. 4: Left and right position MSU05 Modules and the microcontroller

SENSORS SYSTEM

The set of ultrasonic sensors are controlled by a microcontroller the PIC16F84[17] manufactured by Microchip. Two peer of Ultrasonic sensors modules MSU05. The MSU05 is a module based on a microcontroller, an ultrasonic transmitter and an ultrasonic receiver sensors of 40 KHz type MURATA. The module is activated by a brief pulse, it sends a some 40KHz pulses[15-16]. The pulses reach an obstacle and then come back. The module computes the travel time of the pulses; it then generates a pulse with a width proportional to the distance from the obstacle as shown in Fig. 4.

Tow modules are placed in front of the wheelchair in such a way, that the area between 2-3 m in front of the vehicle is covered. And two modules are placed in left and right side in order to control the left and right turn of the wheelchair. If there is any obstacle that leads to a lock situation in a left or right turn then system predicts this study and avoids taking an action in these study.

RESULTS

First, some tests were done on each method and the rate of recognition was registered as shown in Fig. 5. In the CZEXM method, the rate is lower; however it gives better the rate for one speaker training, it is speaker dependant. DTWE gives better results specially if there is any distortion in the speaker rhythm. LPC, Energy and Cepstral coefficients give better rates than those cited earlier. From the tests results for each method separately, the balance vector is computed. Secondly, tests were made using the hybrid method. The recognition rate is

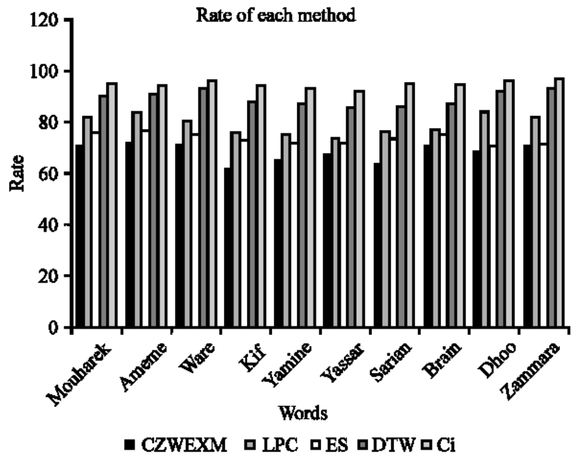

Fig. 5: Recognition rate for each method

improved. Finally, some test were done on the set of ultrasonic modules, the error is less than two cm for an obstacle distance of 2-3 m.

CONCLUSION AND FUTUR WORK

In this study a fusion information method to control a wheelchair by a severely handicapped person is presented and designed. A methodological approach to the implementation of isolated word recognition system based on a mixture of techniques used in speech recognition was adopted. This can be implemented easily within a specialised microcontroller, the CMOS RISC microcontroller as example the PIC18F4555[17].

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. 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 and the system need to be portable, thus not bulky and easy to control. Advances have been made in the technology of smart wheelchairs during the development of the HPWC. Performance of the HPWC has demonstrated its potential as an effective approach to providing independent mobility to a wide range of users who can not independently operate a powered wheelchair system. This speech command system can be enhanced by eliminating other kind of outdoor environment noises. The future works in the improvement of the HPWC will allow for different operating levels ranging from simple obstacle avoidance to fully autonomous navigation. The reinforcement learning algorithms will be investigated. Additionally, the HPWC will provide a means for development and testing of shared control methods, where a human operator and machine share control of a system. 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 robot movements or tele-operation control

REFERENCES

1. Bourhis, G. and Y. Agostini, May, 1998. The vahm robotized wheelchair: System architecture and human-machine interaction. J. Intelligent and Robotic Systems, 22: 39-50.
2. Cooper Rory, A., 1999. Engineering manual and electric powered wheelchairs. Critical Reviews in Biomedical Engin., 27: 27-73.
3. Graf, B., 2001. Reactive Navigation of an Intelligent Robotic Walking Aid. In Proceeding of Roman'01, Romania, pp: 353-358.
4. Rao, R.S., K. Rose and A. Gersho, 1998. Deterministically Annealed Design of Speech Recognizers and Its Performance on Isolated Letters, Proceedings IEEE ICASSP'98, Paris, France, pp: 461-464.
5. Glotin, H. and F. Berthommier, 2000. Test of several external posterior weighting functions fur multiband full combination ASR, *In*: ICSLP 2000, pp: 333-336.
6. Beritelli, F., S. Casale and A. Cavallaro, Dec., 1998. A Robust Voice Activity Detector for Wireless Communications Using Soft Computing, IEEE Journal on Selected Areas in Communications (JSAC), special Issue on Signal Processing for Wireless Communications, 16: 81-86.
7. Niles, L. and H. Silverman, 1990. Combining hidden markov models and neural network classifiers. In: CASSP'90, pp: 417-420.
8. Hagen, A., A. Morris and H. Bourlard, 1990. Different Weighting Schemes in the Full Combination Sub-Bands Approach in Noise Robust ASR. In: Proceedings of the ESCA Workshop on Robust Methods for Speech Recognition in Adverse Conditions, pp: 199-202.
9. Reynods D.A., R.C. Rose and M.J.T. Smith, 1992. PC Based TMS230C30 Implementation of the Gaussian Mixture Model Text-Independent Speaker Recognition System, ICSPAT, pp: 93-99.

10. Fezari, M., M. Bousbia-Salah and M. Bedda, 2006. Hybridl technique to enhance voice command for an electric wheelchair, in Asian J. Inform. Tech., 5: 139-144.
11. Bergasa, L.M., M. Mazo, A. Gardel, M.A. Sotero and J.C. Garcia, 1999. Guidance of a Wheelchair for Handicapped People by Head Movements. In Proc. 7th Intl. conference on Emerging Technologies and factory Automation (ETFA'99), Barcelona, Spain, pp: 105-111.
12. Hans M., B. Graf and R.D. Schraft, 2002. Robotics Home Assistant Care-O-bot: Past-Present-Future. In Proceeding of ROMAN'02, Romania, pp: 120-128.
13. Mazo, M., F. Rodriguez, J. Lazaro, J. and Al., 1995. Wheelchair for physically disabled people with voice, ultrasonic and infrared sensor control, Autonomous Robots, 2: 203-224.
14. Fezari, M., A. Redjati, M. Bedda and M. Bousbia-Salah, Dec., 2005. A Voice Command System for Robot Guidance. In Proceedings of the International Workshop on Text, Image and Speech Recognition, TISR'05, Annaba, Algeria, pp: 12-18.
15. Schilling, K., H. Roth, R. Lieb and H. Stuzle, 2004. Sensors to Improve the Safety for Wheelchair Users, EROS'04, proceedinds Symposium on Robotics on Systems, Annecy, France, pp: 791-796.
16. Bousbia-Salah, M. and M. Fezari, 2005. An Electronic Travel Aid for Blind people. In the 17th IMACS Congress, Paris, France /July 11-15, 2005. pp: 173-183.
17. Data sheet PIC16F84 from Microchip inc. User's Manual, 1998. http://www.microchip.com/datasheet.