HMM/GMM Based Voice Command System, A Way to improve a wireless control for a Didactic Manipulator Arm TR45

Mohamed FEZARI, Hamza Attoui

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

Abstract

A speech control system for a didactic manipulator arm, the TR45, is designed as an agent in a tele-manipulator system command. Robust Hidden Markov Model (HMM) and Gaussian Mixture models (GMM) are applied in spotted words recognition system with Cepstral coefficients with energy and differentials as features. The HMM and GMM are used independently in automatic speech recognition agent to detect spotted words and recognize them.

A decision block will generate the appropriate command and send it to a parallel port of the PC (personal Computer). To implement the approach on a real-time application, a Personal Computer parallel port interface was designed to control the movement of robot motors with a wireless communication components5. The user can control the movements of robot arm using a normal speech containing spotted words.

Keywords: Human-machine interaction, HMM,GMM, Voice command, Robot arm and Robotics.

1. Introduction

Manipulator robots are used in industry to reduce or eliminate the need for humans to perform tasks in dangerous environments. Examples include space exploration, mining, and toxic waste cleanup.

However, the motion of articulated robot arms differs from the motion of the human arm. While robot joints have fewer degrees of freedom, they can move through greater angles. For example, the elbow of an articulated robot can bend up or down whereas a person can only bend their elbow in one direction with respect to the straight arm position. The motion of articulated robot arms differs from the motion of the human arm. While robot joints have fewer degrees of freedom, they can move through greater angles. For example, the elbow of an articulated robot can bend up or down whereas a person can only bend their elbow in one direction with respect to the straight arm position [1]-[3].

. There have been many research projects dealing with robot control and tele-operation of arm manipulators, among these projects, there are some projects that build intelligent systems [10-12]. Since we have seen human-like robots in science fiction movies such as in "I ROBOT" movie, making intelligent robots or intelligent systems became an obsession within the research group.

In addition, speech or voice command as Human-robot interface has a key role in many application fields and various studies made in the last few years have given good results in both research and commercial applications [3-4] and [11] just for speech recognition systems. This paper presents a new approach to the problem of the recognition of spotted words within a phrase, using statistical approaches based on HMM and GMM [5] and [7]. By combining the two methods, the system achieves considerable improvement in the recognition phase, thus facilitating the final decision and reducing the number of errors in decision taken by the voice command guided system.

Speech recognition systems constitute 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 has started to import some of the existing techniques developed in AI for dealing with uncertain information.

HMM is a robust technique developed all applied in pattern recognition. Very interesting results were obtained in isolated words speaker independent recognition system, especially in limited vocabulary. However, the rate of recognition is lower in continuous speaking system. The GMM is also a statistical model that has been used in speaker recognition and in isolated word recognition systems. The two techniques were experimented independently and then combined in order to increase the recognition rate. The approach proposed here is to design a system that gets specific words within a large or small phrase, process the selected words (Spots) and then execute an order [6][7]. As application, a set of four reduction motors were activated via a wireless designed system installed on a PC parallel port interface. The application uses a set of twelve commands in Arabic words, divided in two subsets one subset contains the names of main parts of a robot arm (arm, fore-arm, wrist (hand), and gripper), the second subset contains the actions that can be taken by one of the parts in subset one (left, right , up , down, stop, open and

close). A specific word, "yade" which means arm, is also used at the beginning of the phrase as a "password". The application should be implemented on a DSP or a microcontroller in the future in order to be autonomous [8] .

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

Voice command needs the recognition of spotted words from a limited vocabulary used in Automatic Guided Vehicle (AGV) system [13] and in manipulator arm control [14].

2. Application Description

The application is based on the voice command for a set of four reduction motors. It therefore involves the recognition of spotted words from a limited vocabulary used to recognise the part and the action t of a robot arm.

The vocabulary is limited to twelve words divided into two subsets: object name subset necessary to select the part of the robot arm to move and command subset necessary to control the movement of the arm example: turn left, turn right and stop for the base (shoulder), Open close and stop for the gripper. The number of words in the vocabulary was kept to a minimum both to make the application simpler and easier for the user.

The user selects the robot arm part by its name then gives the movement order on a microphone, connected to sound card of the PC. The user can give the order in a natural language phrase as example: "**Yade**, gripper open execute". A speech recognition agent based on HMM technique detects the spotted words within the phrase, recognises the main word "Yade" witch is used as a keyword in the phrase, it recognises the spotted words, then the system will generate a byte where the four most significant bits represent a code for the part of the robot arm and the four less significant bits represent the action to be taken by the robot arm. Finally, the byte is sent to the parallel port of the PC and then it is transmitted to the robots via a wireless transmission.

The application is first simulated on PC. It includes three phases: the training phase, where a reference pattern file is created, the recognition phase where the decision to generate an accurate action is taken and the appropriate code generation, where the system generates a code of 8 bits on parallel port. In this code, four higher bits are used to codify the object names and four lower bits are sued to codify the actions. The action is shown in real-time on parallel port interface card that includes a set four stepper motors to show what command is taken and a radio Frequency emitter.

3. The Speech Recognition Agent

The speech recognition agent is based on HMM. In this paragraph, a brief definition of HMM is presented and speech processing main blocks are explained.

However, a pre-requisite phase is necessary to process a data base composed of twelve vocabulary words repeated twenty times by fifty persons

(Twenty five male and twenty five female). So before starting in the creation of parameters, 50*20*12 "wav" files are recoded in a repertory. Files from 35 speakers are saved on DB1 to be used for training and files from 15 are speakers are used for tests and then saved in DB2.

The training phase, each utterance (saved wav file) is converted to a Cepstral domain (MFCC features, energy, and first and second order deltas) which constitutes an observation sequence for the estimation of the HMM parameters associated to the respective word. The estimation is performed by optimisation of the likelihood of the training vectors corresponding to each word in the vocabulary. This optimisation is carried by the Baum-Welch algorithm [7].

3.1 HMM basics

A Hidden Markov Model (HMM) is a type of stochastic model appropriate for non stationary stochastic sequences, with statistical properties that undergo distinct random transitions among a set of different stationary processes. In other words, the HMM models a sequence of observations as a piecewise stationary process. Over the past years, Hidden Markov Models have been widely applied in several models like pattern [6], or speech recognition [6][9] . The HMMs are suitable for the classification from one or two dimensional signals and can be used when the information is incomplete or uncertain. To use a HMM, we need a training phase and a test phase. For the training stage, we usually work with the Baum-Welch algorithm to estimate the parameters (Π_i ,*A*,*B*) for the HMM [7, 9]. This method is based on the maximum likelihood criterion. To compute the most probable state sequence, the Viterbi algorithm is the most suitable.

A HMM model is basically a stochastic finite state automaton, which generates an observation string, that is, the sequence of observation vectors, $\boldsymbol{O}=\boldsymbol{O}_1,..\boldsymbol{O}_t$,… ,$\boldsymbol{O}_T$. Thus, a HMM model consists of a number of N states $S=\{S_i\}$ and of the observation string produced as a result of emitting

a vector $\boldsymbol{O}_t$ for each successive transitions from one state S_i to a state S_j. $\boldsymbol{O}_t$ is d dimension and in the discrete case takes its values in a library of M symbols.
The state transition probability distribution between state S_i to S_j is $A=\{a_{ij}\}$, and the observation probability distribution of emitting any vector $\boldsymbol{O}_t$ at state Sj is given by $B=\{b_j(\boldsymbol{O}_t)\}$. The probability distribution of initial state is $\Pi=\{\pi_i\}$.

$$a_{ij} = P(q_{t+1} = S_j / q_t = S_i) \quad (1)$$

$$B=\{b_j(\boldsymbol{O}_t)\} \quad (2)$$

$$p_i = P(q_0 = S_i) \quad (3)$$

Given an observation O and a HMM model $\lambda=(A,B,\Pi)$, the probability of the observed sequence by the forward-backward procedure $P(O/\lambda)$ can be computed [10]. Consequently, the forward variable is defined as the probability of the partial observation sequence $O_1O_2,....O_t$ (until time t) and the state S at time t, with the model λ as $\alpha(i)$. and the backward variable is defined as the probability of the partial observation sequence from t+1 to the end, given state S at time t and the model λ as $\beta(i)$. the probability of the observation sequence is computed as follow:

$$P(O/l) = \sum_{i=1}^{N} a_t(i) * b_t(i) = \sum_{i=1}^{N} a_T(i) \quad (4)$$

and the probability of being in state I at time t, given the observation sequence $\boldsymbol{O}$ and the model λ is computed as follow:

$$p_i = P(q_0 = S_i) \quad (5)$$

The flowchart of a connected HMM is an HMM with all the states linked altogether (every state can be reached from any state). The Bakis HMM is left to right transition HMM with a matrix transition defined as:

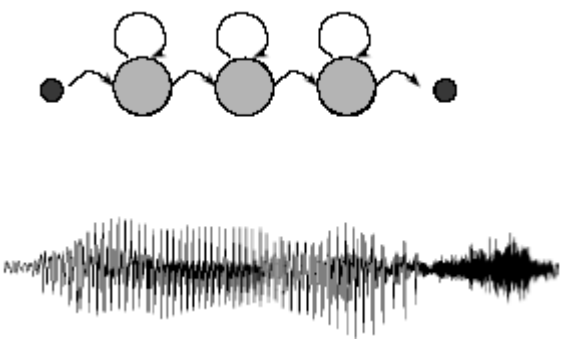


Fig. 1 : Presentation of left-right (Bakis) HMM

3.2 GMM Model basics

The GMM can be viewed as a hybrid between parametric and non- parametric density models. Like a parametric model, it has structure and parameters that control the behavior of density in known ways. Like non-parametric model it has many degrees of freedom to allow arbitrary density modeling. The GMM density is defined as weighted sum of Gaussian densities:

$$P_{G,M}(x) = \sum W_m g(x, m_m, C_m) \quad (6)$$

Here m is the Gaussian component (m=1…M), and M is the total number of Gaussian components. wm are the component probabilities ($\Sigma wm = 1$), also called weights. We consider K-dimensional densities so the argument is a vector x = (x1, ... , xK)T . The component pdf, g(x, μm, Cm), is a K-dimensional Gaussian probability density function (pdf). Equation (7)

$$g(x, m_m, C_m) = 1 * e^{-1/2(x-m_m)^T C_m^{-1}(x-m_m)} / (2\Pi)^{k/2} |C_m|^{1/2} \quad (7)$$

where is the mean vector, and Cm is the covariance matrix. Now, a Gaussian mixture model probability density function is completely defined by a parameter list given by θ = {w1, μ1, C1... w1, μ1, C1}m m=1…M
Organizing the data for input to the GMM is important since the components of GMM play a vital role in the making of word models. For this purpose, we use K- Means clustering technique to break the data into 256 cluster centroids. These centroids are then grouped into sets of 32 and then passed into each component of GMM. As a result we obtain a set of 8 components for GMM. Once the component inputs are decided, the GMM modelling can be implemented.

EM Algorithm

The expectation maximization (EM) algorithm is an iterative method for calculating maximum likelihood distribution parameter estimates from incomplete data (elements missing in feature vectors). The EM update equations are used which gives a procedure to iteratively maximize the log-likelihood of the training data given the model. The EM algorithm is a two step process:

$$y(m,t) = \frac{w_m^i g(x_t, m_m^i, C_m^i)}{\sum_{m=1..M} w_j^i g(x_t, m_j^i, C_j^i)} \quad (8)$$

Estimation Step in which current iteration values of the mixture are utilized to determine the values for the next iteration.
Maximization step in which the predicted values are then maximized to obtain the real values for the next iteration. Are presented in equations: 9, 10 and 11.

$$m_m^{i+1} = \frac{\sum_{t=1..T} y_{m,t} X_t}{\sum_{t=1..T} y_{m,t}} \quad (9)$$

$$W_m^{i+1} = \sum_{t=1..T} y_{m,t} \quad (10)$$

$$l_{m,j}^{i+1} = \frac{\sum_{t=1..T} y_{m,t} (x_{t,j} - m_{m,j}^{i+1})^2}{\sum_{t=1..T} y_{m,t}} \quad (11)$$

EM algorithm is well known and highly appreciated for its numerical stabilities under a threshold values of λmin. Using the final re-estimated w, μ and C the value of L_{GMM} is calculated with respect to all the word models available with the recognition engine as:

$$L_{GMM} = \frac{1}{T} \sum_{t=1..T} \log P_{GM}(X_t) \quad (12)$$

3.3 HMM/GMM model

The HMM/GMM hybrid model has the ability to find the joint maximum probability among all possible reference words W given the observation sequence O. In real case, the combination of the GMMs and the HMMs with a weighted coefficient may be a good scheme because of the difference in training methods. The ith word independent GMM produces likelihood LiGMM, I = 1, 2,…, W, where W is the number of words. The ith word independent HMM also produces likelihood LiHMM, I = 1, 2,…, W. All these likelihood values are passed to the so – called likelihood decision block, where they are transformed into the new combined likelihood L' (W):

$$L'(W) = (1 - x(W))LiGMM + x(W)LiHMM \quad (13)$$

where x(W) denotes a weighting coefficient.

The value of x is calculated during training of the Hybrid Model. In Hybrid Testing, the subset of training data is used and it's HMM & GMM likelihood values are calculated which are combined using weighing coefficient. Static values of weighted coefficient are also used in order to get higher recognition rate.

In that case the conception of 12 HMM models one per vocabulary word. And 12 GMM models, on for each word. The resulting of both models is taken by the decision block .

3.4 Speech processing phase

Once the phrase is acquired via a microphone and the PC sound card, the samples are stored in a wav file. Then the speech processing phase is activated. During this phase the signal (samples) goes through different steps: pre-emphasis, frame-blocking , windowing, feature extraction and MFCC analysis.

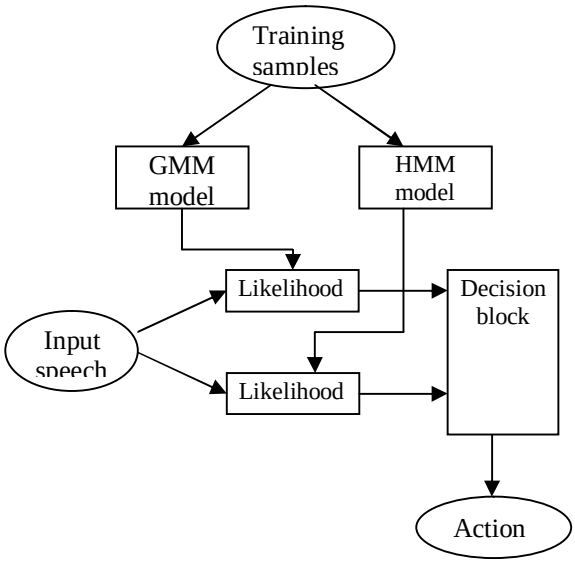


Fig. 2. speech recognition agent based on HMM/GMM model.

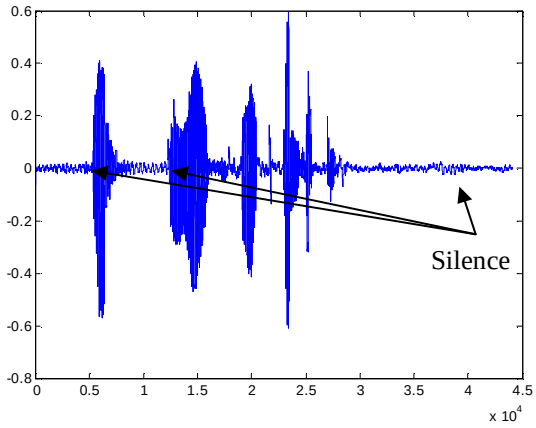


Fig.3. Phrase test "yade diraa fawk tabek" and silence at the beginning and at the end

a) Pre-emphasis step

In general, the digitized speech waveform has a high dynamic range. In order to reduce this range pre- emphasis is applied. By pre-emphasis [1], we imply the application of a high pass filter, which is usually a first - order FIR of the form $H(z) = 1 - a \times z^{-1}$. The pre-emphasis is implemented as a fixed- coefficient filter or as an adaptive one, where the coefficient a is adjusted with time according to the autocorrelation values of the speech. The pre-emphasis block has the effect of spectral flattening which renders the signal less susceptible to finite precision effects (such as overflow and underflow) in any subsequent processing of the signal. The selected value for a in our work is 0.9375.

b) Frame blocking

Since the vocal tract moves mechanically slowly, speech can be assumed to be a random process with slowly varying properties [1]. Hence, the speech is divided into overlapping frames of 20ms every 10ms. The speech signal is assumed to be stationary over each frame and this property will prove useful in the following steps.

c) Windowing

To minimize the discontinuity of a signal at the beginning and the end of each frame, we window each frame frames [1]. The windowing tapers the signal to zero at the beginning and end of each frame. A typical window is the Hamming window of the form:

$$W(n) = 0.54 - 0.46 * \cos\left(\frac{2\pi n}{N-1}\right) \quad 0 \le n \le N-1 \qquad (7)$$

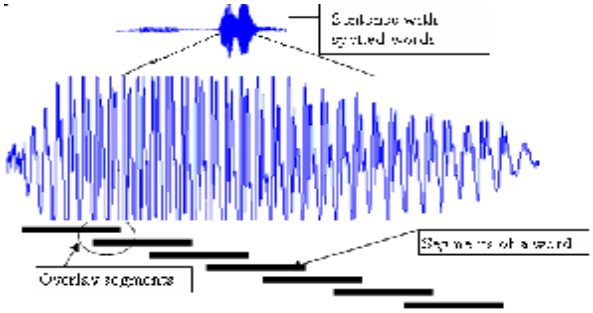


Fig. 4. Windowing

d) Feature extraction

In this step, speech signal is converted into stream of feature vectors coefficients which contain only that information about given utterance that is important for its correct recognition. An important property of feature extraction is the suppression of information irrelevant for correct classification, such as information about speaker (e.g. fundamental frequency) and information about transmission channel (e.g. characteristic of a microphone). The feature measurements of speech signals are typically extracted using one of the following spectral analysis techniques: MFCC Mel frequency filter bank analyzer, LPC analysis or discrete Fourier transform analysis. Currently the most popular features are Mel frequency Cepstral coefficients MFCC [7].

e) MFCC Analysis

The Mel-Filter Cepstral Coefficients are extracted from the speech signal . The speech signal is pre-emphasized, framed and then windowed, usually with a Hamming window. Mel-spaced filter banks are then utilized to get the Mel spectrum. The natural Logarithm is then taken to transform into the Cepstral domain and the Discrete Cosine Transform is finally computed to get the MFCCs. as shown in the block diagram of Figure 3.

$$C_k = \sum_{i=1}^{N} \log(E_i) * \cos\left[\frac{p\,k(i-1/2)}{N}\right] \qquad (8)$$

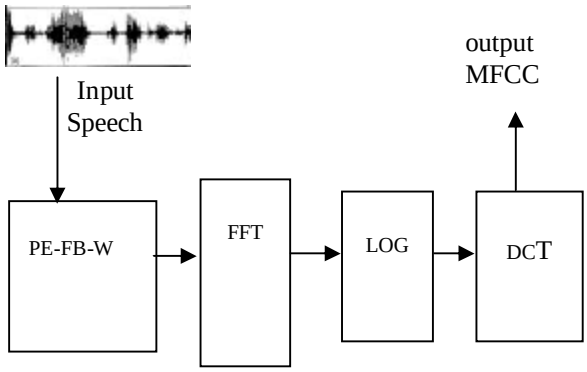


Fig. 5: MFCC block diagram

Where the acronyms signify:
-PE-FB-W: Pre-Emphasis, Frame Blocking and windowing.
- FFT: Fast Fourier Transform
- LOG: Natural Logarithm
- DCT: Discrete Cosine Transform

4. Parallel Interface Circuit

The speech recognition agent based on HMM will detect words, and process each word. Depending on the probability of recognition of the object name and the command word a code is transmitted to the parallel port of the PC. 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 transmitted is composed of 8 bits, four bits 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".

A parallel port interface was designed to display the real-time commands. It is based on the following TTL IC (integrated circuits): a 74LS245 buffer, a microcontroller PIC16F84 and a radio frequency transmitter from RADIOMETRIX TX433-10 (modulation frequency 433 Mhz and transmission rate 10 Kbs).

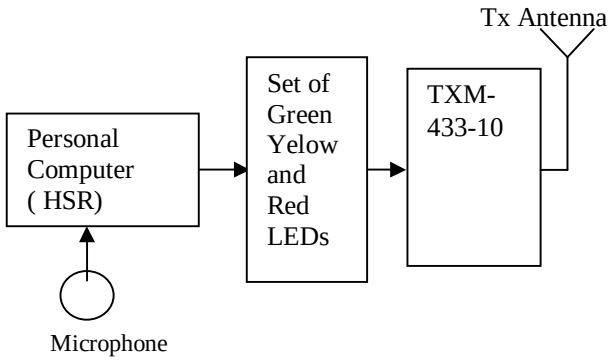
4.1 TR45 Manipulator Arm Description

The robot arm is composed of four feedback controlled movements for the elements: base, upper-limb, limb and wrest the movement command is realised by a moto-reductor block (1/500) powered by +12 and – 12 volts. The copy of voltage is given by a linear rotator potentiometer fixed on the moto-reductor block and powered by +10 and -10 volts.

One open loop controlled movement, for the gripper, with the same type of command.
Displacement Characteristics:
Base :290^0
Upper limb : 108^0
Lim : 280^0
Wrist : 380^0
Gripper : 100^0

TABLE 1.
THE MEANING OF THE VOCABULARY VOICE COMMANDS, ASSIGNED CODE AND CONTROLLED MOTOR.

1) Yade (1)	Name of the manipulator (keyword)
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



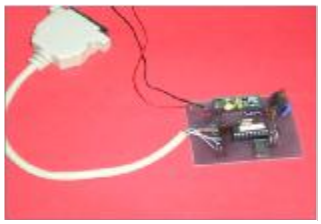


Fig. 6. A Parallel interface circuit and a photo of the designed card.

As in Figures 6.b and 6.c, the structure of the mechanical hardware and the computer board of the robot arm in this paper is similar to MANUS [10-12]. However, since the robot arm needs to perform simpler tasks than those in [13] do, the computer board of the robot arm consists of a PIC16F876 , with 8K-instruction EEPROM (Electrically Programmable Read Only Memory) , three timers and 3 ports [15], four power circuits to drive the reductor motors and one H bridges driver using BD134 and BD133 transistors for DC motor to control the gripper, a RF receiver module from RADIOMETRIX which is the SILRX-433-10 (modulation frequency 433MHz and transmission rate is 10 Kbs) [16 as shown in figure 6.b. Each motor in the robot arm performs the corresponding task to a received command (example: "yamin", "kif" or Fawk") as in Table 1. Commands and their corresponding tasks in autonomous robots may be changed in order to enhance or change the application. In the recognition phase, the speech recognition agent gets the sentence to be processed, treats the spotted words, then takes a decision by setting the corresponding bit on the parallel port data register and hence the corresponding LED is on. The code is also transmitted in serial mode to the TXM-433-10.

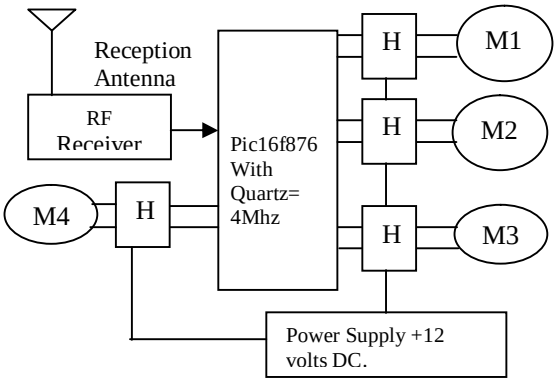


Fig. 6.b. Robot Arm block diagram

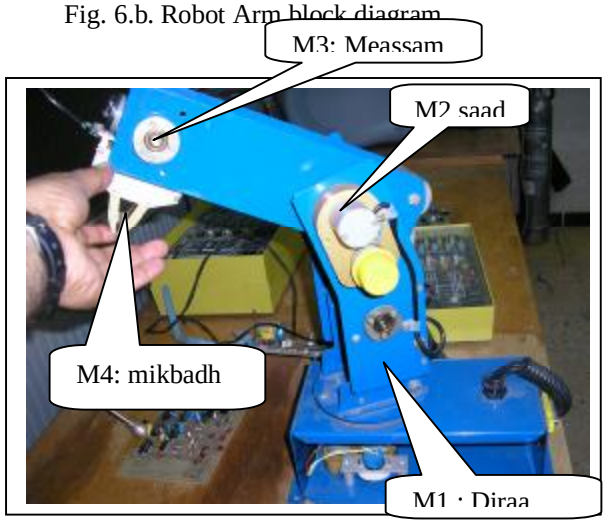


Fig. 6.c. Overview of the Robot arm and Parallel interface

5. Manipulator Arm Interface

6. Experiments on the System

The speech recognition agent is tested within the laboratory of L.A.S.A, there were two different conditions to be tested: of line, by using DB2 and on real-time. There are three types of testes : HMM and GMM models then the HMM/GMM model is testes on-line and in real-time. The results are presented in figures After testing the recognition of command and object words 100 times in the following conditions: a) off-line witch means test words are selected from DB1 b) on real-time witch means some users will command the system in real-time . the results are shown in figures 7.a and 7.b.

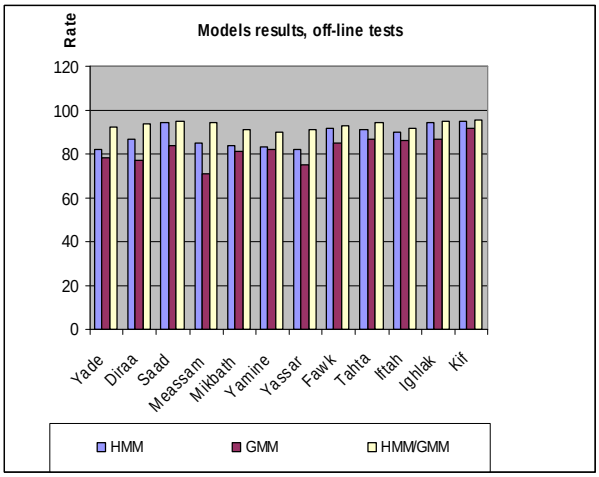


Fig. 7.a. HMM, GMM and HMM/GMM models results.

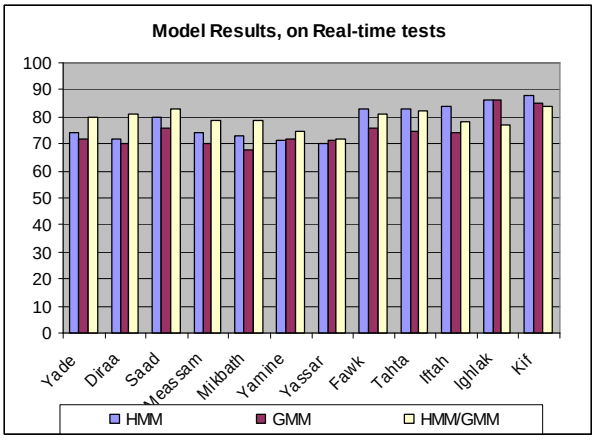


Fig. 7.b. HMM , GMM and HMM/GMM models results.

The effect of the environment is also taken into consideration and here the results on the HMM/GMM model just by changing the microphone, we notice that with the microphone Mic1 used in recording the data base we get better rate than using anew microphone Mic2.

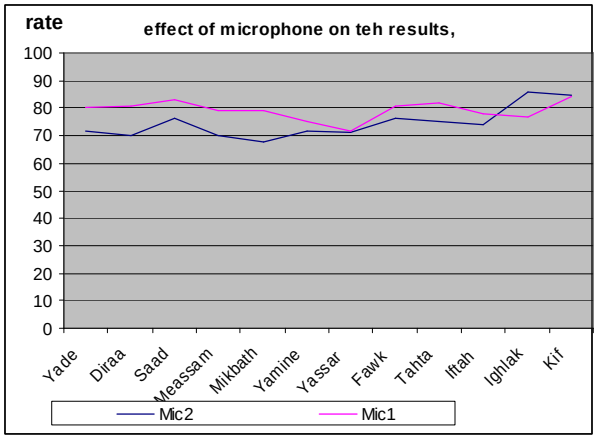


Fig. 7.c. Microphone effect on results

7. Conclusion and Future Work

A voice command system for robot arm is proposed and is implemented based on a hybrid model HMM/GMM for spotted words. The results of the tests show that a better recognition rate can be achieved using hybrid techniques and especially if the phonemes of the selected word for voice command are quite different. The effect of the used microphone for tests is proved in the results presented in 7.c. However, a good position of the microphone and additional filtering may enhance the recognition rate.

The HMM based model gives better results than GMM independently, by combining GMM and HMM and using as features MFCC and differentials we increased the recognition rate. The application is speaker independent. However by computing parameters based on speakers' pronunciation the system can be speaker dependant.

Spotted words detection is based on speech detection then processing of the detected. Once the parameters were computed, the idea can be implemented easily within a hybrid design using a DSP and a microcontroller since it does not need too much memory capacity. Finally, Since the designed electronic command for the robot arm consists of a microcontroller and other low-cost components namely wireless transmitters, the hardware design can easily be carried out. [18] [19].

References

[1] Jeppe Mikkelsen, "A Machine Vision System Controlling a Lynxarm Robot along a Path ," University of Cape Town, South Africa, October 28, 1998.

[2] Beritelli F., Casale S., Cavallaro A., 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*, Vol. 16, n. 9,1998.

[2] Bererton, C. & Khosla, P, Towards a team of robots with reconfiguration and repair capabilities, *Proceedings of the 2001 IEEE International Conference on Robotics and Automation*, pp. 2923– 2928. 2001.

[4] Rao R.S., Rose K. and Gersho A., Deterministically Annealed Design of Speech Recognizers and Its Performance on Isolated Letters, *Proceedings IEEE ICASSP'98*, pp. 461-464, 1998.

[5] Gu L. and Rose K, Perceptual Harmonic Cepstral Coefficients for Speech Recognition in Noisy Environment. *Proc ICASSP 2001*, Salt Lake City, 2001.

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

[7] Rabiner L. R. Rabiner. Tutorial on Hidden Markov Models and Selected Applications in Speech Recognition. *Readings in Speech Recognition, chapter A*, pp. 267-295, 1989.

[8] Hongyu L. Y. Zhao, Y. Dai and Z. Wang, , A secure Voice Communication System Based on DSP, *IEEE 8th International Conf. on Cont. Atau. Robotc and Vision, Kunming, China, 2004*, pp. 132-137, 2004.

[9] Ferrer M.A. , I. Alonso, C. Travieso, (2000). Influence of initialization and Stop Criteria on HMM based recognizers , *Electronics letters of IEE*, Vol. 36, pp.1165-1166, 2000.

[10] Kwee Hok, Intelligent control of Manus Wheelchair. *In proceedings Conference on Reabilitation Robotics, ICORR'97*, Bath 1997, pp. 91-94, 1997.

[11] Yussof, H.; Yamano, M.; Nasu, Y.; Mitobe, K. & Ohka M. Obstacle Avoidance in Groping Locomotion of a Humanoid Robot, *International Journal of Advanced Robotic Systems*, Vol. 2**, No. 3,** pp. 251 – 258, 2005.

[12] Buhler C., Heck H., Nedza J. and Schulte D, MANUS wheelchair-Mountable Manipulator-Further Devepolements and Tests. *Manus usergroup Magazine*, Vol. 2 ,no.1 , pp. 9-22, 1994, 1994.

[13] Heck Helmut, (1997). User Requirements for a personal Assistive Robot, *In proc. Of the 1st MobiNet symposium on Mobile Robotics Technology for Health Care Services, Athens*, pp. 121-124, 1997.

[14] E. Rodriguez, B. Ruiz, A. G. Crespo, F. Garcia. Speaker Recognition Using a HMM/GMM Hybrid Model". In: Proceedings of the First International Conference on Audio- and Video-Based Biometric Person Authentication, pp. 227- 234, April, 2003

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

[16] Data sheet PIC16F876 from Microchip inc. User's Manual, 2001, http://www.microchip.com.

[17] Radiometrix components, TXm-433 and SILRX-433 Manual, HF Electronics Company. http://www.radiometrix.com.

[18] Kim W. J., et al., Development of A voice remote control system. *In Proceedings of the 1998 Korea Automatic Control Conference, Pusan, Korea*, , pp. 1401-1404, 1998.

[19] Fezari M., M. Bousbia-Salah and M. Bedda, Hybrid technique to enhance voice command system for a wheelchair, *in proceedings of Arab Conference on Information Technology ACIT'05, Jordan*, 2005.

Authors' information

[1]Mohamed FEZARI,
Faculty of Engineering, Department of Electronics, University of Annaba, Laboratory of Automatic and Signals, Annaba, BP.12, Annaba, 23000, ALGERIA

[2]Hamza Attoui,
Faculty of Engineering, Department of Electronics, University of Annaba, Laboratory of Automatic and Signals, Annaba, BP.12, Annaba, 23000, ALGERIA

First A. Author: Mohamed Fezari is a Senior Lecturer in Electronics and computer architecture at the University of Badji Mokhtar Annaba, Algeria. He got a Bachelor degree in Electrical engineering from university of Oran, 1983. He received an MSc degree in Computer science from University of California Riverside, 1987. He holds PhD degree in Electronics from the University of Badji Mokhtar Annaba. Dr. Fezari has more than 20 years of experience in teaching and industrial projects development with private companies. He also leads and teaches at both BSc and MSc levels in computer science and Electronics. His research interests are: Speech processing, DSP, microcontroller, Robotics and Human machine interaction and rehabilitation.