***DC MOTOR SPEED CONTROL BASED ON SPEECH RECOGNITION USING LINEAR PREDICTIVE CODING (LPC) AND ADAPTIVE NEURO FUZZY INFERENCE SYSTEMS (ANFIS)***

Muhammad Akil1,2,Ingrid Nurtanio1, Rhiza Samsoe’oed Sadjad 1

1 Department of Electrical Engineering, Hasanuddin University, Makassar

2 Department of Mechatronic Engineering, Bosowa Politechnic, Makassar

e-mail: akilmuhammad1989@yahoo.co.id, *ingrid\_unhas@ yahoo.com*, *[rhiza@unhas.ac.id](mailto:rhiza@unhas.ac.id)*.

**Abstract – Speech recognition is an interesting research object, not only it can be used as a speech recognition function but also it can be used as a control. Firstly the sound which used as an instruction in control must be extract to get the feature. There are many feature extraction methods, but in this research uses Linier Predictive Coding (LPC). LPC method is used to get a characteristic based on the inputs of sound command such as “Nyala”, “Mati”, “Lambat”, “Sedang”, and “Cepat”. The databases of feature extraction are gotten by record the sound command repeatedly until 10 times from each the sound instruction. Then, the database is training using ANFIS with 50 epochs and getting the training error as 0,00050384 . The result of the first training data is showing between the training data plotting and the ANFIS output are very closer together, so ANFIS can plot the input and output well. The result of training data test after giving the sound command input is showing that the output of ANFIS is correct with the kind of sound instruction as given before. This output can be send to Arduino controller as an input from serial communication, so the input can be processed by controller to put out the PWM signal and transmit to DC motor driver. The examination of the system can expend PWM value as the sound command input.**

***Keywords --*** *Control, Speech Recognition, LPC, ANFIS, PWM.*

1. **INTRODUCTION**

Human voice is a very precious gift, with the voice, a person easily communicates with others so that the process of discussion or exchange ideas can be fulfilled. The sound is unique with different tones. In general, the sound formation process comes from forming airflow from the lungs and then modulated to produce a specific sound [1]. In the field of biometrics, the difference in sound can be used as a reference to identify a person based on the physiological and behavioral characteristics that the individual possesses [2].

Sound signals coming out of the throat can not be used as an input in the control system, because the signal is still shaped in audio wave according to the size of the tone level. Thus, it is necessary to study the features or characteristics of the signal. There are several methods of doing the voice signal feature extraction methods include LPC (Linear Predictive Coding) [2,3,5,7,8] and MFCC (Mel-Frequency Cepstrum Coeffisients) [8] followed by a couple of voice pattern recognition methods the extraction of which Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [6], Artificial Neural Network (ANN) [3] and others.

With the development of today's technology, speech recognition can be incorporated as a feature that can be formed using the methods of artificial intelligence in order to generate output that can be used to control something, for example the control of the movement of the robot arm in real-time using Neuro Fuzzy [2], controlling the movement of the robot car using ANN methods (Artificial Neural Network) [3], and as a cue to play the music player with Propagation Neural Network Feedback method [6].

Based on previous studies, researchers were able to conclude that the application of speech recognition systems are still focused on controlling the *On-Off* course, so we wanted to do research on the Control of DC Motor Speed Based on Speech Recognition Method Using Linear Predictive Coding (LPC) as an extraction features of voice signal and followed by a learning method using Adaptive Neuro-Fuzzy Inference Systems (ANFIS).

The purpose of this paper is (1) To control the DC motor speed using speech recognition with LPC and ANFIS method, (2) To determine differences in the characteristics of the extracted speech recognition using Linear Predictive Coding (LPC) and (3) to obtain the pattern of speech recognition by using Adaptive Neuro Fuzzy Inference System (ANFIS). This paper contains about speech recognition method using Linear Predictive Coding (LPC) and ANFIS as a voice characteristic pattern recognition, and also discussed the DC motor speed control using the technique of Pulse Width Modulation (PWM).

1. **LITERATURE RIVIEW**
2. **Feature Extraction Using** ***LPC*** ***(Linear Predictive Coding)***

Feature extraction in Automatic Speech Recognition  is the calculation sequence of feature vectors that provide representation of the sound signal which is given. Usually this is done in three stages [8]. The first stage is called the analysis of voice which analysis spectra from the speech sound and produce raw features that describe the spectrum of short greeting in overall interval. The second stage compiles an extended feature vector consisting of static and dynamic features. The final step alters the feature vector that has been expanded into several complete and powerful vectors that will be supplied to the identifier.

The LPC method used to extract the features from the original voice with good quality and efsien to be used in the calculation [1]. Linear Predictive Coding (LPC) analysis with estimating formant, separating the formant of the signal, which is called inverse filtering process, and then estimate the intensity and frequency of the speech signal left, which is called residue. Because the speech signals vary over time, these estimates are performed for every small piece of signal, it is called a frame. The procedure for obtaining LPC coefficients are shown in Figure 1 [9].

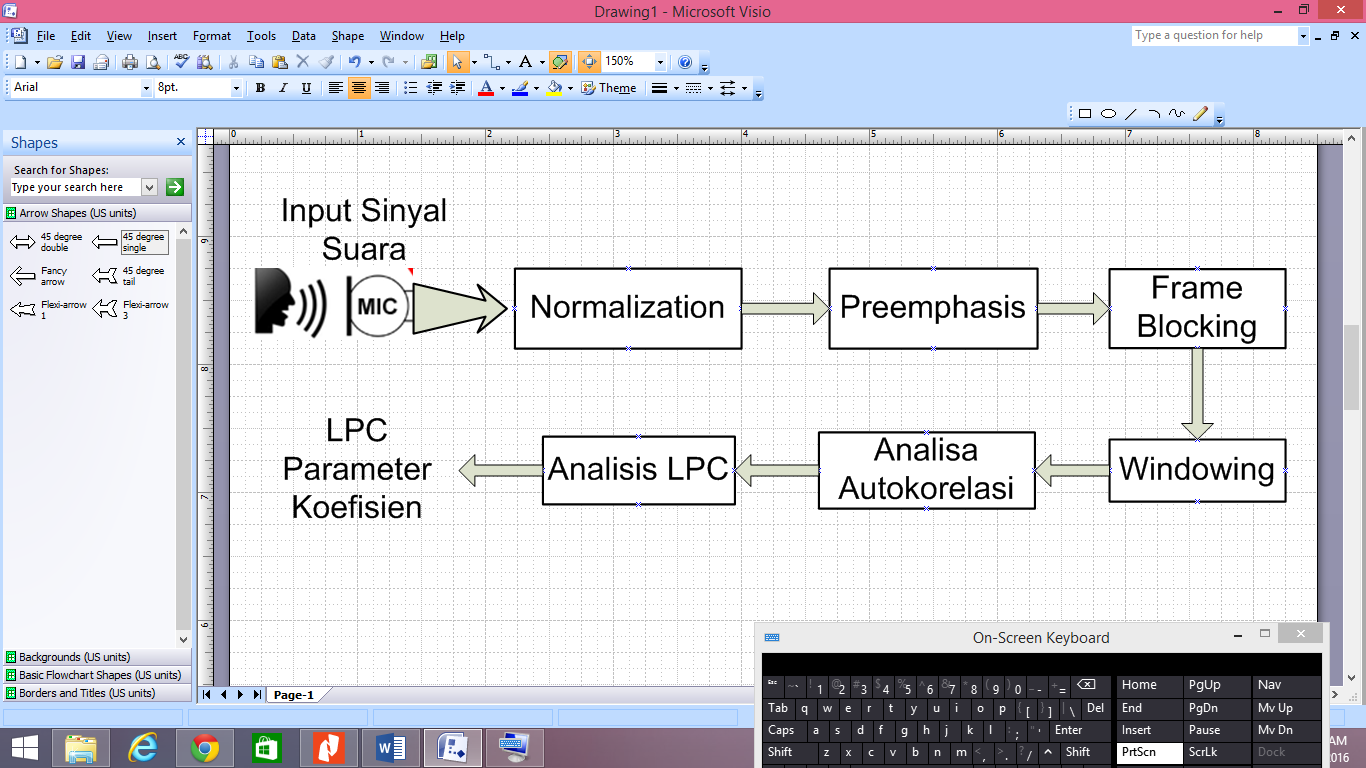


Figure 1. Block diagram of LPC method

LPC calculation steps are as follows [5]:

1. Pre-Emphasis

First of all an analysis of the speech sample is performed by passing the sound through a filter in order to obtain a smoother spectral frequency of speech signal and make it less of a range to the precision effect. The coefficient of the filter should be between 0.9 and 1. Where the relative spectral shape is in a high value to low-lying areas and tends to drop sharply for frequencies above 2000 Hz. Filter preemphasis is based on the relationship of input / output in the time domain which is expressed in the different equations such as Following [1]:

(1)

Where a is the constant pre-emphasis filter, typically it is about 0.9 <a <1.0.

1. Frame Blocking

After pre-emphasis phase, the results of the voice signal is divided into frames consisting of M samples from 20 to 40 seconds. In this process, the sound signal is segmented into several frames in overlap. This is done in order to has no missing signal (deletion). This process will continue until all signals have entered into one or more frames.

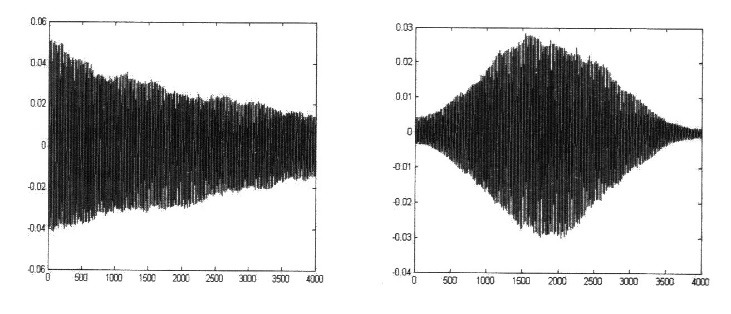
1. *Windowing*

The analog signal that has been converted into a digital signal is read frame by frame and for each frame, it is done windowing with a particular window function. Windowing process aims to minimize the lack of continuity signal at the beginning and end of each frame. If we define the window as , where is the number of samples in each of its frame, then the result of windowing is the signal [6]:

(2)

Where usually uses window hamming which shapes:

(3)



|  |  |
| --- | --- |
| 1. Before windowing | 1. After windowing |
| Figure 2. Illustration of windowing process [6] | |

1. Autocorrelation Analysis

Autocorrelation analysis on each frame windowing results with equation [1]:

(4)

Where in each is an order from LPC which is frequently used in the range of 8 to 16.

1. LPC analysis

The next step is the *LPC* analysis, which converts the each frame of p+1 autocorrelation into subsets *"*LPC parameters”.

for . Later this set becomes LPC coefficients, or another LPC  transformation. The formal methods to change from the autocorrelation coefficients into theLPC parameter set is referred to the Durbin method and can be formed form the algorithm as follows [1]:

(5)

(6)

(7)

(8)

(9)

With is the result of autocorrelation, is an *error*, is the reflection coefficient, is a predictive coefficient for .

1. LPC parameter conversion into *LPC* coefficients

To change LPC parameter into cepstral coefficient to get better performance and resistance from noise as on the equation [1]:

(10)

(11)

This cepstral coefficient is the coefficient of logarithmic spectrum representation.

1. ***ANFIS (Adaptive Neuro-Fuzzy Inference Systems)***

ANFIS is a method for tuning the basic rules of the Fuzzy system with a learning algorithm based on training data set in Artificial Neural Network[10][2]. ANFIS has a faster profit and significantly more accurate compared to many other neural network system [10]. ANFIS associated with Sugeno  type Fuzzy models have 2 inputs and a single output as shown in Figure 5. Specifically ANFIS only supports on Sugeno type system. Based on this system we can [11]:

* Entering data (training*,* testing and checking)
* Producing or entering an early FIS models.

The rule sets of the first order Sugeno Fuzzy system as follows

Rule :

F:\KULIAH S2 ELEKTRO UNHAS\Semester 3\Gambar\Drawing7.emf

Figure 3. ANFIS Structure

In ANFIS system as shown above, there are five layers processe with functions and equations of each layers are respectively described as follows [1]:

**Layer 1: Fuzzyvication**

The output of node i at layer 1 is denoted as *.* Each node in layer 1 is adaptive to the output:

(12)

(13)

Information:

* and  are the values of the input node
* and  is a fuzzy set.

So each node on layer 1 functions to generate membership levels.

**Layer 2: Product**

Each node in this layer serves to calculate the activation force (firing strength) on each *rule* as a product of all the inputs that enter or as the operator t-norm (triangular-norm):

(14)

So that

(15)

(16)

The output from this layer acts as a weight function.

**Layer 3: Normalization**

Each node in this layer is nonadaptive, its function is only to calculate the ratio between the firing strength of the rule firing and the total strength of all the rule:

(17)

**Layer 4: Defuzzyfication**

In this layer, each node is adaptive with the function:

(18)

Information:

* is an output of layer 3
* is a parameter set of Sugeno model fuzzy in the first order.

**Layer 5: Total Output**

At this stage, all over the input signal (the output of layer 4) are added together so that it is given a single node with denoted by .

(19)

The entire layer will build an adaptive-networks which function the same as a first-order Sugeno fuzzy model.

1. **Pulse Width Modulation (PWM)**

One of the DC motor speed control technique is to use a PWM (Pulse Width Modulation). PWM signal is generated by using timer feature on arduino or other microcontroller. PWM is a very efficient technique in providing supplay tension between on and off in full [15]. In microcontroller there are 3 types of timers that can be used to generate PWM value, each timer is connected with certain Pin in arduino (Pin which support in PWM).

The PWM value can be done by assigning an analog value to a pin that has been given a special sign as an output of PWM signal. The value is from 0 to 255 because the PWM register is used 8 bits. The 0 value implies that no voltage is issued by the Arduino PWM Pin, if the value is raised then the PWM Pin will produce the average voltage up to the maximum voltage. For more details can be seen in Figure 4.

C:\Users\Tosh\Desktop\Graphic1.emf

Figure 4. The output signal of PWM

1. **METHOD**
2. **Block Diagram**

The overall system block diagram can be seen in figure 5.

F:\KULIAH S2 ELEKTRO UNHAS\Semester 3\Gambar\Drawing6.emf

Figure 5. Block diagram of DC motor speed control system based on speech recognition

Based on the block diagram, the first stage performed is to make a sound recording using a microphone. The recorded sound is extracted feature by using LPC method, the extraction result of the characteristic is made database so that it can be used as the input on speech recognition method. The voice recognition output will be connected to the controller section through the serial communication line, in this section the controller will identify the input type and will be processed to produce PWM signal output. The PWM signal is then fed to the DC motor driver so that it can rotate at a speed corresponding to the PWM value.

1. **Flowchart**

Systematically the sequence or steps taken in the research can be seen in Figure 6. The recording process to produce characteristics using LPC while the pattern recognition method or feature used ANFIS. At this stage the data is trained and tested to produce a minimum error.

F:\KULIAH S2 ELEKTRO UNHAS\Semester 3\Gambar\Flowchart 1.emf

Figure 6. Flowchart System

1. **RESULTS** **AND DISCUSSION**

## **Voice Recording Process and LPC Feature Extraction**

The first step is to record the voice signal that will be used as a voice command. There are 5 types of voice commands that are "Nyala", "Lambat", "Sedang", "Cepat", and "Mati". Each voice command will be done recording process repeatedly 5 times so that the total recording as many as 15 pieces.

In order to produce sound characteristics, the recorded sound signals will be carried out by several processes as discussed in the previous chapter which are the stages of LPC method calculation. The number of characteristics generated from each recorded voice instruction is 6 characteristics

The following is the result of "Lambat" sound feature extraction using LPC method.

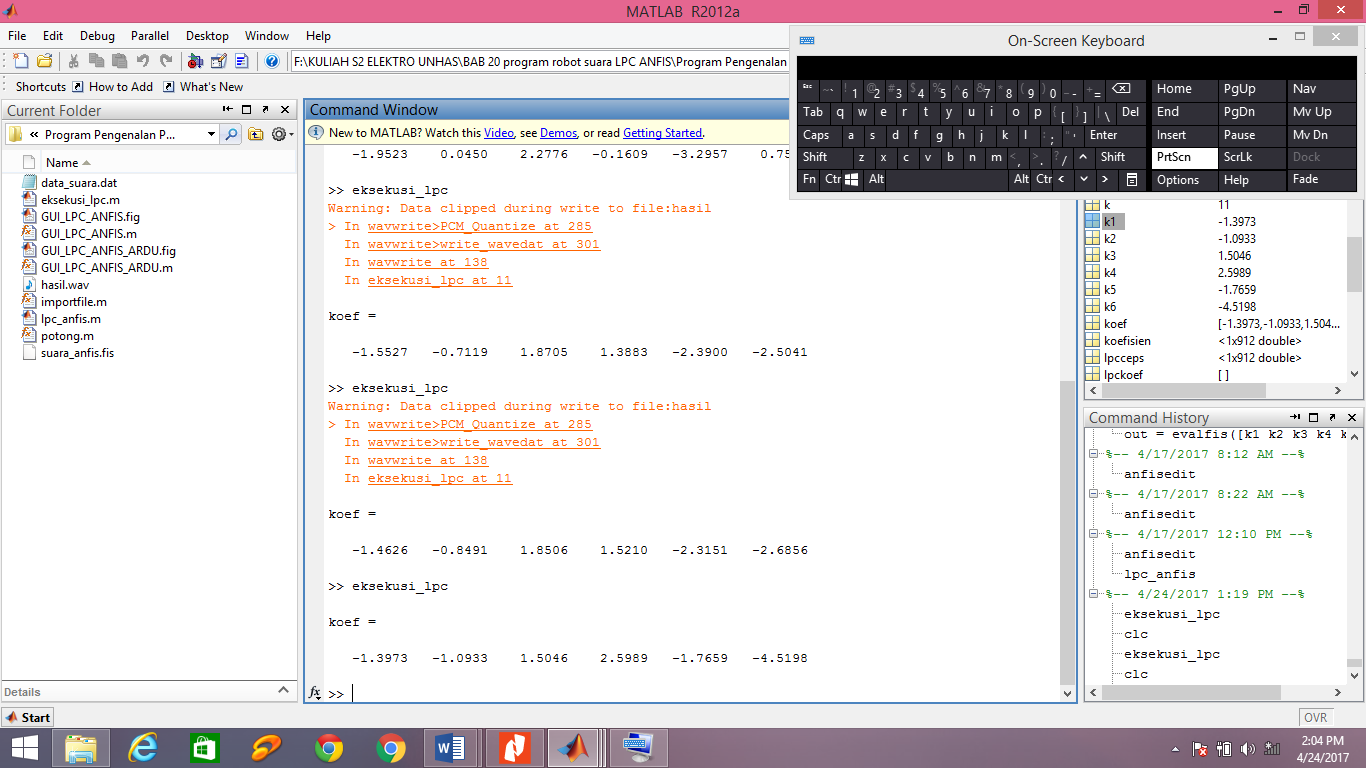


Figure 7. The result of "Lambat" sound feature extraction with 6 coefficients

F:\KULIAH S2 ELEKTRO UNHAS\Semester 3\Gambar\Sinyal_Suara.emf

Figure 8. Sound Signal Instruction "Lambat"

F:\KULIAH S2 ELEKTRO UNHAS\Semester 3\Gambar\Suara_Pre_Emphasiis.emf

Figure 9. The Pre-Emphasis Sound Signal

Sound recording results that have been done repeatedly stored in a database. Each voice instruction is given different target values ​​so that ANFIS can map patterns based on the input sound signal. The "Nyala" sound signal is given the value of the "1" tag, "Lambat" is given the target value "2", "Sedang" is given the target value "3", "Cepat" is given the target value "4" and "Mati" target is "5" .

Table 1. Database features voice signals and targets

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Ciri 1 | Ciri 2 | Ciri 3 | Ciri 4 | Ciri 5 | Ciri 6 | Target |
| -1.7677 | -0.1021 | 1.6886 | 0.5121 | -1.8426 | -1.7719 | 1 |
| -1.6023 | -0.4476 | 1.9782 | 0.6927 | -2.3422 | -1.4263 | 1 |
| -1.4706 | -0.6969 | 2.1161 | 0.4289 | -2.1715 | -0.8736 | 1 |
| -1.4669 | -0.3786 | 1.4647 | 0.6932 | -1.2341 | -1.6367 | 1 |
| -1.8475 | -0.5192 | 3.4149 | 0.1234 | -5.5467 | -0.4700 | 1 |
| -1.9268 | -0.0228 | 2.2992 | 0.1266 | -3.5794 | 0.2417 | 2 |
| -1.9403 | -0.1134 | 2.6546 | -0.1136 | -3.9561 | 0.4861 | 2 |
| -1.9031 | 0.0491 | 1.9926 | -0.1322 | -2.4741 | 0.6009 | 2 |
| -1.4234 | -0.4422 | 1.0431 | 1.1871 | -1.0936 | -1.6250 | 2 |
| -1.6134 | -0.1516 | 1.3231 | 0.6505 | -1.7152 | -0.8273 | 2 |
| -1.9292 | 0.4036 | 1.5264 | -0.7654 | -1.2847 | 1.4278 | 3 |
| -1.9142 | 0.2327 | 1.9155 | -0.6792 | -2.1662 | 1.3777 | 3 |
| -1.1542 | -1.0435 | 0.7742 | 1.3284 | 0.2779 | -0.9242 | 3 |
| -1.7457 | -0.2297 | 1.6724 | 0.3809 | -1.4830 | -0.4979 | 3 |
| -1.7362 | 0.0500 | 1.3486 | 0.0621 | -1.4437 | 0.1584 | 3 |
| -2.0382 | 0.2790 | 1.9352 | 0.1780 | -3.0410 | -0.0875 | 4 |
| -1.6658 | 0.0152 | 1.5230 | -0.0143 | -1.7191 | 0.2726 | 4 |
| -1.4411 | -0.3502 | 1.2465 | 0.8021 | -1.4006 | -1.1076 | 4 |
| -1.3236 | -0.5604 | 0.8890 | 1.2725 | -0.8888 | -1.4343 | 4 |
| -1.9617 | 0.0151 | 2.5053 | -0.3193 | -3.5598 | 0.9157 | 4 |
| -0.9764 | 0.0329 | -0.2511 | -0.4143 | 0.1425 | 0.6363 | 5 |
| -0.9651 | -0.1267 | 0.0767 | -0.6665 | 0.2460 | 0.7129 | 5 |
| -0.5480 | -0.2493 | -0.6381 | -0.3819 | 0.1689 | 0.5791 | 5 |
| -1.3828 | 0.3527 | 0.0325 | -0.5897 | 0.5769 | 0.5592 | 5 |
| -1.3811 | 0.2215 | 0.1857 | -0.3169 | 0.0431 | 0.4620 | 5 |

## **Training data with ANFIS**

ANFIS is used to perform database training based on the target that has been made. The type of FIS used is sugeno. After entering the database file to be trained, automatically ANFIS editor defines that there are 6 types of input and 1 output used as the target. The total membership of each input is given 3 conditions so that overall there are 18 number of function members. Before conducting the training data, firstly determine the method and amount of epoch training FIS. FIS training optimization method is hybrid and the epoch number is 50. This epoch determine whether data is trained to have a very small error (close to zero).

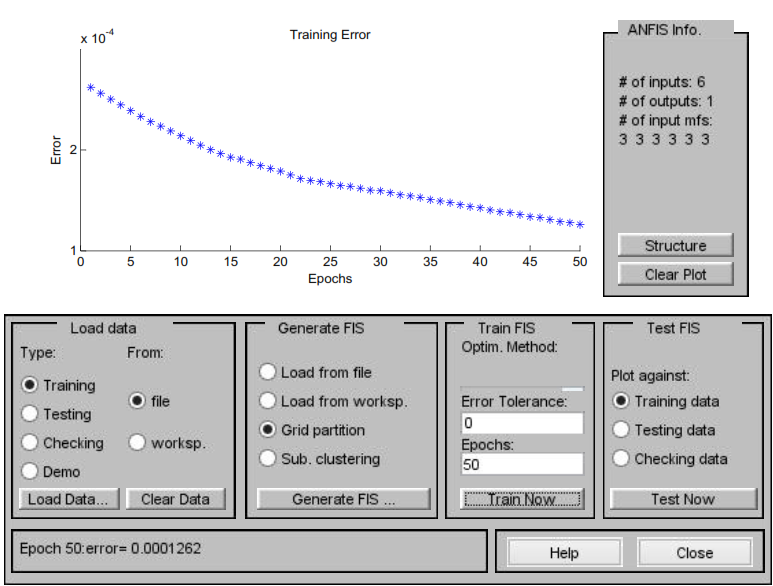


Figure 10. Database training with epoch is 50

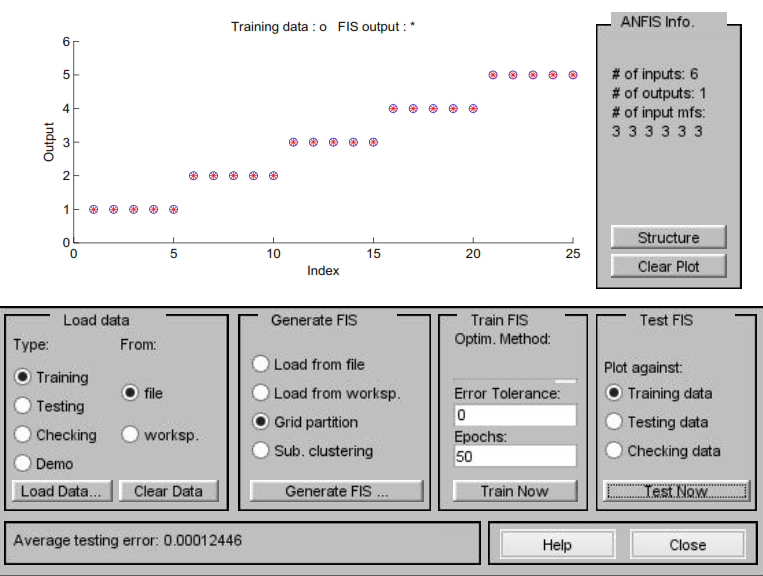


Figure 11. training and testing result

Based on the figure 10, we can be seen that the average number of *errors* generated after the initial data with the number of training epochs 50 is 0.00011963. Furthermore, the data that had been trained earlier will be tested and the results can be seen in Figure 11. The results of the plot between the training data and FIS *output* is shown in Figure 11.

## **Testing System**

1. **Voice Recognition Test**

After doing the training and trial test on the database in which the errors occurred very small then the next stage is to conduct a trial using the new data with the same kind of voice commands. Testing is done by entering a new voice command from both the respondent in the database and outside the database.

Table 2. Testing of the respondens in the database

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No.** | **Voice Command Type** | ***Output Program*** | ***Output***  ***detected*** | **Target** | ***Error*** |
| 1 | Nyala | 0 – 1,25 | 1,0783 | 1 | -0,783 |
| 2 | Lambat | 1,25 – 2,25 | 1,5251 | 2 | 0,4749 |
| 3 | Sedang | 2,25 – 3,25 | 2,6547 | 3 | 0,3453 |
| 4 | Cepat | 3,25 – 4,25 | 3,3082 | 4 | 0,6918 |
| 5 | Mati | 4,25 – 5,25 | 5,0478 | 5 | -0,478 |
|  |  |  | Total Error | | 0,251 |
|  |  |  | Error Average | | 0,0502 |

Table 3. Testing the respondens outside of the database

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **No.** | **Voice Command Type** | ***Output Program*** | ***Output***  ***detected*** | **Target** | ***Error*** |
| 1 | Nyala | 0 – 1,25 | 0,6025 | 1 | 0,3975 |
| 2 | Lambat | 1,25 – 2,25 | 2,0617 | 2 | 0,0617 |
| 3 | Sedang | 2,25 – 3,25 | 3,1359 | 3 | -0,1359 |
| 4 | Cepat | 3,25 – 4,25 | 3,6391 | 4 | 0,3609 |
| 5 | Mati | 4,25 – 5,25 | 4,5498 | 5 | 0,4502 |
|  |  |  | Total Error | | 1,1344 |
|  |  |  | Error Average | | 0,22688 |

1. **Testing the control circuit**

DC motor speed control circuit uses arduino with PWM signal as a representation of fast or slow motor rotation. Based on the test results of the control circuit with the voice input, the PWM signal coming out of the arduino corresponds to the voice command input. When given voice instruction “Lambat” then arduino will receive serial character data from the computer, then it will be processed and produce the PWM signal with a value is 125 and with a duty cycle is 49%.

To get the duty cycle value at other PWM signal output is using the same formula and the result can be seen in table 4.

Table 4. Output of PWM Signal based on voice command input

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **No.** | **Voice Recognition Type** | ***Output Program*** | **Signal PWM** | **Duty Cycle** |
| 1 | Nyala | 0 – 1,25 | 62,5 | 24,5% |
| 2 | Lambat | 1,25 – 2,25 | 125 | 49% |
| 3 | Sedang | 2,25 – 3,25 | 187,5 | 73,5% |
| 4 | Cepat | 3,25 – 4,25 | 250 | 98% |
| 5 | Mati | 4,25 – 5,25 | 0 | 0 |

The DC motor speed control circuit can be seen in figure 12.

F:\KULIAH S2 ELEKTRO UNHAS\Semester 4\Skema Tesis.EMF

Figure 12. Control Circuit Schematic

## After performing the control circuit simulation, we obtain the PWM signal form for each output as shown in Figures 13, 14, 15, and 16.

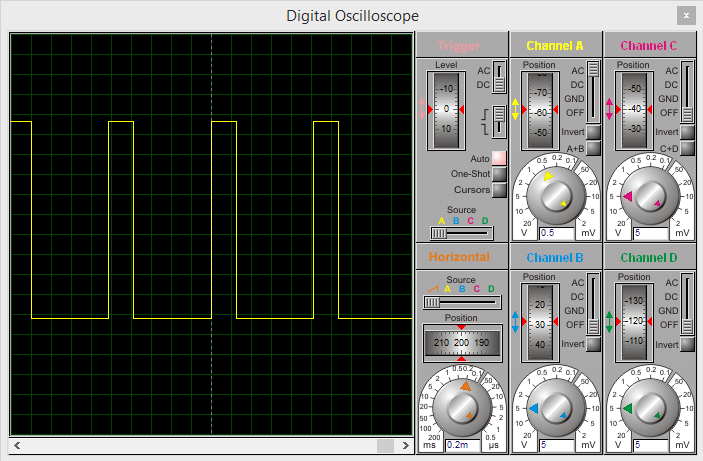


Figure 13. PWM Signal with duty cycle 24,5%

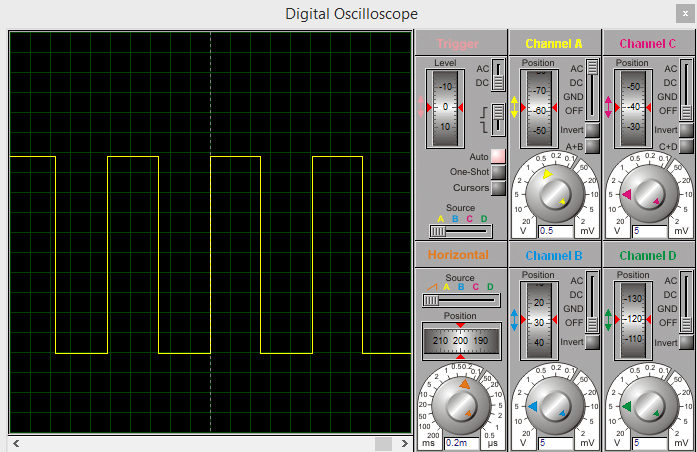


Figure 14. PWM Signal with duty cycle 49%

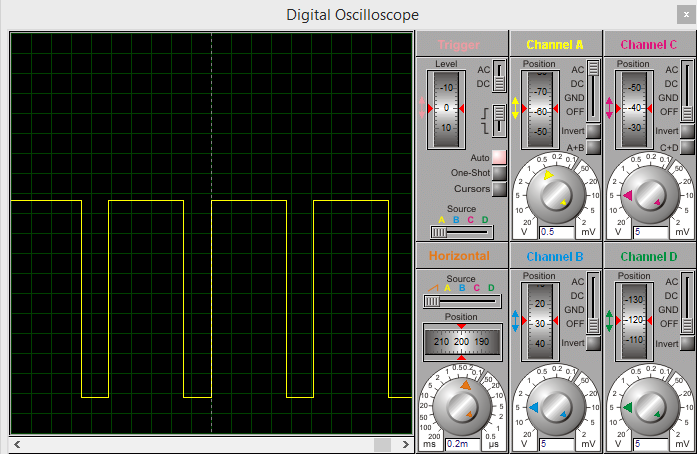


Figure 15. PWM Signal with duty cycle 73,5%

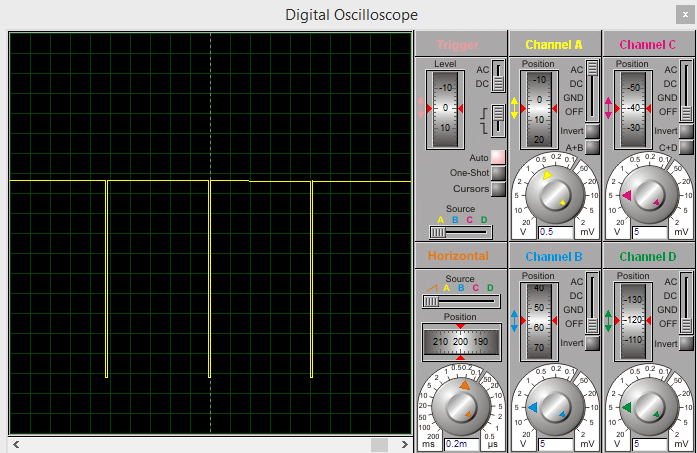


Figure 16. PWM Signal with duty cycle 98%

1. **CONCLUSION**

Overall the purpose of this study can be achieved from the results obtained in the previous discussion about the characteristic extraction of the recording using LPC method issued a unique feature for each voice instruction. Each of these characteristics can be well recognized by ANFIS as a sound pattern recognition instrument. From the results of experiments derived from the respondents resulted in an average error of 0.0502, while outside the respondents produced an average error of 0.22688. While in a control circuit experiment using serial communication between computers and arduino, data can be sent according to the input signal that enters and outputs the correct PWM signal in order to run the DC motor with changing speed as the voice input signal. So it can be concluded the circuit works 100%.

**REFFERENCE**

|  |  |
| --- | --- |
| [1] | Sanjaya, M. 2016. Intelligent Speech Recognition based Robot using Matlab and Arduino. Yogyakarta: Andi. |
| [2] | Ronando, E. and Isa, IM. Speech Recognition Word 2012. As a Movement Control Robot Arm In *Real-Time* *Linear Predictive Coding* Method of *Neuro-Fuzzy.* Journal of Science and Arts ITS Vol. 1, No. 1. PP A51-A56. |
| [3] | S. Wijoyo, *"Speech Recognition Using Linear Predictive Coding and Artificial Neural Network for Controlling Movement of Mobile Robot"* *2011 Int.* *Conf.* *Inf.* *Electron.* *Eng.* , Vol. 6, pp. 179-183, 2011. |
| [4] | Kurniawan, P. et al. 2012. Design and Analysis of Voice Recognition Music Player Based Application System Using Artificial Neural Network Back Propagation (JST-BP). Thesis. Telkom University. |
| [5] | Gupta, H. et al. 2016. *LPC and LPCC Method of Feature Extraction in Speech Recognition System.* *International Conference* - *Cloud System and Big Data Engineering (Confluence).* *IEEE Conference Publications.* *Page* 498-502 |
| [6] | Sanjaya, M. and Salleh. Z. 2014. Implementation of Pattern Recognition Using Voice *Mel-Frequency Cepstrum Coefficients* *(MFCC)* and *Adaptive Neuro-Fuzzy Inferense System* *(ANFIS)* For Automatic light control. Al-HAZEN *Journal of Physics* Volume 1, No.1, page 43-54 |
| [7] | Thiang and Wanto. 2010. *Speech Recognition Using LPC and HMM Applied for Controlling Movement of Mobile Robot.* *National Seminar on Information Technology.* |
| [8] | Dave. N. 2013. *Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition.* *International Journal of Advanced Research in Education & Technology (IJARET).* Volume 1, Issue VI. |
| [9] | Yusnita MA, 2011. *Malaysian English Accents Identification using LPC and Formant Analysis.* *IEEE International Conference on Control Systems, Computing and Engineering.* Page 472-476. |
| [10] | Pandey, B. et al. *Multilingual* 2010. *Speaker Recognition Using ANFIS.* *2nd International Conference on Signal Processing Systems (ICSPS) IEEE.* Page V3-714 - V3-718 |
| [11] | Helmi, N. and Helmi B. Hoda. 2008. *Speech Recognition with Fuzzy Neural Network for Discrete Words.* *Fourth International Conference on Natural Computation IEEE.* *Page* *265-269.* |
| [12] | Thiang and Saputra, H. 2005. Word Recognition System by Using *Linear Predictive Coding* and *Nearest Neighbor Classifier.* Journal of Electrical Engineering volume 5, Number 2. Page 19-24. |
| [13] | Kumar, P. Sharma, Lakshmikantha, BR and Shanmukha, K. Sundar. 2011. *Real Time Control of DC Motor Drive using Speech Recognition.* *International Conference on Power Electronics, 2010 (IICPE2010)) IEEE.* *Page 1-5* |
| [14] | Rai, N. and Rai, B. 2014. *An Ann Based Mobile Robot Control Through Voice Command Recognition Using Nepali Language.* *International Journal of Applied Control, Electrical and Electronics Engineering (IJACEEE) Vol 2, No.4.* |
| [15] | Kumar, Ravi KS. Et al. 2015. "Microprocessor Based Closed Speed ​​Control Loop of DC Motor Using PWM". International Conference on Control, lnstrumentation, Communication and Computational Technologies (ICCICCT). IEEE |