A DC Motor Speed Control Using The LPC-ANFIS Speech Recognition System

Muhammad Akil [1], Ingrid Nurtanio [2], Rhiza Samsoe'oed Sadjad [3]
[2,3] Department of Electrical Engineering, Hasanuddin University, Makassar
[1] Department of Mechatronic Engineering, Bosowa Politechnic, Makassar
akilmuhammad1989@yahoo.co.id [1], ingrid@unhas.ac.id [2], rhiza@unhas.ac.id [3].

***Abstract* -- The aim of this research is to design an implementation of the speech recognition system to control the speed of a DC motor. The Linear Predictive Coding (LPC) method is used in the speed recognition system, tuned by the Adaptive Neuro-Fuzzy Inference Systems (ANFIS) method. There are 5 (five) samples of voice signals in Bahasa Indonesia recognized by this system, i.e.: "Nyala", "Lambat", "Sedang", "Cepat" and "Mati". Every voice signal is repeated 5 (five) times until as many as 25 samples are recorded. Their voice characteristics are extracted using the LPC method represented by the LPC coefficients stored in a database system. The ANFIS method is implemented in 50 iterations to tune and to train the LPC coefficients until the least error, i.e. 0,00012446 is obtained. Voice samples originated from the internal database system are 83% successfully recognized by this system. However; samples extracted from the human voice signals of different persons - different sex from the person whose voice signals are recorded in the database system, and from various ages - are only 78,8% successfully recognized by the system. The output of the speech recognition system is coded into the ASCII Codes and converted into the PWM signal to control the speed of a DC motor.**

Keywords -- Speech recognition, voice signal, LPC, database system, ANFIS, speed control, DC motor, PWM.

I. INTRODUCTION

One of the technique to control of DC motor speed is using the PWM (Pulse Width Modulation) Technique. PWM is a pulse width modulation which value varies from 0% to 100% commonly known as dutycicle, the greater the value the faster the speed. This technique is very efficient to provide power supply between on and off full-time [15]. The PWM values are generated with control circuits that have timer features 8 bit or more, so it can be analogized that the value of 0 means stop and the maximum value of 255 means full velocity.

This research will discuss how to create dutycicle using speech recognition, so the DC motor speed can be triggered from giving the sound input in the system. In general there are 5 types of voice commands that will be used as an input in voice recognition. These sounds represent the size of a dutycicle that will be fed to a controlling circuit based on its type. The types of voice commands are "Nyala", "Lambat", "Sedang", "Cepat" and "Mati". The "Nyala" instruction will send output dutycicle of 24%, "Lambat" represented by dutycicle of 49%, "Sedang" in 73.5%, "Cepat" in 98%, and "Mati" in 0%.

In order for the voice instructions to be recognized by the system, firstly it is necessary to extract the features to obtain a recognizable value of the computer. Based on the previous research, there are many methods that used to extract of sound signal such as LPC (Linear Predictive Coding) method [2,3,5,7,8] and MFCC (Mel-Frequency Cepstrum Coeffisient) [8], then they are followed by some methods of voice pattern recognition from the extraction results such as Adaptive Neuro-Fuzzy Inference Systems (ANFIS) [6], Artificial Neural Network (ANN) [3] and others. In this research, it is used LPC method because this is a good method to get the calculation result or prediction of sound parameter precisely [5].

In the development of today's technology, voice recognition can serve as a characteristic that can be established using artificial intelligence methods to produce the outputs that can be used to control anything such as real-time arm-robot motion control using Neuro Fuzzy [2], controlling the movement of automobile robots with the ANN (Artificial Neural Network) method [3], and as a gesture to play music player with the Back Propagation Artificial Neural Network method [6].

Based on the previous researches, the researcher can conclude that the application of voice recognition system is still focused on On-Off Control only, so we want to do the research on DC Motor Speed Control based on Speech Recognition Using Linear Predictive Coding (LPC) Method as a feature extraction of voice signal and Followed by learning methods using Adaptive Neuro-Fuzzy Inference Systems (ANFIS).

The purpose of this paper is (1) To control the speed of DC motor using speech recognition by LPC and ANFIS method, (2) To know the difference of voice recognition extraction using Linear Predictive Coding (LPC) method and (3) To get voice recognition pattern Using Adaptive Neuro Fuzzy Inference Systems (ANFIS). This paper is consists of voice recognition method using Linear Predictive Coding (LPC) and ANFIS as a voice recognition pattern, and it also discussed about DC motor speed control using Pulse Width Modulation (PWM) technique.

II. SYSTEM DESIGN

A. Diagram Block

The overall system diagram block can be seen in Fig. 1.

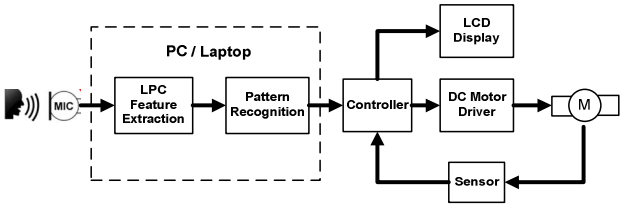


Fig. 1. Diagram block of DC Motor speed control system based on speech recognition

Fig. 1 shows the parts of the DC motor speed control system as a whole. There are 2 important parts such as computer/PC and the control circuit. Particularly, the first part is the computer/PC, this section will be done the process of recording and feature extraction of the voice using LPC. The LPC method is used to extract the features from original sounds with a good quality and efficiency for using in calculations [1]. The procedure for obtaining the LPC coefficient is shown in Fig. 2 [9].

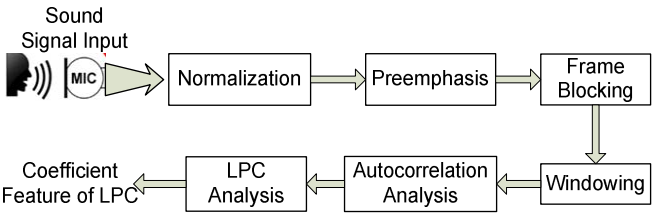


Fig. 2. LPC method block diagram

LPC calculation steps are as follows [5]:

1. Pre-Emphasis
2. Frame Blocking
3. Windowing
4. Autocorrelation Analysis
5. LPC Analysis
6. Conversion LPC parameter to LPC Coefficient

The feature extraction results are incorporated into the pattern recognition block, in this section the ANFIS (Adaptive Neuro-Fuzzy Inference System) is used. ANFIS has significantly faster and more advantages than many other neural network systems [10]. Specifically, ANFIS only supports Sugeno type system. Based on this system we can [11]:

• Entering data (training, testing and checking)

• Produce an early FIS model or enter an initial FIS model.

There are five process layers in ANFIS namely Fuzzification, Product, Normalization, Defuzzification and Output Total. All these layers will build adaptive-networks that have function same as the first-order Sugeno fuzzy model.

In the second part is the control circuit. This section is a hardware of the system which consists of several electronic circuits such as controller, LCD, Driver motor, sensor and DC Motor as control.

B. Methods

The systematic order or steps taken in the research can be seen in Fig. 3.

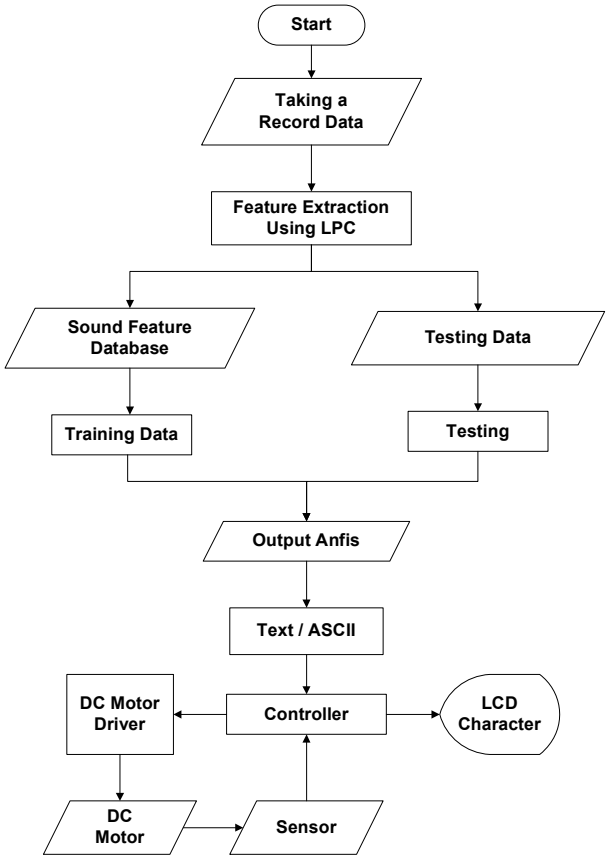


Fig. 3. Flowchart System

Based on Fig. 3, the first step is recording each type of voice instruction. The recording process is done 5 times repeatedly for each voice, so the total sound recordings are 25. Of course, the determination of this amount is unlimited and can be done more than the specified amount. Voice recording process is done in the room that is not too crowded.

The method that used to generate the sound characteristics of the recording is the LPC method. This method will exclude the sound characteristic of the extraction after recording. The sound recording process is done by one person and the result will be stored in the database. Furthermore, the database will be trained using ANFIS. There are two types of data such as training data and testing data. Data stored in the database will be trained to produce a small error which closes to zero. The results of training data will be tested on several respondents to determine the ability of the system in recognizing the data from outside of the database. The number of respondents from outside of the database are selected by sex and age.

III. RESULTS AND DISSCUSSION

A. *Sound Recording Process and LPC Feature Extraction*

The first step records the voice signal that will be used as a voice command. There are 5 types of voice commands: "Nyala", "Lambat", "Sedang", "Cepat", and "Mati". Each voice command will be recorded repeatedly in 5 times so that the total recording as many as 25 pieces.

To produce sound characteristics, the recorded voice signal will be performed on several stages of process that have discussed in the previous chapter like the stages of LPC method calculation. The number of characteristics that is generated from each recorded voice instruction is 6 characteristics.

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

```
>> eksekusi_lpc

koef =

   -1.3973   -1.0933    1.5046    2.5989   -1.7659   -4.5198

fx >>
```

Fig. 4. The Result of the sound feature extraction "Lambat" with 6 coefficients

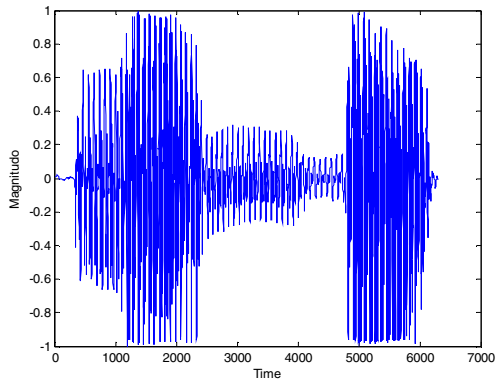


Fig. 5. "Lambat" instruction sound signal

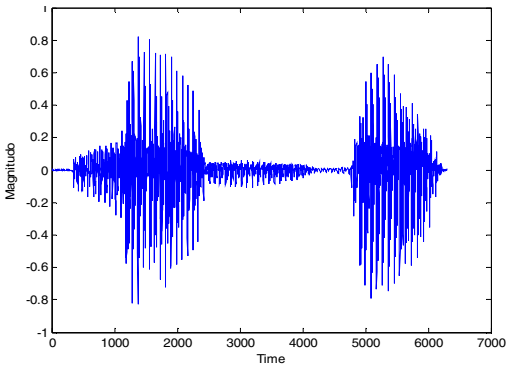


Fig. 6. Pre-Emphasis sound signal

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 target "1", "Lambat" is given the target value of "2", "Sedang" is given the target value "3", "Cepat" is given the target value "4" and "Mati" with the target "5". Below is the entire database of sound feature that had recorded and their targets.

TABLE I. SOUND SIGNAL FEATURE DATABASE AND THE TARGET

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6	Target	Sound
-1.7677	-0.1021	1.6886	0.5121	-1.8426	-1.7719	1	Nyala
-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	Lambat
-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	Sedang
-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	Cepat
-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	Mati
-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	

B. *Data Training with ANFIS*

ANFIS is used to perform database training based on the target that has been made. The type of FIS which is 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 there are 18 number of function members. Before doing the training data, firstly determine the method of training FIS and the number of epoch, the FIS training optimization method that used is Hybrid while the number of epoch is 50. This epoch is very determining whether the trained data has a very small error (close to zero).

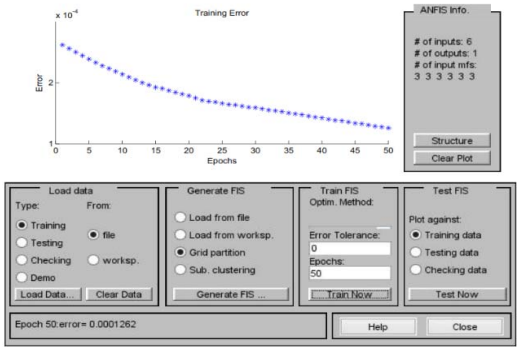


Fig. 7. Database training with epoch 50

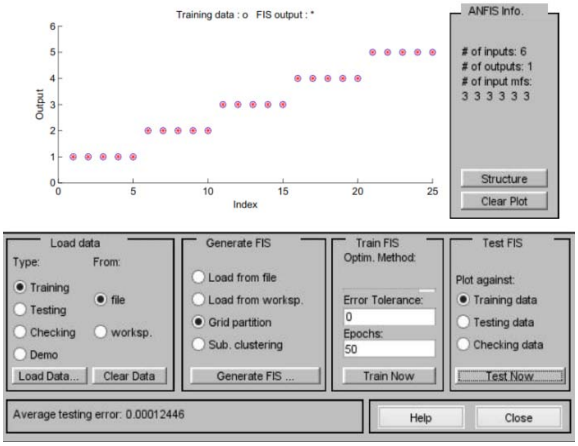


Fig. 8. The training data testing result

Based on Fig. 7, it can be seen that the average number of errors which is generated after the initial training data with the number of epoch 50 is 0.00012446. Then, data that have been trained before will be tested and the results can be seen in Fig. 8. The result plot between training data and FIS output is shown in Fig. 8.

C. System Testing

1) Voice Recognition Testing

After doing the training and the testing on the database where errors that appear very small, then the next stage is to test using new data with the same type of voice command. Testing is done by entering a new voice command from both the respondent in the database and outside the database. The number of respondents tested is 20 for men and women with age between 18-25 years old.

TABLE II. TESTING OF THE RESPONDENS IN THE DATABASE (1=DETECTED, 0=UNDETECTED)

Nu.	Name	Testing					
		Repeating	Nyala	Lambat	Sedang	Cepat	Mati
1	Muhammad Akil (Man, 27 Years Old)	1	1	1	0	1	1
		2	1	1	1	1	0
		3	1	0	1	1	0
		4	1	0	1	1	1
		5	1	1	1	1	1
		6	1	1	0	1	1
		7	1	1	1	1	1
		8	1	1	1	1	1
		9	1	1	1	1	1
		10	1	1	0	1	1
		11	1	1	1	1	1
		12	1	1	1	1	0
		13	1	1	1	1	0
		14	1	0	1	0	1
		15	1	0	1	0	1
		16	1	1	1	1	1
		17	1	1	1	1	0
		18	1	1	1	1	1
		19	1	1	0	1	0
		20	1	1	0	1	1
Total			**20**	**16**	**15**	**18**	**14**

The test is done by repeating each voice instruction until 20 times so that the total number of test data taken is 100. The total number of detected instructions is 83% while the error number is 17%.

The testing process on the respondents outside of the database is done by repeating 5 times for each voice instruction. Here's the percentage for overall result of success and error testing system.

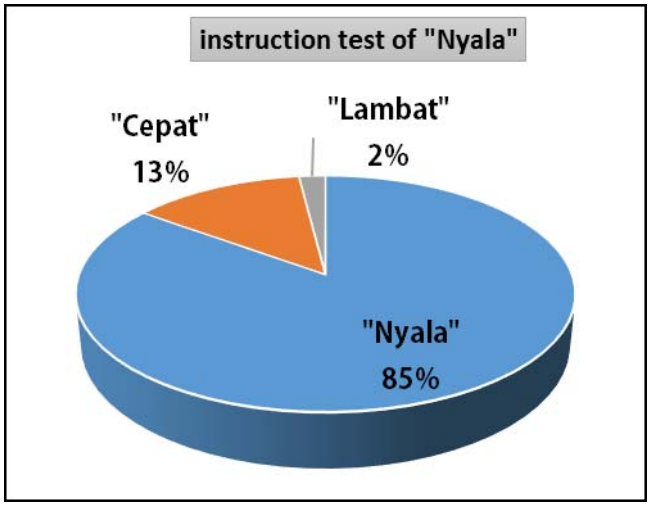


Fig. 9. Graph of percentage of successful Instruction test of Nyala

The "Nyala" instruction is detected at 85% but at the time of trial there is a mistake that the system detects the other instruction type results amount 13% "Cepat" and "Lambat" instructions only 2%. This error arises because the voice input during the test approaches the intonation of the Cepat and Lambat instructions.

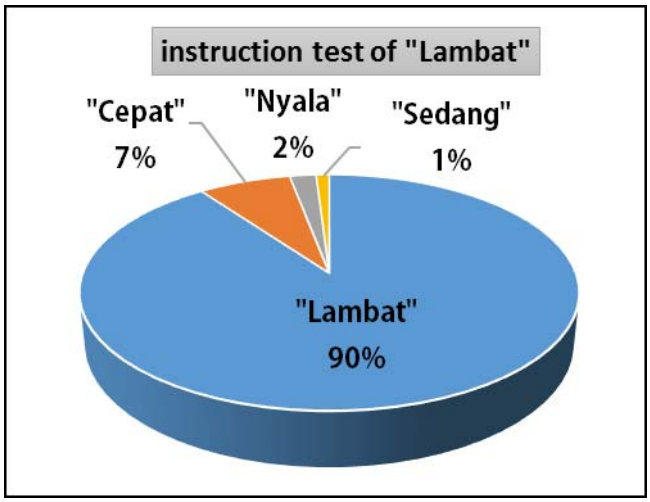
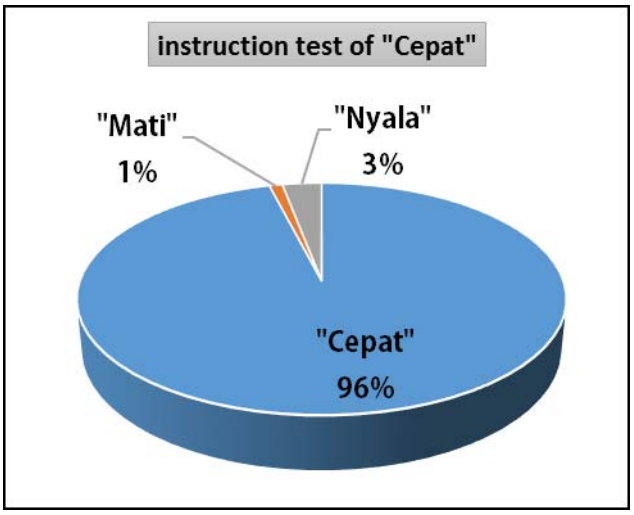
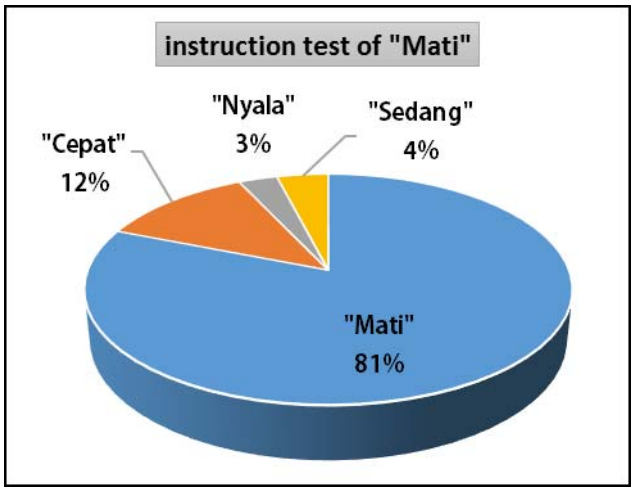


Fig. 10. Graph of successful percentage of Lambat Instruction trial

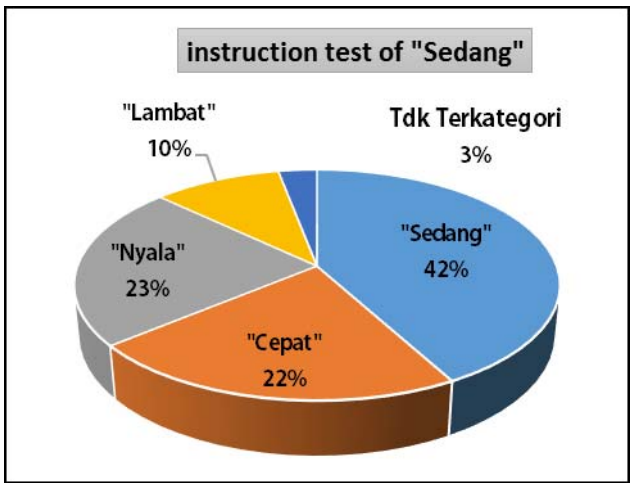
Lambat instruction test results show a success rate of 90% with 10% detection errors, the errors that occur is the system detects Cepat 7% instructions, and 2% Nyala, and Sedang only 1%.



(a) The Successfully testing graph of Cepat instructions



(b) The Successfully testing graph of Mati instructions



(c) The Successfully testing graph of Sedang instructions

Fig. 11. Graph of Percentage of Success Testing Type of Instruction (a) Cepat (b) Mati and (c) Sedang

The Overall can be shown the number of percentage of success outside the database test.

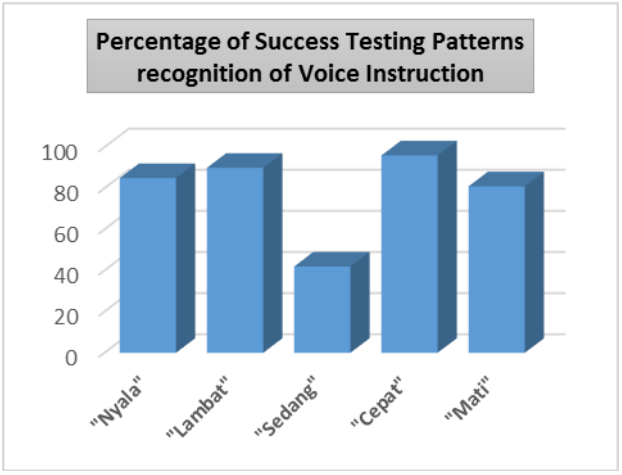


Figure 12. Percentage of success Testing of respondents outside the database

Based on figure. 12, respectively the best type of instruction detected is when the test is Cepat, Lambat, Nyala and Mati instruction, but there is a very low result on Sedang instruction, the result is lower by 42%.

The most common mistake is during the sound recording process by entering a type of sound that resembles intonation with another sound instructions so that the results obtained are different types of sound insertion. Then at the time of recording the respondent was late to mention the type of voice instruction so that the result is cut and caused the result is not perfect.

2) Testing of Control Circuit

A 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 from the Arduino corresponds to the voice command input. When it is given the voice instruction "Sedang" the Arduino will receive the serial character data from the computer, then it will be processed and produce the PWM signal with value is 187,5 and with a duty cycle is 73,5%.

$$Duty\ Cycle = \frac{T_{on}}{T_{total}} x\ 100\%$$

$$Duty\ Cycle = \frac{187,5}{255} x\ 100\%$$

$$Duty\ Cycle = 73,5\%$$

To get the duty cycle value at other PWM signal output is using the same formula and the result can be seen in Table 3.

TABLE III. OUTPUT OF PWM SIGNAL BASED ON VOICE COMMAND INPUT

No.	Voice Recognition Type	*Output Program*	Signal PWM	Duty Cycle
1	Nyala	1	62,5	24,5%
2	Lambat	2	PWM-15	decrease
3	Sedang	3	187,5	73,5%
4	Cepat	4	PWM+10	increase
5	Mati	5	0	0

IV. CONCLUSION

Based on the results obtained on the voice instruction test, the feature of voice feature extraction with the number of voice features in each type of instruction is 6 characteristics. The results obtained at the time of testing with respondents in the database by 83% with the vote data 100 times. While the test on the respondents outside the database for each type of voice instruction obtained an average percentage of 78.8%.

REFFERENCE

[1] Sanjaya, M, Intelligent Speech Recognition based Robot using Matlab and Arduino, Yogyakarta, Andi, 2016.

[2] Ronando, E. and Isa, IM, "Speech Recognition 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", Int. Conf. Inf. Electron. Eng. , Vol. 6, pp. 179-183, 2011

[4] Kurniawan, P. et al, , "Design and Analysis of Voice Recognition Music Player Based Application System Using Artificial Neural Network Back Propagation (JST-BP)", Thesis, Telkom University, 2012.

[5] Gupta, H. et al, "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, 2016.

[6] Sanjaya,M, and Salleh,Z, "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, 2014.

[7] Thiang and Wanto, "Speech Recognition Using LPC and HMM Applied for Controlling Movement of Mobile Robot", National Seminar on Information Technology, 2010.

[8] Dave. N, "Feature Extraction Methods LPC, PLP and MFCC In Speech Recognition". International Journal of Advanced Research in Education & Technology (IJARET), Volume 1, Issue VI, 2013.

[9] Yusnita MA, "Malaysian English Accents Identification using LPC and Formant Analysis", IEEE International Conference on Control Systems, Computing and Engineering, Page 472-476, 2011.

[10] Pandey, B. et al, "Multilingual 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, "Speech Recognition with Fuzzy Neural Network for Discrete Words", Fourth International Conference on Natural Computation IEEE, Page 265-269, 2008.

[12] Thiang and Saputra, H, "Word Recognition System by Using Linear Predictive Coding and Nearest Neighbor Classifier", Journal of Electrical Engineering volume 5, Number 2, Page 19-24, 2005.

[13] Kumar, P. Sharma, Lakshmikantha, BR and Shanmukha, K. Sundar, "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., "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, 2014.

[15] Kumar, Ravi KS. Et al., "Microprocessor Based Closed Speed Control Loop of DC Motor Using PWM", International Conference on Control, Instrumentation, Communication and Computational Technologies (ICCICCT) IEEE, 2015