

Classification and Real Time Testing of Speaker Recognition System

N K Kaphungkui¹, Dr. Aditya Bihar Kandali²

Assistant Professor, Department of Electronics and Communication Engineering, DUIET, Dibrugarh University, Assam, India¹

Associate Professor, Department of Electrical Engineering, Jorhat Engineering College, Assam, India²

ABSTRACT: Speaker Recognition is a process of validating the person's identity to have access in the system by using the person's speech/voice characteristic which is unique to each individual. It is a system of man-machine interface. The objective of this work is to classify fourteen (14) speakers with high accuracy rate using Mel frequency Cepstral Coefficient (MFCC) as speech features extraction method and Back Propagation Neural Network as speech classifier. After classification, the classifier is again tested with real time new speech data to recognize each registered speaker's. Training function with Scaled conjugate gradient approach is used for back propagation neural network. A voice database from a group of seven male and seven female is created uttering the same sentence repeatedly. The accuracy obtained from the classification is 93.1% with small misclassification of 6.9%. The overall precision, sensitivity and specificity scores are 93.02%, 93.21% and 99.46% respectively which are good enough. MATLAB tool is used for the entire simulation.

KEYWORDS: Speaker Recognition, MFCC, Back Propagation Neural Network, Confusion Matrix, Accuracy, Training, Testing.

I. INTRODUCTION

Speaker recognition is the process of identifying the person from the characteristics of his/her voice signal. The broad classification of Speaker recognition is shown in Figure 1 [2], [13]. In an open set type the system can have many trained speakers but in a closed set, the system has only a specified number of speakers which are registered into the system. In a Text-Dependent system type, the same spoken words or sentence is used for training session as well as the testing session by the system. So here, prior information is known to the speaker about the system. Hence for successful recognition, the system required cooperation from the test speaker [1]. Whereas in text independent type (unconstrained mode) as no prior information about the utterances is available to the test speaker, he/she can speak anything while testing the network [4].

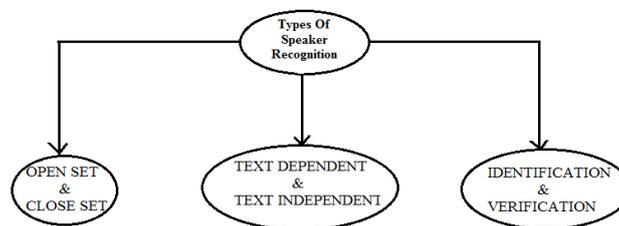


Figure.1 Speaker Recognition classification system

Finally Speaker identification type deal with the finding of unknown speaker's identity by comparing the given utterance with the voice database registered into the system. It follows the one to many comparison procedure (1:N). But in Speaker verification system, it deals with to either accept or dismiss the person claimed to be as a speaker following the procedure of 1:1 mapping process [17]. Extraction of speech features and speech feature matching are two important integral part of speaker recognition system. Several common methods used for extraction of unique speech features are Linear Prediction coefficients (LPC), Mel-Frequency Cepstral Coefficients (MFCC), Gammatone Frequency Cepstral Coefficients (GFCC), Linear Predictive Cepstral Coefficients (LPCC) and Perceptual Linear Predictive (PLP). In a similar manner, for Speech features matching several classifiers such as Dynamic Time Warping (DWT), Gaussian Mixture Model (GMM), Hidden Markov Model (HMM), K- Mean Clustering and Vector

Quantization (VQ) are employed. The training and testing phase are two important steps which govern all types of speaker recognition system [19]. A Researcher have reported that efficiency of 81.8% is achieved for class ten classification by using 70% of total samples for training and 30% of total samples for system testing with the combination of MFCC, pitch and rms in feed forward neural network (FFNN) [12]. With respect to accuracy of the classification, Artificial Neural Network give results better than fuzzy logic based systems when the speech is recorded in a controlled and clean environment. Accuracy of 74% is obtained with ANN against 72% accuracy rate with fuzzy logic [6]. Others work have also achieved the highest accuracy of 92% for 10 speakers classification with the combination of Mel Frequency Cepstrum Co-efficients and BPNN in text dependent speaker recognition system. [15]. A moderate accuracy is obtained with the new approach of combining LPC and MFCC for Assamese Speaker Recognition Using Artificial Neural Network for 10 users [3].This work will be implemented for 14 users with the combination of MFCC and BPNN to achieve an accuracy of 93.1%.

II. MEL FREQUENCY CEPSTRAL COEFFICIENT

Many experts who work in the field of speech have reported that the extraction of speech features with MFCC approach is gaining more success and popular. This is because MFCC technique is modelled as human auditory system which fail to perceived frequency over 1KHz and showing more robust against noisy environment [5], [8]. Mel Frequency Cepstral Coefficients (MFCC) algorithm is also more preferred due to the generation of user’s voice coefficients which are unique to every individual [9]. A novel approach of using weighted MFCC is also presented for speaker recognition system and concludes that the rate of recognition is superior to non-weighted MFCC [20]. Speaker recognition base on the combination of 16-order MFCC and LPCC as speech classifier is also reported [18]. The overall steps involving for the extraction of speech feature based on MFCC are shown in Figure. 2 [7].

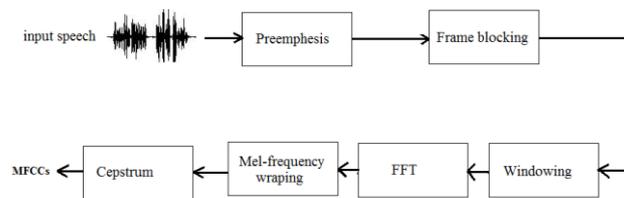


Figure. 2 Steps of MFCC computation

In this work MFCC technique is used for the extraction of the speech feature. Each of the collected speech samples is represented with a unique 13 coefficients for all the speakers. The parameters of MFCC feature extraction is set as frame duration of 26msec, frame shift of 10msec, Pre-emphasis coefficient of 0.97, number of filter banks channel of 20, filter’s Lower frequency limit of 300 Hz , filter’s Upper frequency limit of 3.7 KHz and Cepstral sine lifter parameter of 22. A data base is created from 14 speakers by uttering the same sentence repeatedly. The data base consist of seven female f1, f2, f3, f4, f5, f6, f7 and seven male speakers as m1, m2, m3, m4, m5, m6 and m7. The speech samples are recorded in a relatively noise free environment with the same recording device.The length of the speech samples variesfor different speakers as it depends on a time taken to utter the sentence. The speech length is more when more time is taken to utter the sentence and vice versa.The sampling of the voice is at 48 KHz with a 16 bit, bit depth. The resultants MFCC of speech are in matrix form consisting of fix 13 rows and variable column for different speaker as shown in Figure. 3. This is represented as M x N where M is fix rows and N is variable columns.

Sl	Speakers	MFCC Size (M x N)	Sl	Speakers	MFCC Size (M x N)
1	MFCC of f1	13 x 1181	8	MFCC of m4	13 x 1038
2	MFCC of m1	13 x 1457	9	MFCC of f5	13 x 1089
3	MFCC of f2	13 x 1111	10	MFCC of m5	13 x 916
4	MFCC of m2	13 x 939	11	MFCC of f6	13 x 1205
5	MFCC of f3	13 x 1283	12	MFCC of m6	13 x 820
6	MFCC of m3	13 x 1063	13	MFCC of f7	13 x 1031
7	MFCC of f4	13 x 980	14	MFCC of m7	13 x 1096

Figure.3 MFCC computation Table

The value of N will be more if a speaker takes more time to complete the utterance of the sentence. The pictorial representation of a particular speaker’s MFCCs is shown in Figure 4.

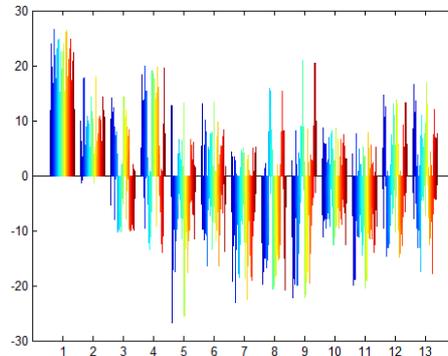


Figure. 4. Resultant Bar Plot of MFCC

III. BACK PROPAGATION NEURAL NETWORK

The basic structure of multi layer Perceptron neural network consists of three layers that is i/p layer, hidden layer and o/p layer as shown in Figure. 5.

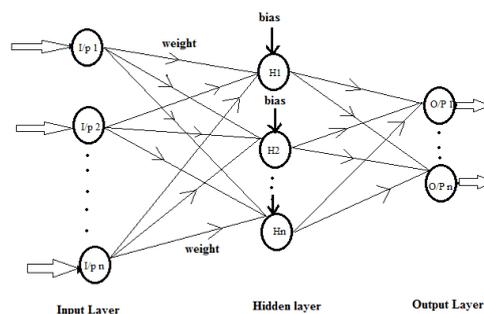


Figure. 5 Multilayer Perceptron Neural Network Structure

Two basic steps govern a Back propagation algorithms to obtain a desire output i.e feed forward and back propagation. In feed forward step, the data propagates from inputs layer towards output nodes along with the network initial weights and bias. If the output is not equal to the set target then back propagation step takes place by updating the weights backward from output layer towards input layer. This is a training algorithm for supervised learning of Artificial Neural Networks (ANN). At the initially stage, when designing the network, random values of weights are assigned to predict the desired set target. Hence if any variation exist between the output and the desire target value, the weights are updated to get a minimize error. The weights minimizing the error function between the output and the target is the solution of the neural network leaning. The overall step by step execution of BPNN is summarized as follows

- Network Error Calculation – obtain the difference between network output and desire output in feed forward direction
- Error Minimization – Checking whether the network error is minimum or not
- Updating the network weights – if huge difference occurs, the network’s weight is updated and the error is checked again. The whole process is repeated until the least minimum error is obtained.
- Training session of Network is over and ready to predict the target – The network training stop when it obtained the least mean square error (MSE).

In Figure. 6 the detail flowchart of Back Propagation Algorithm is shown. The algorithm of back propagation aim to minimize the value of the mean square error (MSE) functions of the model by updating the weights of the network [10], [11], [16].The final weights which minimize the error function will be the solution of the neural network learning problem. High accuracy is obtained and performance is better when more number of training data set is used for training the network model [14].

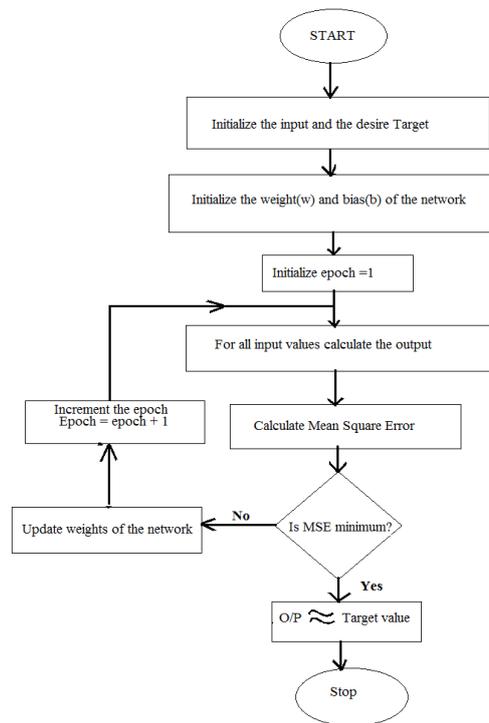


Figure.6 Flowchart of BPNN

Ideally the training session of the network is completed when the network’s actual output is equal to the desired set target value. To achieved this condition, the loop runs continuously by incrementing the epoch number along with the updating of weights.

IV. CONFUSION MATRIX

Confusion Matrix is a table for performance measurement in pattern classification. It summarizes the prediction results of a classification problem in machine learning. The number of correct prediction as well as incorrect predictions is summarized with certain count values by all the classes. A multiple class confusion matrix obtained from actual vs. predicted values is shown in Figure. 7. The total number of test examples of any class is the sum of a corresponding row, which means the sum of true positive (TP) and false negative (FN) of that particular class. For example, total number of test examples of class 3 is the sum of $E_{31}+E_{32}+TP_3+E_{34}...+E_{3n}$. So, the total number of FN’s of a certain class would be the sum of all the errors in a row of that class without the TP. From the above example the false negative of class 3 would be $E_{31}+E_{32}+E_{34}...+E_{3n}$. In a similar manner, the total number of FP’s of a certain class would be the sum of all the errors in that class’s column without the True Positive. So, for class 3 the False Positive would be $E_{13}+E_{23}+E_{43}+...+E_{n3}$. And the total number of True Negative (TN) for a certain class is given by the sum of all the rows and columns by excluding that class’s row and column. Hence TN of class 3 is the sum of all elements by excluding the shaded area of Figure.7

n x n confusion matrix

		PREDICTED					
		classes	1	2	3	4 ...	n
ACTUAL	1	TP1	E12	E13	E14... ..	E1n	
	2	E21	TP2	E23	E24... ..	E2n	
	3	E31	E32	TP3	E34... ..	E3n	
	4	E41	E42	E43	TP4 ...	E4n	
	⋮	⋮	⋮	⋮	⋮	⋮	
	n	En1	En2	En3	En4... ..	TPn	

Figure. 7 Multiple class confusion matrix



The various important parameters which defined a confusion matrix are Accuracy, Precision, Sensitivity and Specificity. Each of these parameters are expressed briefly as under. Accuracy is the degree of measure of being in a correct state. It defined the closeness of the predicted values to the true values. The accuracy of a Confusion matrix's is expressed as the ratio of the sum of all correct classifications to the total number of correct classifications as well as the misclassification. The less the misclassification the more is the accuracy

$$\text{Accuracy} = (\text{Sum of the entire Diagonal elements}) / (\text{Sum of all elements}) \tag{1}$$

Precision is calculated as the ratio of true positive value to the sum of true positive value and the false positive value.

$$\text{Precision} = TP / (TP + FP) \tag{2}$$

For example, Precision of class 3 is expressed as, Precision3 = TP3 / (TP3 + E13 + E23 + E43... + En3)

Sensitivity is the portion of data's that are truly positive and are correctly classified. Sensitivity is given by the ratio of true positive value to the sum of true positive value and the false negative value.

$$\text{Sensitivity} = TP / (TP + FN) \tag{3}$$

Class 1 sensitivity is calculated as, Sensitivity 1 = TP1 / (TP1 + E12 + E13 + E14... + E1n)

The specificity is the proportion of actual negative data's that are correctly classified. Specificity is given by the ratio of true negative value to the sum of true negative value and false positive value.

$$\text{Specificity} = TN / (TN + FP) \tag{4}$$

Specificity of class 3 is calculated as, Specificity3 = TN3 / (TN3 + E13 + E23 + E43 + ... + En3)

V. IMPLEMENTATION AND SIMULATION RESULT

The speaker classifier model is implemented with a multi-layer perceptron neural network consisting of 13 inputs nodes, 310 hidden neurons and 14 output nodes for fourteen class classifications as shown in Figure 8. The neural network input data for classification is obtained from MFCC computation of fourteen different speakers. The basic 13 MFCC of each speaker are concatenated and given as input data to the neural network. The input size in matrix form is 13x15209 (Row x Column) and the neural network desire target consisting of 0's and 1's is set at 14x15209. The simulation result of fourteen class's pattern classification is shown in Figure. 9.

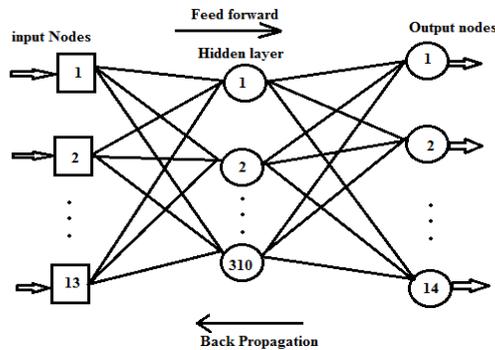


Figure. 8. Neural Network Classifier Model

	f1	m1	f2	m2	f3	m3	f4	m4	f5	m5	f6	m6	f7	m7
f1	1000 7.1%	0	0	0	0	0	0	0	0	0	0	0	0	0
m1	0	1427 9.4%	11	15	7	4	3	11	16	13	14	5	2	2
f2	0	0	1000	10	0	0	17	2	16	4	11	0	10	1
m2	0	0	0	1000	5.7%	2	4	0	0	0	0	1	0	1
f3	14	3	18	0	1100	7.2%	3	17	0	18	0	0	13	1
m3	0	1	17	19	7	300	17	15	9	2	3	6	21	12
f4	10	4	0	2	13	16	200	2	10	1	7	3	5	4
m4	0	0	0	0	0	0	0	1000	0.3%	2	3	0	0	1
f5	17	2	18	2	11	5	19	300	0.4%	1	6	1	7	1
m5	0	0	0	0	0	0	0	0	1000	0.0%	0	0	0	0
f6	0	0	0	0	0	0	0	0	0	1000	7.3%	0	0	0
m6	2	4	11	2	0	0	4	5	0	1	4	0	2	4
f7	14	5	14	0	19	3	7	2	21	9	5	2	143	1
m7	0	0	0	0	0	0	0	0	0	0	0	0	0	1000

Figure.9 Confusion Matrix of 14 class Classification

Class	Metric	Value
1	Precision of f1 (class1)	91.9%
1	Sensitivity of f1	93.5%
2	Precision of m1 (class2)	97.9%
2	Sensitivity of m1	92.7%
3	Precision of f2 (class3)	90%
3	Sensitivity of f2	90.8%
4	Precision of m2 (class4)	92%
4	Sensitivity of m2	97.3%
5	Precision of f3 (class5)	92%
5	Sensitivity of f3	91.9%
6	Precision of m3 (class6)	92.8%
6	Sensitivity of m3	88.1%
7	Precision of f4 (class7)	90.3%
7	Sensitivity of f4	91%
8	Precision of m4 (class8)	91.6%
8	Sensitivity of m4	97.3%
9	Precision of f5 (class9)	89%
9	Sensitivity of f5	91%
10	Precision of m5 (class10)	94.9%
10	Sensitivity of m5	98.4%
11	Precision of f6 (class11)	94.9%
11	Sensitivity of f6	94.6%
12	Precision of m6 (class12)	97.2%
12	Sensitivity of m6	93.2%
13	Precision of f7 (class13)	91.5%
13	Sensitivity of f7	90.8%
14	Precision of m7 (class14)	96.4%
14	Sensitivity of m7	94.4%

Figure. 10 Precision and Sensitivity Table



The overall Accuracy from the confusion matrix is found to be 93.1% and the Precision and Sensitivity for each of the speakers is extracted from the confusion matrix which is shown in Figure.10. Precision of all the classes score 90% and above except for f5 which is 89% and sensitivity too is above 90% and above but except for m3 which is 88.1%.The specificity of fourteen class classification is also calculated as per equation 4 and shown in the Figure.11. The score varies from 99.15% to 99.83% which is exceptionally high.

Sl	Specificity	TN/(TN+FN) %	Sl	Specificity	TN/(TN+FN) %
1	Specificity f1	99.30%	8	Specificity m4	99.38%
2	Specificity m1	99.78%	9	Specificity f5	99.15%
3	Specificity f2	99.21%	10	Specificity m5	99.67%
4	Specificity m2	99.47%	11	Specificity f6	99.55%
5	Specificity f3	99.26%	12	Specificity m6	99.83%
6	Specificity m3	99.45%	13	Specificity f7	99.37%
7	Specificity f4	99.33%	14	Specificity m7	99.71%

Figure.11 Specificity score of all the speakers

The classifier’s performance representation in terms of its parameters namely precision, sensitivity and specificity of all the fourteen classes is shown in Figure.12. The Receiver output characteristic (ROC) is another parameter which summarizes the overall performance of the classifier. This is a graph generated by plotting the True Positive Rate along the y-axis against the False Positive Rate along x-axis. The more the resultants lean sharply towards the true positive rate; the better is the classification of the classifier. The best validation performance is 0.0048851 at 1000 epochs which is shown in Figure. 13.

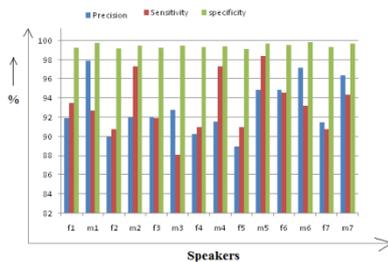


Figure. 12 Performance Parameters of Different speakers

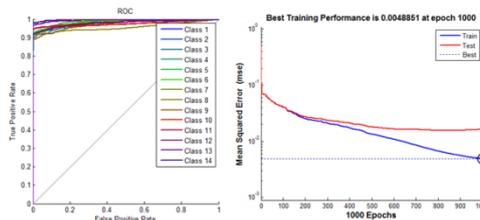


Figure.13 Classifier's ROC and Validation Performance

For real time testing of the classifier, the voice samples of all the speakers are collected again by uttering the same sentence which is used for while training the network. The MFCC features of each speaker are then computed again for testing the network. The score of diagonal elements which is extracted from the confusion matrix while testing with the speech data of different speakers is shown in Figure. 14. The score for the right speaker will always be maximum in the diagonal element. Hence by examining to whom this maximum score belongs, the corresponding unknown speaker can be readily recognized.

Sl No	Test speaker	Diagonal Score when Testing with f1	Diagonal Score when Testing with m1	Diagonal Score When Testing with f2	Diagonal Score When Testing with m2	Diagonal Score When Testing with f3	Diagonal Score When Testing with m3	Diagonal Score When Testing with f4
1	f1	(890)	0	69	0	43	2	28
2	m1	10	(1301)	34	57	0	18	15
3	f2	53	10	(400)	6	36	46	79
4	m2	5	0	6	(703)	2	3	0
5	f3	70	2	235	3	(940)	5	35
6	m3	0	8	17	23	4	(886)	9
7	f4	38	4	81	7	31	22	(767)
8	m4	0	15	65	142	6	0	0
9	f5	80	30	73	0	80	11	1
10	m5	0	7	0	8	2	0	5
11	f6	8	20	49	32	46	6	12
12	m6	0	6	41	0	0	22	23
13	f7	18	0	58	0	40	16	20
14	m7	0	11	0	0	0	26	19

Sl No	Test speaker	Diagonal Score when Testing with m4	Diagonal Score when Testing with f5	Diagonal Score When Testing with m5	Diagonal Score When Testing with f6	Diagonal Score When Testing with m6	Diagonal Score When Testing with f7	Diagonal Score When Testing with m7
1	f1	61	47	11	0	0	0	3
2	m1	23	27	6	0	17	0	64
3	f2	4	43	0	70	29	76	23
4	m2	73	0	72	0	47	0	5
5	f3	26	137	2	0	0	15	8
6	m3	46	23	2	17	14	27	48
7	f4	1	0	10	11	1	33	11
8	m4	(684)	0	158	0	8	9	0
9	f5	24	(643)	0	36	6	25	10
10	m5	12	10	(621)	0	0	0	18
11	f6	15	30	42	(1042)	0	9	6
12	m6	37	1	7	0	(663)	7	0
13	f7	3	102	7	0	58	(829)	15
14	m7	47	31	11	16	5	12	(899)

Figure. 14 The diagonal score of confusion matrix while testing the network with different speakers

VI. CONCLUSION

When sufficient number of training data is provided to the neural network, its performance in terms of correct classification is improved. Hence more number of data set is reserved for training the network. In this work the total data set is divided into two parts as training data set and testing data set. Out of the total data set 85 % of it is used for training and another 15% for testing the network. The classifier accuracy obtained is 93.1% with only 6.9% misclassification. The classifier's performance is acceptable good as the overall precision, sensitivity and specificity of all the fourteen classes score 93.02 %, 93.21% and 99.46% respectively. The classifier correctly recognized the speaker when testing with new inputs from the registered speakers. The maximum score in the diagonal element of confusion matrix when testing the network with new speech data always corresponds to the right speaker as shown in Figure 14.

REFERENCES

1. Abdelmajid Hassan Mansour¹, Gafar Zen Alabdeen Salh², HozayfaHayderZeen Alabdeen³ "Voice recognition Using back propagation algorithm in neural networks" International Journal of Computer Trends and Technology (IJCTT) ISSN: 2231-2803 volume 23 Number 3, pp 132-139 May 2015.
2. AbhinavAnand, RuggeroDonidaLabati, MadasuHanmandluy, Vincenzo Piuri, Fabio Scotti, "Text-Independent Speaker Recognition for Ambient Intelligence Applications by Using Information Set Features" IEEE International Conference on Computational Intelligence and virtual Environments for Measurement system and Applications (CIVEMSA), Year: 26-28 June, 2017, Pages: 30 – 35.
3. Bhargab Medhi¹, Prof. P.H.Talukdar² "Assamese Speaker Recognition Using Artificial Neural Network" International Journal of Advanced Research in Computer and Communication Engineering Vol. 4, Issue 3, pp 321-324, March 2015.
4. Dany Isha¹, Antoine Abche² and Elie Karam², Georges Nassar³, Dorothée Callens³, "A Text-dependent Speaker-Recognition System" 2017 IEEE International Instrumentation and Measurement Technology Conference (I2MTC), Year: 22-25 May 2017 page 1-6, IEEE Conference Publications.
5. E.C. Gordon, Signal and Linear System Analysis. John Wiley & Sons Ltd., New York, USA, 1998.
6. J Sirisha Devi "Language and Text Independent Speaker Recognition System using Artificial Neural Networks and Fuzzy Logic" International Journal of Recent Technology and Engineering (IJRTE), Volume-7, Issue-6, pp 327-330, March 2019.
7. Jorge MARTINEZ*, Hector PEREZ, Enrique ESCAMILLA, Masahisa Mabo SUZUKI, "Speaker recognition using Mel Frequency Cepstral Coefficients (MFCC) and Vector Quantization (VQ) Techniques" CONIELECOMP 2012, 22nd International Conference on Electrical Communications and Computers, 27-29 Feb. 2012 pages: 248 - 251, IEEE Conference Publications.
8. Khan Suhail Ahmad¹, Anil S. Thosar², Jagannath H. Nirmal³ and Vinay S. Pande⁴ "A Unique Approach in Text Independent Speaker Recognition using MFCC Feature Sets and Probabilistic Neural Network" 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR) Year: 5-7 June 2015 Pages: 1 - 6, IEEE Conference Publications.
9. Koustav Chakraborty, Asmita Tale Prof. Savitha Upadhyaya, "Voice Recognition Using MFCC Algorithm" International Journal of Innovative Research in Advanced Engineering (IJIRAE) Volume 1 Issue 10 (November 2014), page 158-161.
10. Md. Ali Hossain¹, Md. Mijanur Rahman², Uzzal Kumar Prodhon³, Md. Farukuzzaman Khan⁴ "Implementation Of Back-Propagation Neural Network For Isolated Bangla Speech Recognition" International Journal of Information Sciences and Techniques (IJIST) Vol.3, No.4, pp 1-9 July 2013.
11. N.AYSHWARYA¹, G.LOGESHWARI², G.S.ANANDHA MALA³ "FEED FORWARD BACK PROPAGATION NEURAL NETWORK FOR SPEAKER INDEPENDENT SPEECH RECOGNITION", International Journal of Industrial Electronics and Electrical Engineering, ISSN: 2347-6982 Volume-2, Issue-8, pp 36-39, Aug.-2014.
12. Neha Chauhan "Speaker recognition using pattern recognition neural network and feedforward neural network, International Journal of Scientific & Engineering Research, Volume 8, Issue 3, pp 1444-1446, March-2017.
13. Parvati J. Chaudhary¹, Kinjal M. Vagadia² "A Review Article on Speaker Recognition with Feature Extraction International Journal of Emerging Technology and Advanced Engineering, Volume 5, Issue 2, February 2015 page 94-97.
14. Roger Achkar*, Mustafa El-Halabi*, Elie Bassil*, Rayan Fakhro*, Marny Khalil* "Voice Identity Finder Using the Back Propagation Algorithm of an Artificial Neural Network" Complex Adaptive Systems, Publication 6, Conference Organized by Missouri University of Science and Technology 2016 - Los Angeles, CA.



15. S. S.Wali, S.M. Hatture, S. Nandyal “MFCC Based Text-Dependent Speaker Identification Using BPNN” International Journal of Signal Processing System VOL.3, No.1, pp 30-34, June 2015.
16. Siddhant C. Joshi¹, Dr. A.N.Cheeran², “MATLAB Based Back-Propagation Neural Network for Automatic Speech Recognition” International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering, Vol. 3, Issue 7, pp 10498-10504, July 2014.
17. Varun Sharma, Dr. P K Bansal, “A Review On Speaker Recognition Approaches And Challenges” International Journal of Engineering Research & Technology (IJERT) Vol. 2 Issue 5, May – 2013 page 1581-1588.
18. XinxingJing¹, Jinlong Ma², Jing Zhao³, Haiyan Yang⁴ “Speaker Recognition Based on Principal Component Analysis of LPCC and MFCC” 2014 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC) 5-8 Aug. 2014, Pages: 403 - 408 IEEE Conference Publications.
19. Yi Wang, Dr. Bob Lawlor, “Speaker Recognition Based on MFCC and BP Neural Networks” 28th Irish Signals And systems Conference Year: 20-21 June 2017, Pages: 1 - 4, IEEE Conference publication.
20. Zhang Wanli, Li Guoxin, “The Research of Feature Extraction Based on MFCC for Speaker Recognition” Proceedings of 2013 3rd International Conference on Computer Science and Network Technology, 12-13 Oct. 2013, Pages: 1074 - 107IEEE Conference Publications.