International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-8 Issue-11, September 2019

To Improve Voice Recognition System using GMM and HMM Classification Models

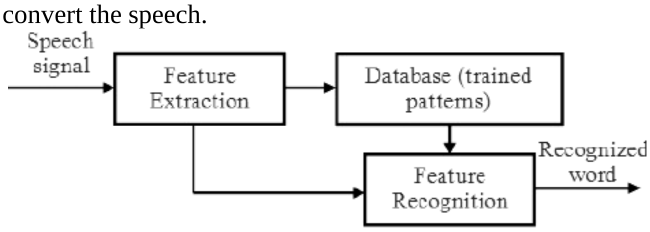
Sonali Nemade, Yogesh Kumar Sharma, Ranjit D. Patil

***Abstract:** In this paper, the researcher study automatic speech recognition technology for the individual. We propose a new voice recognition system using a hybrid model GMM-HMM. HMM and GMM is a non-linear classification model. Each state in an HMM can be thought of as a GMM. HMM is consider observation for state. It is also known as time series classification model. In this model, samples have been trained independently and parameters consider jointly which provides better performance than other classification models. Speech recognition system consider two types of learning patterns such as supervised learning and unsupervised learning. In this context speaker dependent and speaker independent used for identifying the efficient and effective voice. In this paper researcher considered supervised learning model for recognize efficient voice. This new voice recognition system identifies incorrect phonemes and verifies the correctness of voice pronunciation. Using the GMM-HMM hybrid model produces better performance and effectiveness of voice.*

***Keywords :** GMM classifier, HMM classifier, MFCC, deep neural network, artificial Intelligence..*

I. INTRODUCTION

Voice recognition system evaluates the individual voice, such as the flow and frequency of their voice. Voice recognition is also known as speaker recognition. Voice recognition is a program to interpret spoken commands. In speech recognition use Artificial Intelligence and deep machine learning techniques. Now a day's voice recognition implements on a computer or machine with ASR (automatic speech recognition) software programs. ASR includes the user to train the ASR program to recognize individual voice so that it can more effectively and accurately convert the speech.

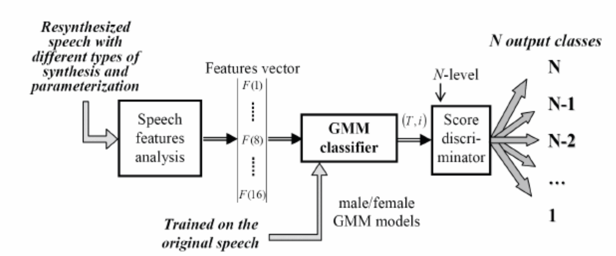


Automatic speech recognition system use pattern matching digital audio files of spoken words. Voice recognition is also called as automatic speech recognition (ASR) system. In this process interpreting human speech on a computer.

Manuscript published on 30 September 2019.
***Correspondence Author(s)**
Mrs. Sonali Nemade, Assistant Professor Dr. D. Y. Patil Arts , Commerce and Science College Pimpri , Pune.
Dr. Yogesh Kumar Sharma, Associate Professor and Research Coordinator, Department of Computer Science and Engineering, Shri JJT University, Jhunjhunu, Rajasthan
Dr.Ranjit D. Patil, Vice-Principal and H.O.D in Dr. D.Y. Patil ACS College, Pimpri, Pune

In ASR system audio speech into an error-free and word transcription of that speech.

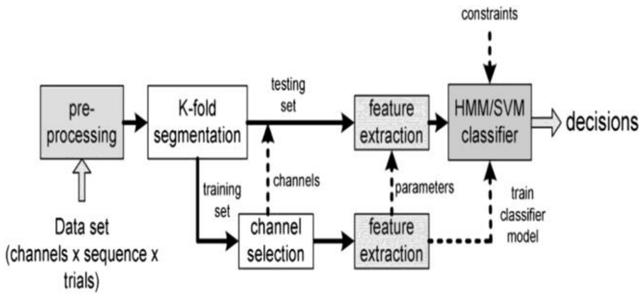
A. *GMM classification model:*



Block diagram of GMM classification model

The Gaussian Mixture Model Classifier (GMM) is a supervised and probabilistic learning classification model. In this classification model to classify the N-dimensional signals. GMM is representing a distributed subpopulation within the overall population. GMM used two-component first data points and the second component is equi-probability. In data point component of the GMM model used a mixture of a finite number of Gaussian distribution. The Gaussian Mixture Model Classifier is similarly to the k-means clustering algorithm with a k data set. Another way to implementing the GMM classification model using the Expectation maximization algorithm to trained voice data.

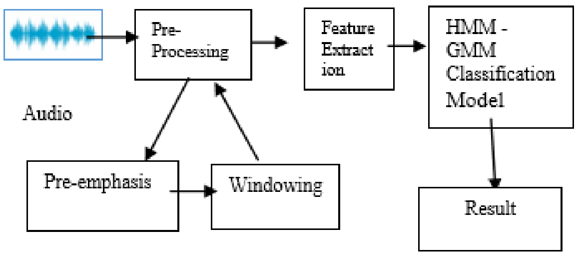
B. *HMM classification model:*



Block diagram of HMM classification model

HMM, classifier used a different approach first separate data sets into class, second train each HMM class and finally test set compare using likelihood. HMM, classification is used statistical pattern recognition. Hidden Markov Model (HMM) used hidden state representation is a probabilistic graphical model. Each state sets the parameters for training data according to a training rule. For the implementation of HMM classifier for voice recognition feature or parameters are broken up into a sequence of quasi-stationary segments. These segments consider as independently and in isolation. In the HMM classifier, the frame rate is large for a speech given frame.

II. PROPOSED METHODOLOGY



Black Diagram of Propose Enhance Voice Recognition System

A. *GMM-HMM Hybrid Classification Model:*

HMM-GMM hybrid system can be used different attributes such as stochastic modeling for acoustic variation of voice, the system can handle time series effectively and it calculates very efficiently. Only HMM classifier is used for voice recognition some limitations are presented for example HMMs trained with maximum likelihood rules provide less discriminative power. The HMM consider GMM in each state of HMM. State are usually considered as separate GMMs for the transition matrix. In transition matrix learned from training data and determine probabilities from one state to another. The following are the parameters used to enhance voice recognition system.

Pre-emphasis:

In enhanced voice recognition system remove the noisy data by using different techniques such as DC offset removal, silence words removal.

Windowing:

In the proposed model, the next step is windowing. Each frame is to minimize the discontinuities of signals beginning to end of the frame. For spectral analysis used window function of each frame. In this step window as

$$b(m) = a(m)w(m), \qquad 0 \leq m \leq M-1$$

Here M is the number of samples in each frame.

Feature extraction

The next step is feature extraction is converted voice audio/video signal to feature vector coefficients for correct recognition. In the feature, extraction identifies speaker information using frequency band and transmission channel. For feature, extraction used the MFCC extraction technique and DWT extraction technique for better performance of voice recognition.

Pattern classification

In the pattern recognition process consider different large individual groups for voice recognition. Pattern recognition compares unseen test patterns for each voice class and comparing the similarity between them. In GMM-HMM hybrid model provide better performance. In this pattern, classification identified the speaker and verification using a hybrid model. In this process used deep learning in artificial neural network models for recognize voice.

III. DISCUSSION

In this paper, voice is recognized with background noise. The author proposed a speaker-independent model for voice recognition. In this model no need to training phase for each user. In this proposed model HMM-based model produced a better result than the GMM but used hybrid GMM-HMM model produced a better result than the other. GMM-HMM hybrid model using good feature extraction techniques such as MFCC and DWT for increased the voice recognition rate. However, by computing parameters based on speakers' pronunciation, the system can be speaker-dependent.

There should be proper comments of the reviewers for the purpose of acceptance/ rejection. There should be minimum 01 to 02 week time window for it.

IV. CONCLUSION

In this paper, we investigate a hybrid model GMM-HMM. In this method system trained and tested. The GMM-HMM systems use GMMs as observation functions for HMMs. The HMM of each class is modeled by one transition state, which means it is a GMM. The best number of mixtures is evaluated on the validation fold and the performance of the mixture yielding the best total performance will be shown.

REFERENCES

1. C. Clavel, T. Ehrette, and G. Richard, "Events detection for an audio-based surveillance system," in Proceedings of the 2005 IEEE International Conference on Multimedia and Expo (ICME), Amsterdam, The Netherlands, Jul. 2005, pp. 1306–1309.
2. J. Schr¨oder, S. Goetze, V. Gr¨utzmacher, and J. Anem¨uller, "Automatic acoustic siren detection in traffic noise by part-based models," in Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Vancouver, Canada, May 2013, pp. 493–497.
3. P. Laffitte, D. Sodoyer, C. Tatkeu, and L. Girin, "Deep neural networks for automatic detection of screams and shouted speech in subway trains," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, Mar. 2016, pp. 6460–6464.
4. S. P¨aßler and W. Fischer, "Food intake monitoring: Automated chew event detection in chewing sounds," IEEE Journal of Biomedical and Health Informatics, vol. 18, no. 1, pp. 278–289, 2014.
5. H. Bourlard and C. J. Wellekens, "Links between Markov models and Multilayer Perceptron" IEEE Trans. on Pattern Analysis and Machine Intelligence, vol. 12, no. 12, pp. 1167– 1178,1990. Forman, G. 2003.
6. Utpal Bhattacharjee, "A comparative study of LPCC and MFCC feature for the recognition of Assamese phonemes"IJERT, ISSN:2278-0181, Vol.2 issue 1, January-2013
7. Zhu PANG et al. Discriminative training of GMM-HMM acoustic model by RPCL learning, Front. Electr. Electron. Eng. China 2011, 6(2): 283–290
8. J. Schr¨oder, J. Anem¨uller, and S. Goetze, "Classification of human cough signals using spectro-temporal Gabor filterbank features," in Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), Shanghai, China, Mar. 2016, pp. 6455–6459.
9. S. Matos, S. S. Birring, I. D. Pavord, and D. H. Evans, "Detection of cough signals in continuous audio recordings using hidden Markov models," IEEE Transactions on Biomedical Engineering, vol. 53, no. 6, pp. 1078–1083, 2006.
10. J. Schr¨oder, M. Brandes, D. Hollosi, J.Wellmann, M. Wittor f, O. Jung, V. Gr¨utzmacher, and S. Goetze, "Foreign object detection in tires by acoustic event detection," in DAGA 2015, Nuremberg, Germany, Mar. 2015, pp. 1266–1269.

AUTHORS PROFILE

Mrs. Sonali Nemade, M.Sc. Comp. Sci. M.Phil Comp. Sci. Pursuing Ph.D. Assistant Professor Dr. D. Y. Patil Arts , Commerce and Science College Pimpri , Pune.

Retrieval Number: K21780981119/19©BEIESP
DOI: 10.35940/ijitee.K2178.0981119
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)

Dr. Yogesh Kumar Sharma, MCA Comp. Sci., Ph.D., Associate Professor and Research Coordinator, Department of Computer Science and Engineering, Shri JJT University, Jhunjhunu, Rajasthan

Dr.Ranjit D. Patil, M.C.S, M.Phil. NET and Ph.D. Vice-Principal and H.O.D in Dr. D.Y. Patil ACS College, Pimpri, Pune