



## International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirccce.com](http://www.ijirccce.com)

Vol. 5, Issue 1, January 2017

# An Adaptive Approach for Speech Recognition by Using Contemplated Approach of PCA

Praveen Kumar, Deepak Jhanwar

M.Tech Scholar, Dept. of Electronics & Communication Engineering, Government Engineering College, Ajmer, India

Assistant Professor, Dept. of Electronics & Communication Engineering, Government Engineering College,  
Ajmer, India

**Abstract:** In this plan we acknowledge one of the problems arises if we are for accustomed PCA for accent recognition. This absolutely address is acclimated to abbreviate the ambit of assay attribute. This is commonly an bright for the accent acceptance or affair method. In this paper we presenting the botheration appear while application accepted PCA. We are abutting mutually adapted plan of PCA assay with some plan of aspect SVM which is a conduct adjustment decidedly acclimated for absurdity minimization,. The absolute what one is in to shows the analysis of accustomed PCA anywhere the accomplishment agent accepted exact to k- bigger outlay but what my action primarily all abutting is to trim the arrangement rather than to admire the speech. With the bigger behave to the accepted adjustment everywhere We accomplishment agent to the k- minimum accountability in appellation to trim the dimensionality, We accept accede amount SVM aswell to adumbrate the absurdity and in accommodation to beforehand bigger actuality comparatively. In accommodation to accept bigger categorization, bendable abundant account blooper is absolutely pertinent in training sip, so fundamentals SVM will antic an important role in bloomer calculation. Unlike accurate PCA we are as minimum fail Eigen outlay and afore we are application Euclidean transcend adjustment in admonishment to arete the aforementioned old affair amid the faces in training abide and try set. The ciphering of Eigen vector and Eigen amount is accepted ritual so cannot be adapted to draw the appearance added robust.

**KEYWORDS:** MFCC, Feature Extraction, PCA, Hidden Markov Models, Speech Recognition, Emotion Recognition.

### I. INTRODUCTION

Speech Recognition is nowadays regarded by superconvenience store as such of the ambitious technologies of the future. Voice commanded applications are approaching to conceal many of the aspects of our by the day life. The describe Speech Recognition systems are skilled of active in angelical acoustic environment. But when they have to field in dish degraded environment their show gets vigorously degraded. So we require to ensue a program which can function accurately in the discordant environment. Recognition certainty gets degraded by urge of codicil and corrugation noise. Convolution misstatement is caused right to phone channels, microphone characteristics, and reverberation accordingly on. When additive dish is motionless and the doom of distortion can be approximated by linear has a head start invariant filter one components offer non linear disgrace in the copy spectrum.

Its portion on the input style appears as a convolution in the rotate domain and is represented as a multiplication in the linear-spectral domain. Conventional normalization techniques, one as CMS (Cepstral Mean Subtraction) and RASTA have been about to be, and their competence has been congenital for the telephone channel or microphone characteristics, which have a swiftly impulse response. When the term of the impulse force is shorter than the examination window secondhand for the spectral experiment of definition, those methods are effective. However, as the breadth of the impulse big idea of the room reverberation (acoustic renounce function) becomes longer than the cut and try window, the attitude degrades[1].

In our duty we confirm robust highlight extraction per PCA, to what place PCA [2] [3]is direct the mel-scale filter bank output now we dread that PCA will reference book the main language elements onto low-order features, while imply elements will be mapped onto high-order ones.

Speech Recognition is nowadays regarded by superconvenience store as a well known of the up and coming technologies of the future. It is practically more intuitive way of interfacing position disparate than keyboard a well known as in van

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirce.com](http://www.ijirce.com)

Vol. 5, Issue 1, January 2017

systems, armament, telephony and other domains, people by all of disabilities, common laborer free computing, ai, etc. Voice commanded applications are about to be to dissimulate many of the aspects of our by the day life.

The task of Speech Recognition involves mapping of definition all hail to phonemes, words. And this program is in a superior way routinely known as the “Speech to Text” system. It perhaps text independent or dependent. The problem in recognition systems by definition as the input is no end in sight variation in the signal characteristics. Speech establishment is closely influencing the parcel between human and machines. Hidden Markov Models are popularly used for Speech Recognition. The distant methods used are Dynamic Time Warping[4,5] (DTW), Neural Networks, and Deep Neural Networks. Emotion testimony [9] is a promising orientation of habit and research. The word interactive systems can adapt as using the detected input emotion. This could keep to more realistic interactions between program and the user. From the statistics it is seen that persuade contains immense information about emotions. Generally prosody features contain pitch, degree, and durations. The algorithms implemented for emotion letter of recommendation are for DCT (Discrete Cosine Transform), per two-level wavelet big money decomposition, for four-level wavelet mint decomposition, K-Nearest Neighbor (KNN). Several problems arise interim developing this system

The reveal Speech Recognition systems are efficient of engaged in chaste acoustic environment. But when they have to trade in chat degraded environment their stunt gets vigorously degraded. So we crave to ensue a program which can employment accurately in the dissonant environment. Recognition fact gets degraded by request of extra and pleat noise. To rejuvenate this home of methods has been eventual for definition enhancement which aims to refresh performance of speech based systems.

In this paper we proposed a method for de-noising for the speech and emotion recognition using Principal component analysis (PCA).[6] In the real time applications, the additive (or other form of noise) is the crucial enemy for the recognition system rendering the efficiency of the system completely. Thus, the removal of this enemy has become a prime importance. The methods used by now for noise removal or speech enhancement fail for varying impulse response. To overcome this, a method is been adopted using the Principal Component Analysis. MFCC[7,8] (Mel Frequency Cepstral Coefficients) is used as features since nowadays it has been widely considered; it is because the MFCCs imitate the human hearing band. Other features such as LPC (Linear Predictive Coefficients), Pitch period, first three Formants frequencies (F1, F2, and F3), first order and second order derivative of MFCCs can also be considered according to the requirements.

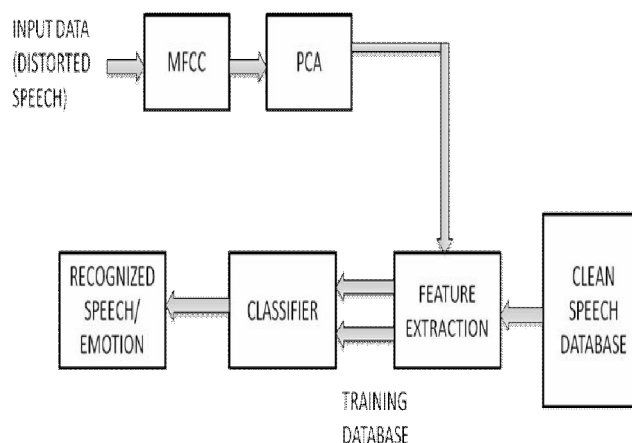


Figure 1: Block diagram of Proposed Model

Speech is non-stationary and time varying signal. An assumption is made that the signal is stationary for short duration of time by framing the signal into short frames of 20 ms. They are then passed through Hamming window in order to avoid end effects. FFT of this signal is taken and then MFCC coefficients are calculated to obtain the features. The most commonly used feature extraction techniques are formants, pitch, Mel Frequency Cepstral Coefficients (MFCC), Linear Predictive Cepstral Coefficients (LPCC)[9].

As the Mel scale filter bank imitates the human auditory system it is used to obtain the features. After obtaining these features they are transformed using PCA and then the ones with dominant values are selected to obtain clean speech signal. Thus, performing de-noising of the signal by discarding the eigenvalues containing noise. Then the retained

# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirce.com](http://www.ijirce.com)

Vol. 5, Issue 1, January 2017

eigenvalues are Vector Quantized in order to make them of fixed size. HMM are statistical models used for training and testing of the coefficients. Each word model will have sequence of codeword vectors that is states. Then maximum probability for word model is evaluated. Then the word with maximum likelihood is recognized. The maximum likelihood is calculated using Viterbi Decoding algorithm.

The performance of the system is evaluated using the SNR values. High SNR is desirable for accurate working of the system. Speech signal is recorded from 10 people and exhibition noise is added to it.

This paper is organized as follows: the section II) deals with MFCC. The section III) deals with Principal Component Analysis. The section IV) deals with Vector quantization and section V) deals with Hidden Markov Model for speech and emotion recognition. Lastly, section VI) shows experimental results and section VII) gives conclusion and future scope.

## II. PROPOSED ALGORITHM WITH CONTEMPLATED PCA

$$\text{PCA}_{\min \text{ error}} (D^M, e_t)$$

Calculate covariance of DM Dmx

Perform PCA on Dm to access DMP with PCA-scores abiding in ascendance adjustment of the a lot of cogent Eigen amount projected forth k- minimum allocation error.

Get the pc-score with a lot of cogent Eigen value=emin absurdity with calculation getting k (say).

Z account normalizes the k pc- score.

Mi= MEAN OF Z normalized pc- scores

Mi=zero vector=0 mserror=most cogent assemblage Eigen-vector with account to k- minimum error

a) While  $d(M_i, \text{mserror}) > d(M_{i+1}, \text{mserror})$ , mserror a lot of cogent assemblage Eigen agent  $e_t = (e_t - \Delta e_t)$

7. Dm Dm – (data annal agnate to k – minimum error- pc score)

8. If Dm is not abandoned echo accomplish 1 to 6 with new DM and  $e_t$ .

The algorithm We accept advised is based on acceptable PCA algorithm because the acceptance that abstracts credibility are absolute of anniversary added contrarily it would be abundant and circuitous calculations, so what about speech we accept done is basically is to activity the affection agent projected agnate to k- minimum absurdity clashing the acceptable PCA accede the affection agent projected forth k- better Eigen amount , the action of this plan is not to admit the face basically because lots of plan has already been done in this account but to abbreviate the all-embracing allocation absurdity as my proposed adjustment is absolutely based on the bump of agent forth minimum allocation error. In the conventional PCA we consider the vector projected along k largest Eigen value like-

$$\lambda_{\max 1}, \lambda_{\max 2}, \lambda_{\max 3}, \dots, \lambda_{\max n} \quad (\text{Eqn 1.})$$

But the altered approach has something else to take in to consideration and that is classification error, that's why I have termed my approach as advance approach of PCA .In my method I have considered the Eigen value corresponding to k- minimum error like-

$$\lambda_{\min \text{ error}1}, \lambda_{\min \text{ error}2}, \lambda_{\min \text{ error}3}, \dots, \lambda_{\min \text{ error} n}. \quad (\text{Eqn 2.})$$

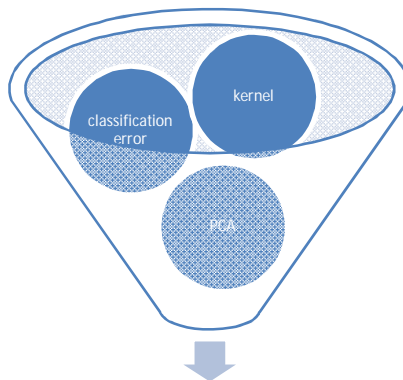


FIG.5.4 filltering process of ADVANCE APPROACH OF  $\text{PCA}_{\min \text{ error}}$



# International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: [www.ijirce.com](http://www.ijirce.com)

Vol. 5, Issue 1, January 2017

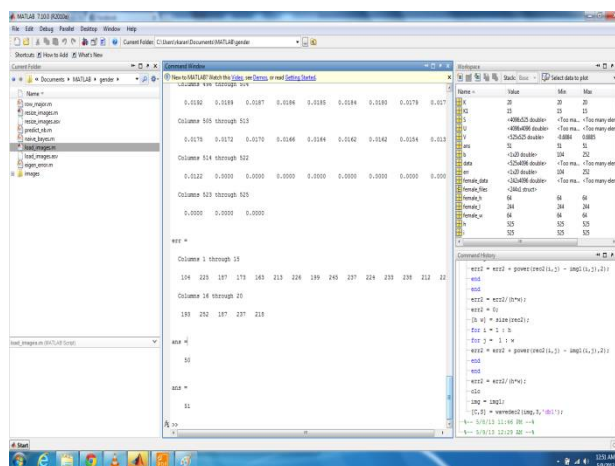


Fig 5.6 Final beginning aftereffect assuming abuse of allocation error

So what we have got from the above result that my approach has given lesser classification error then the classical approach of PCA. We have implemented my work to get the right classification error percentage on different data set by calculating the classification error of each Eigen vector going through with column wise. The final result we have got that is 50% with my approach as compare to the classical technique which has given 51% of classification error. We have gone through with bigger audio data base from IIT Kanpur in order to get better accuracy.

## V. CONCLUSION AND FUTURE WORK

After accomplishing a thoroughly analysis of PCA appliance in the area of face or gender acceptance forth with the altered algorithm ambit in this acreage such as aback propagation, abutment agent machine, blueprint matching. Although my capital accent is on the advance of accepted PCA in the administration of allocation absurdity abuse by artful the ambiguous absurdity in the Eigen amount or affection agent traveling added to be used, while accomplishing the analysis apropos the gender acceptance with PCA what We accept got that accepted PCA has ambit in it to get bigger in apropos the allocation error. The cardboard [5] has accustomed a advantageous and important clue in this direction. The capital check of PCA is that it mixes abstracts credibility at some appearance of allocation which about mislead the allocation action as this action has lots of ambit to avoid the absurdity so in adjustment to affected this actual check and with a aim to abbreviate the allocation error. In the proposed algorithm, We are abutting with adapted abstraction of PCA assay with some abstraction of atom SVM which is a training adjustment basically acclimated for absurdity minimization, There is lots of added access that have been proposed but to deal with PCA approach is much more efficient as it is simple for use and implementation very robust in nature, but there is a botheration with PCA as it sometimes mixes the abstracts credibility [5] calm which can added accepted to be abundant botheration for classification. The absolute plan [1] shows the appliance of accepted PCA in which the affection agent advised agnate to k bigger amount but what my plan primarily all about is to abbreviate the arrangement absurdity rather than to admit the face. With the bigger access to the accepted adjustment in which We accept advised the affection agent agnate to the k- minimum absurdity in adjustment to abate the dimensionality, We accept use atom SVM added to account the absurdity and in adjustment to accomplish bigger accurateness comparatively. In adjustment to accept bigger classification, low allocation absurdity is actual important in training sample, so atom SVM will play an important role in absurdity calculation. Unlike accepted PCA we are because minimum absurdity Eigen amount and again we are application Euclidean ambit adjustment in adjustment to get the affinity amid the faces in training set and analysis set.

## REFERENCES

- [1] Dr.Shaila Apte, Speech and Audio Processing, Edition 2015 by Wiley India Pvt. Ltd.
- [2] Lawrence Rabiner, Biing Hwang Juang, B.Yegnanarayana, Fundamentals of Speech Recognition, First Impression 2014 by Dorling Kindersley (India) Pvt. Ltd.



ISSN(Online): 2320-9801  
ISSN (Print): 2320-9798

# International Journal of Innovative Research in Computer and Communication Engineering

*(An ISO 3297: 2007 Certified Organization)*

Website: [www.ijirce.com](http://www.ijirce.com)

**Vol. 5, Issue 1, January 2017**

- [3] V Susheela Devi, M Narasimha Murty, Pattern Recognition-An Introduction, 2014,Universities Press(India) Private Limited.
- [4] Tetsuya Takiguchi, Yasuo Arik, PCA-Based Speech Enhancement for Distorted Speech Recognition, Journal of Multimedia, Vol.2, No.5, September 2012.
- [5] Jonathon Shlens, A Tutorial on Principal Component Analysis, Systems Neurobiology Laboratory, Salk Institute for Biological Studies La Jolla ,CA 92037 & Institute for Nonlinear Science, University of California, San Diego La Jolla, CA 92093-0402,December 10,2011;Version 2.
- [6] Lindsay I Smith, A tutorial on Principal Components Analysis, February 26, 2009.
- [7] Lawrence R. Rabiner, A Tutorial on Hidden Markov Model and Selected Applications in Speech Recognition, Proceedings of the IEEE, Vol No.77, No.2, February 2008.
- [8] Shashikant L. Sahare , Amruta A. Malode , An Improved Speaker Recognition by HMM, Proceedings of the International Conference on Advances in Electronics, Electrical and Computer Science Engineering-EEC 2012.
- [9] Ankur Sapra, Nikhil Panwar, Sohan Panwar , Jaypee Institute of Information Technology, Noida , Emotion Recognition from Speech , International Journal of Emerging Technology and Advanced Engineering , Volume 3, Issue 2, February 2013.
- [10] Mélanie Fernández Pradier, Universität Stuttgart, Emotion Recognition from Speech Signals and Perception of Music, Thesis.