Scientific Journal of Impact Factor(SJIF): 3.134

e-ISSN(O): 2348-4470
p-ISSN(P): 2348-6406

International Journal of Advance Engineering and Research Development

Volume 2,Issue 5, May -2015

DESIGN & DEVELOPMENT OF CONTINUOUS DENSITY HMM (CDHMM) ISOLATED HINDI SPEECH RECOGNIZER

Satish Kumar[1] ,Prof. Jai Prakash[2]

[1]*Research Scholar, Mewar University, Rajasthan , India, satish_suryan@yahoo.com. He is presently working as Lecturer(S.G) in Kasturba institute of Technology, Pitampura, Delhi—110088.*
[2]*Visiting Professor, Mewar University, Rajasthan, India , drjp27@yahoo.co.uk*

ABSTRACT: *This paper describes the insight of the design & development of a Proposed Hindi Speech Recognizer based on the continuous density hidden Markov model (CDHMM) . Here we have proposed a new recognizer which have been used with continuous density hidden Markov modeling to get a proposed CDHMM Hindi Speech Recognizer. The multidimensional Mel frequency cepstral coefficients(MFCC) speech vectors are extracted from raw speech for every given word that are used as a sequence of observation vectors, are uniformly segmented into 6 states . For each state Gaussian Mixture Model(GMM) parameters such as a Mean(* μ_{jk} *) & Covariance (* $\sum_{jk}$ *) matrix & number of Mixtures(* C_{jk} *) are calculated and simultaneously hidden Markov Model parameters such as* $\lambda = (A, B, \pi)$ *are calculated to prepare a GMM-HMM Model known as λcdhmm ,* $\lambda = \{C_{jk}, \mu_{jk}, \sum_{jk}\}$*. Here ,Q=6, Mixture Components(K or M)=16, The covariance matrix used can be full or diagonal type matrix. Investigaions are done in this paper to find the optimal number of Gaussian mixture components that gives maximum accuracy in the context of Hindi speech recognition system. The results of the experimentation have shown that Proposed CDHMM Speech Recognizer gives maximum performance when no. of Gaussian Mixture Model Components used is 16. This method is more powerful and efficient as compared to discrete Speech Recognizer*

Keywords: GMM, Mixture Components, Viterbi, CDHMM, Speech Recognizer.

I. INTRODUCTION

The observations are continuous signals or vectors in some applications. Vector quantization technique may be used to convert a continuous signal to discrete symbols. There can be degradation in doing the discretization. In order to avoid such as a situation continuous observation densities may be used with HMMs. Such Markov models are known as continuous density Hidden Markov Models (CD HMM). A continuous observation density may be used by putting some restrictions on the form of model pdf. That type or form of pdf shall be used whose re-estimation procedures are available in the literature. In other words we can say that a pdf may be used whose parameters may be re-estimated consistently.

A probability density function (pdf) for which reestimation proceeding is formulated is given below.

$$b_j(o) = \sum_{k=1}^{M} C_{jk}.g(o, \mu_{jk}, \Sigma_{jk})$$

This is called M component Gaussian mixture density. A GMM is Gaussian Mixture Model represented by weighted sum of M component Gaussian densities. Where C_{jk} is mixture weight or coefficients of the K^{th} component in State j [1][2].

μ_{jk} is the Mean of K^{th} component in state j . $\sum_{jk}$ is the covariance matrix of the K^{th} component in state j and the mixture weight C_{jk} satisfy the stochastic constraint so

@IJAERD-2015, All rights Reserved

$$\sum_{k=1}^{M} C_{jk} = 1 \qquad 1 \le j \le N$$

Continuous observation vector 'O' is a D dimensional continuous valued data vectors (speech vectors). Each component density is a Gaussian function given by

$$g(C_{jk}, \mu_{jk}, \sum_{jk}) = \frac{1}{(2\pi)^{D/2} \left|\sum_{jk}\right|^{\frac{1}{2}}} \exp\{\frac{-1}{2}(O-\mu_{jk})' \sum_{jk}^{-1}(O-\mu_{jk})\}$$

A GMM is characterized by mean vectors, covariance matrix and mixture weights from all the components densities and is denoted by U_{jk}

$$\lambda = \{C_{jk}, \mu_{jk}, \sum_{jk}\}$$

There are various types of GMMs depending upon the type of covariance matrix such as it can be full rank matrix or diagonal matrix.

2 Design of GMM-HMM based speech recognizer

2.1 Calculation of Forward variable alpha 'α'

The calculation of alpha 'α' the forward variable is done through forward algorithm described below . In designing CDHMM/GMM-HMM the observation vectors are considered as continuous signals.

The observation symbol probability distribution $B = \{b_j(k)\}$ (discrete case) is replaced by a pdf and is given by $B = b_j(x).$ Where x is a continuous observation sequence.

Forward Algorithm

Let us define a forward probability $\alpha_n(S_j)$ which is given by

$$\alpha_n(S_j) = P(X_1, X_2 .. X_n, S(n) = S_j / \lambda$$

where P is the probability of observing the observation sequence $X_1, X_2 .... X_n$ and being in state S_j at time n given the model λ .[3].

* Initialization

$$\alpha_o(S_i) = 1$$

$$\alpha_o(S_j) = o \qquad if \quad S_j \ne S_i$$

* Recursion

$$\alpha_n(S_j) = \sum_{i=1}^{S} \alpha_{n-1}(S_i).\ a_{ij}.\ b_j(X_n)$$

* Termination

$$P(X/\lambda) = \alpha_N(S_E) = \sum_{i=1}^{S} \alpha_N(S_i).\ a_{iE}$$

The Terills diagram for forward recursion is shown in Figure 1.

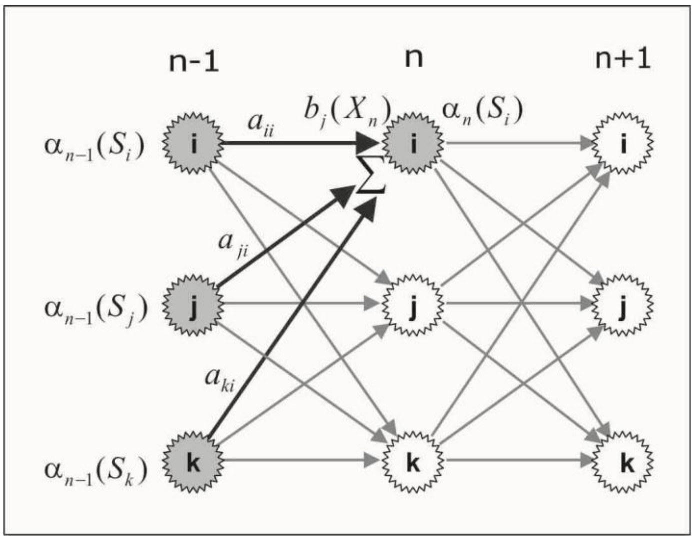

Figure 1 : Showing Forward Recursion

2.2 Calculation of Backward variable 'β'.

To estimate the state occupation probabilities we shall define another set of probabilities called the backward probabilities 'β' and is given below

$$\beta_n(S_j) = p(X_{n+1}, X_{n+2}, X_N \mid S(n) = S_j, \lambda)$$

The probability of future observations given a the HMM is in state S_j at time '*n*'. These can be computed recursively going backwards in time *n* [3].

Backward Algorithm

* Initialization $\beta_N(S_i) = a_{iE}$

@IJAERD-2015, All rights Reserved

* Recursion $$\beta_n(S_i) = \sum_{j=1}^{S} a_{ij}.b_j(X_{n+1}).\beta_{n+1}(S_j)$$

* Termination $$p(X/\lambda) = \beta_o(S_i) = \sum_{j=1}^{S} a_{ij}.b_j(X_1).\beta_1(S_j) = \alpha_N(S_E)$$

The trills diagram for backward variable 'β' Calculation is shown in figure 2.

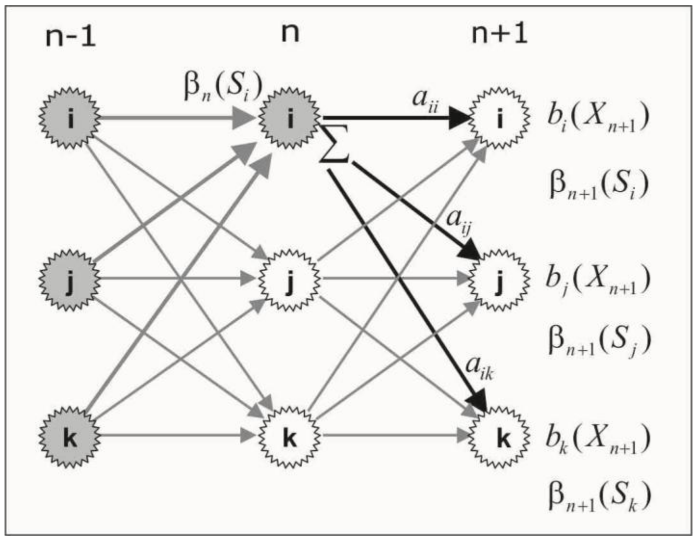

Figure 2 : Showing Backward Recursion

2.3 Viterbi Algorithm

We shall consider the most likely state sequence instead of summing over all possible state sequence. This can be estimated by just changing the summation to a maximization in the recursion step which is given below

$$V_n(S_j) = \max V_{n-1}(S_i).a_{ij}.b_j(X_n)$$

The equation clearly indicates that the likelihood of the most probable path can be achieved by changing the summation into a maximum in the recursion. The path or state sequence backtracking or Viterbi back trace is done via keeping a sequence of back pointers. The terills diagram for Viterbi recursion showing likelihood of the most probable path and back pointers to the previous states on the most probable path is shown in figures 3 and figure 4 respectively [3].

@IJAERD-2015, All rights Reserved

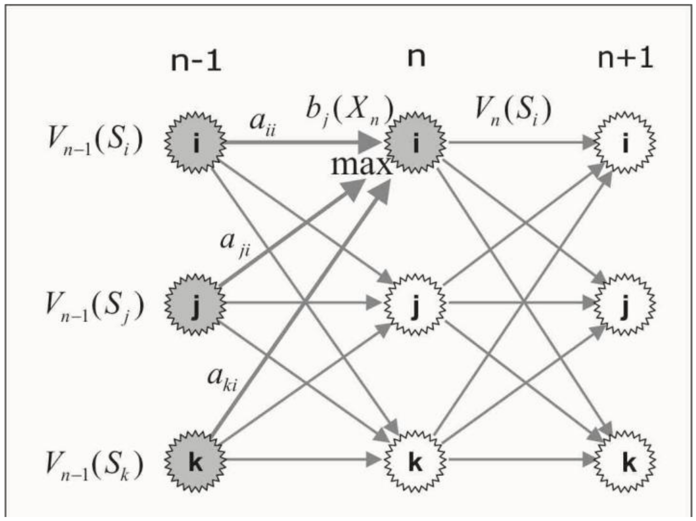

Figure 3 : Showing likelihood of the most probable path (Sequence)

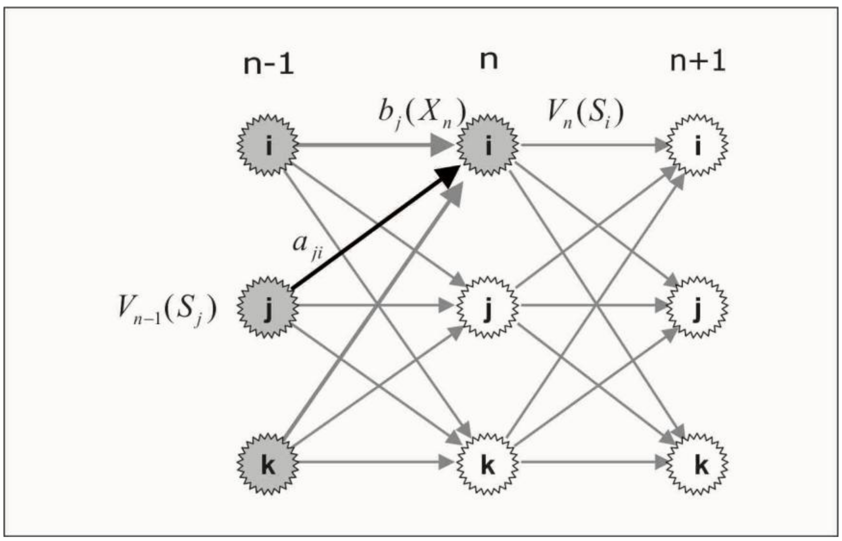

Figure 4 : Showing back pointers to the previous states on the most probable path.

Algorithm

- Initialization

$$V_o(S_i) = 1$$
$$V_o(S_j) = o \quad if \ S_j \neq S_i$$
$$b_{no}(S_j) = o$$

@IJAERD-2015, All rights Reserved

- Recursion

$$V_n(S_j) = \max_{i=1}^{S} . V_{n-1}(S_i).\ a_{ij}.b_j(X_n)$$

$$b._{nn}(S_j) = \arg.\max_{i=1}^{S} . V_{n-1}(S_i).a_{ij}.b_j(X_n)$$

- Termination

$$P^* = V_N(S_E) = \max_{i=1}^{S} V_N(S_i).a_{iE}$$

$$S_N^* = bn_N(q_E) = \arg.\max_{i=1}^{S} V_N(S_i).a_{iE}$$

2.4 **The Forward-Backward Algorithm**

The main objective here is to efficiently find out the parameters of an hidden Markov model (HMM) λ from an observation sequence $X_1, X_2.... X_n$. If we assume a single Gaussian output probability distribution which is given as below[39]

$$b_j(x) = p(X \mid S_j) = g(O, \mu_j, \Sigma_j)$$

The parameter $\lambda = \{C_{jk}, \mu_{jk}, \Sigma_{jk}\}$ defines an hidden Markon model. The transition probability a_{ij} follows a stochastic constraint that is $\sum_j a_{ij} = 1$ and the Gaussian parameters for the state S_j are given as μ_j the mean vector and $\sum_j$ the covariance matrix.

If we known the state-time alignment diagram then each observation vector could be assigned to a specific state. A state-time alignment diagram or terrils diagram can be easily obtained using the Viterbi algorithm thru finding the most probable path.

The maximum likelihood estimate of a_{ij} can be done as follows.

If $C(S_i \to S_j)$ is the count of transitions from state (S_i) to state (S_j).

$$a_{ij} = \frac{C(S_i \to S_j)}{\sum_k C(S_i \to S_k)}$$

Likewise if Z_j is the set of observed acoustic feature vectors assigned to state j we can use the standard maximum likelihood estimates for the mean and co variance matrix.

$$\mu_j = \frac{\sum_{x \in z_j} X}{|z_j|}$$

$$\sum_j = \frac{\sum_{x \in z_j} (x-\mu_j).(x-\mu_j)'}{|Z_j|}$$

EM Algorithm

In Viterbi training which is an approximation we would like to consider all the possible paths. In this case we estimate the probability rathan than having hard time alignment. The state occupation probability is denoted by $\gamma_n(S_{ij})$ which states the probability of occupying State S_j at time n given the sequence of observations.

On comparing with component occupation probability in a GMM we can use this for an iterative algorithm for HMM training the EM algorithm. It has two steps E&M.

E-Step: It estimates the state occupation probabilities E-Step is called Expectation Step.

M-Step: It is called Maximization Step. Here the HMM parameters based on the estimated state occupation probabilities are re-estimated.

2.5 **Calculation of $\gamma_n(S_j)$ (Gamma)**

$\gamma_n(S_j)$ is the probability of occupying state S_j at time 'n' given the sequence of observation. It is expressed in terms of forward $'\alpha'$ and backward $'\beta'$ probabilities.

$$\gamma_n(S_j) = P(S(n)) = S_j \mid X, \lambda = \frac{1}{\alpha_N(S_E)} \alpha_n(j)\beta_n(j)$$

recalling that $p(X \mid \lambda) = \alpha_N(S_E)$

Since $\alpha_n(S_j)\beta_n(S_j) = p(X_1, X_2, ...X_n, S(n) = S_j \mid \lambda).p(X_{n+1}, X_{n+2}....X_N, S(n) = S_j, \lambda)$

@IJAERD-2015, All rights Reserved

$$= p(X_1, X_2 ... X_n, X_{n+1}, X_{n+2} ... X_N, S(n) = S_j / \lambda)$$

$$= p(X, S(n) = S_j | \lambda)$$

$$P\, S(n) = S_j | X, \lambda) = \frac{p(X, S(n) = S_j | \lambda)}{p(X | \lambda)}$$

2.6 **Re-estimation of Gaussian Parameters.**

The sum of state occupation probabilities through time n for a state $'S_j'$ may be called as a soft count which may be used to re-estimate the HMM parameters

$$\overline{\mu}_j = \frac{\sum_{n=1}^{N} \gamma_n(S_j).X_n}{\sum_{n=1}^{N} \gamma_n(S_j)}$$

$$\overline{\sum}_j = \frac{\sum_{n=1}^{N} \gamma_n(S_j).(X_n - \overline{\mu}_j).(X - \overline{\mu}_j)'}{\sum_{n=1}^{N} \gamma_n(S_j)}$$

2.7 **Re-estimation of Transition probabilities and calculation of 'ξ'(Zie)**

We can estimated $\xi_n(S_i, S_j)$, the probability of being in state S_i at time n and state S_j at time $n+1$, given the observations.

$$\xi_n(S_i, S_j) = P\left(S(n) = S_i,\ S(n+1) = S_j | X, \lambda\right)$$

$$= \frac{P\left(S(n) = S_i,\ S(n+1) = S_j, X | \lambda)\right)}{P(X | \lambda)}$$

$$= \frac{\alpha_n(S_i).a_{ij}.b_j(X_{n+1})\beta_{n+1}(S_j)}{\alpha_N(S_E)}$$

The variable $'\xi'$ (zie) is used to re-estimate the transition probabilities as follows:

$$\overline{a}_{ij} = \frac{\sum_{n=1}^{N} \xi_n(s_i, s_j)}{\sum_{k=1}^{S}\sum_{n=1}^{N} \xi_n(s_i, s_k)}$$

E-Step for all time-state pairs. Here the forward probabilities $\alpha_n(S_j)$ and backward probabilities $\beta_n(j)$ are recursively computed. M Step based on the estimated state occupation probabilities re-estimate the HMM parameters mean vectors μ_j & the covariance matrix $\sum_j$ & transition probabilities a_{ij}.

The M component Gaussian mixture model is an appropriate probability function & is given by

$$b_j(x) = p(X \mid S_j) = \sum_{m=1}^{M} C_{jm}.g(o, \mu_{jm}, \Sigma_{jm})$$

We estimate the component-state occupation probabilities. $\gamma_n(S_j, m)$ is the probability of occupying mixture components m of state S_j at time n. We may estimate the mean and co-varience of mixture components m of state S_j at follows.

Mean:

$$\overline{\mu}_{jm} = \frac{\sum_{n=1}^{N} \gamma_n(s_j, m) X_n}{\sum_{n=1}^{N} \gamma_n(S_j, m)}$$

Co-varience: :

$$\overline{\Sigma}_{jm} = \frac{\sum_{n=1}^{N} \gamma_n(S_{j,m}).(X_n - \overline{\mu}_{jm})(X_n - \overline{\mu}_{jm})}{\sum_{n=1}^{N} \gamma_n(S_j, m)}$$

The mixture coefficients or mixture weights are re-estimated as follows:

$$\overline{C}_{jm} = \frac{\sum_{n=1}^{N} \gamma_n(s_j, m)}{\sum_{\ell=1}^{M} \sum_{n=1}^{N} \gamma_n(S_j, \ell)}$$

@IJAERD-2015, All rights Reserved

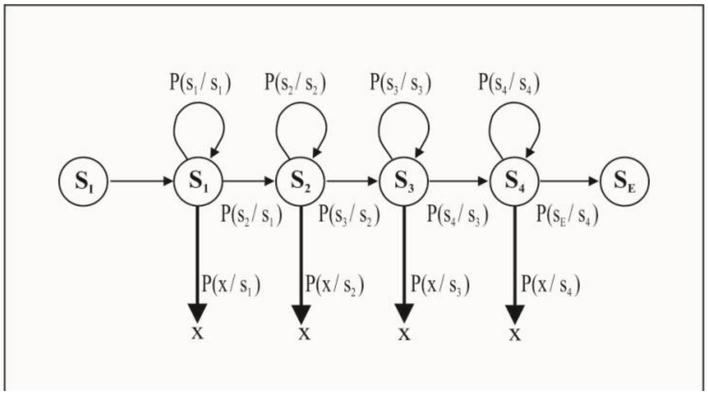

Figure 5 : Showing continuous density HMM Model.

Continuous density HMM an acoustic model is shown in the figure 5. Here the transition probabilities are $a_{kj} = P(S_j \mid S_k)$ and output probability density function $b_j(X) b_j(X) = p(X \mid S_j)$ and an HMM is represented by the parameter $\lambda = [A, B, \pi]$. Where $A = a_{ij}$ and $B = b_j(X) \& \pi_i$ is initial state distribution.

3 Design of proposed CDHMM Hindi Speech recognizer.

Design Specifications of CDHMM:

i) Observation vectors = continuous
ii) Probability density function (pdf)= Gaussian
iii) Number of Mixture Components (M)=16.
iv) No. of states (Q)=6.
v) No. of observation Vectors in a sequence (T) = m2 .
vi) No. of coefficients in a vector (o) = m1.
vii) Type of Co-variance matrix used = diagonal.
viii) Type of data used = continuous observations vectors.
ix) Total No. of words for which system is trained w=10.
x) Number of utterances of each word = 5.

A CDHMM is defined by the parameters $\lambda cdhmm$

$$\lambda cdhmm = [\pi_i, A, B]$$

π_i is the initial state distribution also represented by prior. A is the transition matrix $A = a_{ij}$. B is the probability distribution function, $b = b_j(x)$. It is calculated by using the formulae.

$$b_j(x) = \sum_{k=1}^{M} C_{jk} . g(x, \mu_{jk}, \Sigma_{jk})$$

Where C_{jk} is the mixture weight of the Kth component in state j. μ_{jk} is the mean vector of the Kth component in state j, $\sum_{jk}$ is the covering mixture weight satisfy the stochastic constraint $\sum_{k=1}^{M} C_{jk} = 1, \quad 1 \le j \le N$

Where g represents the Gaussian function.

In order to design an isolated word speech recognizer using CDHMM concept for a vocabulary size of words 'W'=10. We design a Q-State HMM. We represent the speech signal of a given Hindi word as a sequence of continuous observation vectors. A word is represented by a sequence $X_1, X_2.., X_n$ Here the probability of each observation vector is calculated by the Gaussian pdf. The data sequence $X = X_1, X_2, X....X_N$ which is a D dimensional Mel Frequency Cepstral Coefficients(MFCC). The Signal modeling techniques gives an idea how speech samples can be used to generate observation sequence. The Mel Frequency Cepstral Coefficients represents the best approximation of the human ear. The human ear is more sensitive to higher frequencies [4]. A review on Speech recognition techniques concludes that the Mel Frequency Cepstral Coefficients are widely used for developing the front end of a Speech Recognizer [5][6][7]. The Speech vectors are uniformly segmented into Q=6 States as illustrated in figure 6.

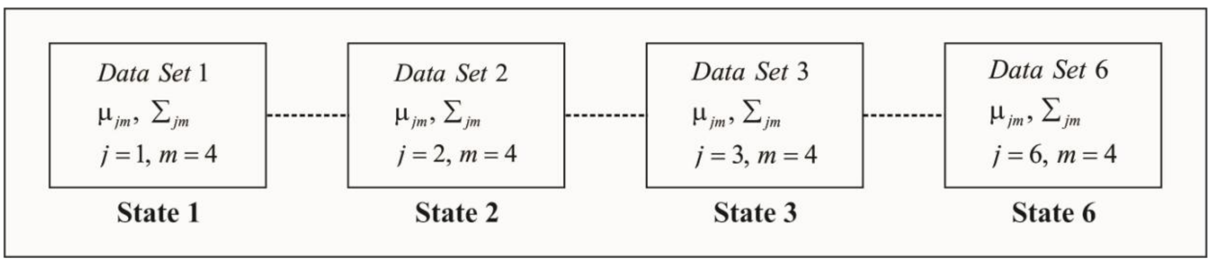

Figure 6 : Showing Uniform Segmentation of Data Sequence

i) No. of mixture 'm' per State = 4

ii) Mean Vector $\mu_{j1}, \mu_{j2}, \mu_{j3}, \mu_{j4}$

iii) Covariance Matrix $\sum_{j1}, \sum_{j2}, \sum_{j3}, \sum_{j4},$

Design Steps

i) Initialize each mean to a random data point

ii) Random guessing of initial state distribution and transition matrix 'A'

$\pi_i, \rightarrow$ Initial State distribution (Prior)

$A = a_{ij} \rightarrow$ Transition Matrix

iii) Calculation of Gaussian parameters $\mu(MU)$ and covarience Σ (Sigma) through initialization of each mean to random data point

@IJAERD-2015, All rights Reserved

Thus for each vocabulary word, we have a training consisting of a number of repetition of sequences (D dimensional continuous observation vectors) of the word by male or female talkers. The first thing is to make individual models of each word. This task can be completed by using the solution to problem 3 to optimally estimate model parameters for each word model. The understand the physical interpretation of the number of states of a word model we make use of the solution to problem 2 to segment each of the word training sequence into states and then studying the Gaussian parameters (Mean vector μ_{jm}) covariance $\sum_{jm}$ (Sigma) and the number of mixture components 'm' per state) that leads to the observations occurring in each state. Our aim is her to do the model refinements such that more states and different data of continuous observation vectors, number of mixtures components per state etc. to improve the capability of the spoken word sequences.

Finally when we have developed the set of CDHMM's for W words & optimized them then the task of recognition of unknown word is performed using the solution to problem to score each word model based upon the given test observation sequence and select the word who has the highest model score or likelihood illustrated in figure 7.

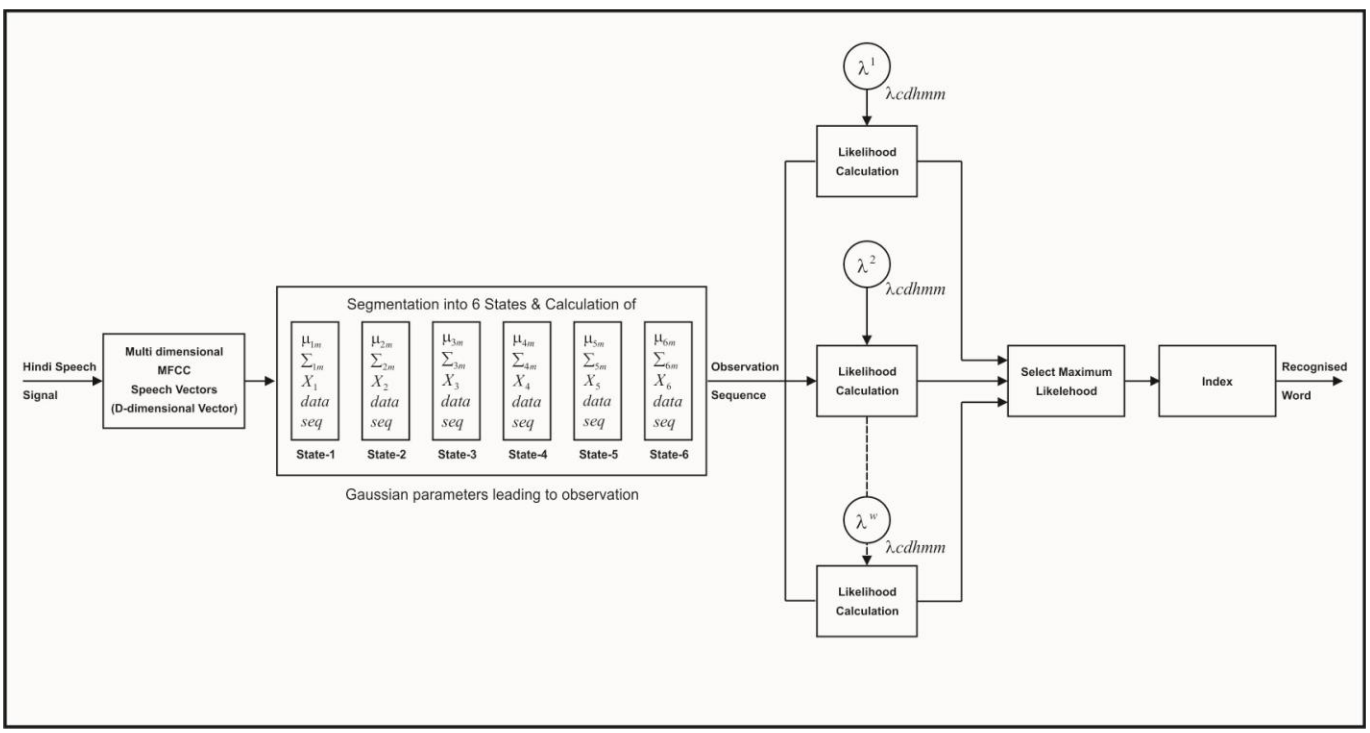

Figure 7: Showing CDHMM Isolated Hindi Speech Recognizer.

4(A) **Results of the experimentation performed on CDHMM Hindi Speech recognizer** for Male Speaker (Data Set1). with M=16 and Q=6 is shown in the Table 1

Table 1: Confusion matrix for various words spoken by Male Speaker(Data Set1).

	Q, States=6 M, Mixture=16	**Likelihood Calculation using CDHMM** Test Utterances									
	Training Utterances	*Hanuman*	*Kalam*	*Kapi*	*Kitab*	*Pen*	*Pustak*	*Raam*	*Ravan*	*Sita*	*Sugriv*
1	*Hanuman*	**-2.7494**	-5.611	-5.0417	-4.3934	-4.3904	-4.3773	-3.7954	-4.5096	-4.6061	-6.7809
2	*Kalam*	-3.9598	**-4.458**	-5.0477	-4.3567	-4.0794	-3.4232	-4.004	-4.199	-4.6498	-7.2015
3	*Kapi*	-3.7282	-5.4826	**-2.8497**	-2.8278	-3.687	-3.2371	-3.9974	-3.4654	-3.0811	-5.2268
4	*Kitab*	-5.2078	-6.6197	-4.3947	**-2.7802**	-5.176	-3.754	-5.2292	-4.7143	-3.7836	-6.1881
5	*Pen*	-5.8057	-7.1286	-5.8101	-6.7023	**-3.0412**	-4.4294	-5.7026	-5.8136	-4.6703	-6.3549
6	*Pustak*	-5.287	-6.0959	-5.2813	-5.0168	-4.6395	**-2.5641**	-5.0255	-4.8351	-3.9623	-7.8272
7	*Raam*	-4.2259	-6.3378	-6.1344	-4.7452	-5.1907	-5.0597	**-3.4213**	-4.8666	-5.3237	-8.5434
8	*Ravan*	-4.9932	-7.9238	-5.4255	-4.5099	-5.5488	-4.9309	-4.7341	**-3.0251**	-4.9129	-9.4172
9	*Sita*	-4.86	-7.458	-3.8955	-3.5213	-4.6613	-4.0002	-4.207	-3.9619	**-2.5542**	-3.8524
10	*Sugriv*	-6.2887	-8.7874	-6.2467	-6.5331	-5.2362	-4.9114	-7.1247	-6.3289	-4.4387	**-3.541**

(B) **Results of the experimentation performed on CDHMM Hindi Speech recognizer** for male Speaker.(Data Set2) with M=16,Q=6 is shown in Table 2.

Table 2 : Confusion matrix for various words spoken by male Speaker(Data Set2)

	Q, States=6 M, Mixture=16	**Likelihood Calculation using CDHMM** Test utterances									
	Training utterances	*Hanuman*	*Kalam*	*Kapi*	*Kitab*	*Pen*	*Pustak*	*Raam*	*Ravan*	*Sita*	*Sugriv*
1	*Hanuman*	**-0.5184**	-1.8749	-1.0063	-1.4911	-1.3287	-1.6552	-1.6233	-0.9683	-2.4191	-2.1865
2	*Kalam*	-1.0342	**-0.3121**	-0.7338	-0.5515	-0.7208	-0.9031	-0.7823	-0.8936	-0.7508	-0.916
3	*Kapi*	-0.8914	-1.1343	**-0.3752**	-0.9918	-0.7177	-0.9972	-1.279	-0.909	-1.5094	-1.4957
4	*Kitab*	-0.924	-0.6494	-0.7287	**-0.3467**	-0.674	-0.9001	-0.6967	-0.7727	-1.0119	-1.1744
5	*Pen*	-1.5243	-	-0.8202	-	**-**	-	-0.826	-	-	-1.1988

@IJAERD-2015, All rights Reserved

			0.8804		0.6575	**0.3399**	1.0148		0.7124	1.1558	
6	*Pustak*	-1.353	-1.1792	-0.6684	-0.6562	-0.9382	**-0.5843**	-1.0788	-1.0661	-1.6513	-1.7959
7	*Raam*	-0.6426	-0.9956	-1.0812	-0.8203	-1.071	-1.5431	**-0.2898**	-0.9584	-1.5636	-1.6113
8	*Ravan*	-1.3753	-1.0747	-0.9608	-1.0096	-1.1267	-1.6133	-1.2818	**-0.6346**	-1.7369	-2.0192
9	*Sita*	-0.8261	-0.6643	-0.809	-0.7576	-0.7223	-1.2219	-0.9362	-0.9732	**-0.4307**	-0.68
10	*Sugriv*	-1.0183	-0.6927	-0.804	-0.7176	-0.869	-1.0477	-0.9738	-0.8099	-0.6394	**-0.5882**

.

5 Results & conclusions.

The proposed CDHMM Hindi speech recognizer makes use of Gaussian pdf and its parameters mean vector μ_{jm}. Here j varies between $1 \leq j \leq S$ and the number of mixture components M=4,8,16 per state. The data set on which the experimentation was performed was MS (Male Speaker,(Data Set1) and MS (Male Speaker(Data Set2) in the age group of 30 yrs. to 50 yrs. The covariance Matrix denoted by Sigma $\sum_{jm} (j = 1 \, to 6 \,\& \, m = 4,8,16)$ was diagonal matrix. The results have shown that the Co variance matrix can be used of 'full' & diagonal type & the likelihood in both the cases as shown in the confusion matrix Table1 and Table 2 are used to recognize the unknown utterance/word. The best results were obtained when M=16 was used as optimal number of mixture components.

REFERENCES

[1]. Lawrence Rabiner and B. H. Juang, "*Fundamentals of Speech Recognition''*, Prentice Hall, Englewood clliffs,NJ,1993.
[2]. Douglas Reynolds." *Gaussian Mixture Models*" MIT Lincoln Laboratory, 244 Wood St. Lexington, MA 02140,USA,dar@ll,mit.edu,
[3]. Steve Renals and Peter Bell."*Hidden Markov Models and Gaussian Mixture Models*" Automatic Speech Recognition-ASR Lectures 4 & 5 28/31 January 2013.
[4]. Joseph W Picone,"*Signal Modeling Technique in Speech Recognition*", Proceedings of the IEEE,vol 81,No.9,September 1993.
[5]. S .K. Gaikwad,"*A Review on Speech Recognition Techniques*", I. J.of Computer Applications",Vol.10 Nov, 2010.
[6]. 6. Satish kumar, Prof Jai Prakash "*Developement of Front End and Statistical Model for Hindi Speech Recognizer: A Practical Approach*" Journal of Indian Research, Vol. 2, Number 3, July- September , 2014.
[7]. Satish kumar, Prof Jai Prakash "*Design & Development of Discrete HMM (DHMM) Isolated Hindi Speech Recognizer*" Journal of Indian Research, Vol. 2, No. 4,October-December , 2014.

@IJAERD-2015, All rights Reserved