



Analytical Evaluation of Hidden Markov Model Techniques and Dynamic Time Wrap for Isolated Word Speech Recognition

Er. Shweta Doda

Department of Computer Science and Engineering
JCD Vidyapeeth, Sirsa, India

Er. Rajni Mehta

Assistance Professor in Dept. Computer Science
JCD Vidyapeeth, Sirsa, India

Abstract- *Speech Recognition is the process of converting a speech signal to a sequence of words, by means of an algorithm implemented as a computer program. Speech is the most natural form of human communication. Speech recognition technology has made it possible for computer to follow human voice commands and understand human languages. The main goal of speech recognition area is to develop techniques and systems for speech input to machine. Dynamic Time Warp and Hidden Markov techniques are used for isolated word recognition. In DTW approach, time warping technique is combined with linear predictive coding analysis. In HMM approach, well known techniques of vector quantization and hidden markov modeling are combined with a linear predictive coding analysis. The objective of the study is to implement small vocabulary isolated speech recognition systems and compare their performances*

Keywords- *Feature Extraction, Performance Evaluation, DTW, HMM*

I. INTRODUCTION

Speech is the most natural and efficient way to exchange information for human beings. To make a real “intelligent computer”, it is important that the machine can hear, understand, and act upon spoken information, and also speak to complete the information exchange. It provides natural, fast, hands free, eyes free and location free means of interaction between the user and computer. Voiced command applications are expected to cover many of the aspects of our future daily life. Car computers, telephones and general appliances may be the more likely candidates for this revolution that may reduce drastically the use of keyboards

Two Models of Speech Recognition System

- ❖ Dynamic Time Wrap (DTW)
- ❖ Hidden Markov Model (HMM)

Dynamic Time Wrap (DTW):- DTW is a method that allows a computer to find an optimal match between two given sequences .It is template based approach. DTW based on two concepts (Symmetrical DTW, Asymmetrical DTW).

Comparison between Asymmetrical and Symmetrical DTW:

1. Complexity:

Complexity is the combination of two things.

Space complexity: Space complexity can be determining the space occupied by the algorithm in the computer memory.

Time Complexity: Whereas the time complexity can be determined by the time required to execute the algorithm.

The overall complexity of the symmetric algorithm is:

$$O((n+|E|) \log n)$$

Where n is the number of nodes.

And the overall complexity of asymmetric algorithm is:

$$O(|E| \log |E|)$$

Where E is the Set of Edges in Graph G.

The time complexity of Asymmetric algorithm is better than symmetric algorithm of Dynamic time wrap. Because the time complexity is depends on the number of comparisons to find the shortest path between the source and destination.

2. Structure:

In asymmetric algorithm different nodes are consider as a forest tree and connect them with the help of this asymmetric algorithm. The structure of asymmetric algorithm is so complex.

Hidden Markov Model (HMM):- It is a mathematical approach to recognize speech. It is a doubly embedded stochastic process with an underlying stochastic process that is not directly observable (it is hidden) but can be observed only through another set of stochastic processes that produce the sequence of observations.

Solutions of HMM Problems

Evaluation: - Evaluation is to find the probability of generation of a given observation sequence by a given model. This probability can be obtained by computing the joint probability of observed sequence and each possible state sequence I of

length T and then summing over all the state sequences. Computationally this method is very costly. However, there is an efficient way of computing this probability using forward and backward probabilities.

Decoding Problem: - Decoding is to find the single best state sequence for the given observation sequence. To uncover the most likely state sequence, remember that the best path to each cell, which is the concatenation of the path to its predecessor state and the best step to the cell.

The Learning Problem: - The learning problem is to adjust the model parameters to maximize the probability of the observation sequence given the model. This is the most difficult task of the Hidden Markov Modeling, as there is no known way to solve for a maximum likelihood model analytically. In this method start with some initial estimates of the model parameters and modify the model parameters to maximize the training observation sequence probability in an iterative manner till the model parameters reach a critical value.

II. RELATED WORK

D. Raj Reddy [April 1976]. In this paper stated that the focus has been to review research progress, to indicate the areas of difficulty, why they are difficult, and how they are being solved. The past few years have seen several conceptual and scientific advances in the field. For the first time use the available extensive analysis of connected speech. Know connected speech recognition is not impossible. The role and use of knowledge are better understood. Almost all systems use knowledge to generate hypotheses and/or verify them. Error and ambiguity can be handled within the framework of search. Stochastic representations and dynamic programming provide a simple and effective solution to the matching problem.

Bishnu S. Atal, and Lawrence R. Rabiner [June 1976] This paper presented that a fairly general framework based on a pattern recognition approach to VUS classification has been described in which a set of measurements are made on the interval being classified and a minimum non-Euclidean distance measure is used to select the appropriate class. Almost any set of measurements can be used so long as there is some physical basis for assuming that the measurements are capable of reliably distinguishing between these three classes. Although a non-Euclidean distance measure was used, other distance measures may be equally appropriate. Finally, a smoothing algorithm was discussed which was appropriate for a digit recognition algorithm in which errors in the analysis were corrected, and unusually short intervals were eliminated.

Lawrence R. Rabiner, S.E. Levinson, and M. M. Sondhi [April 1983] to search out that to present the theory of hidden Markov models from the simplest concepts (discrete Markov chains) to the most sophisticated models (variable duration, continuous density models). The purpose to focus on physical explanation of the basic mathematics; hence to avoid long, drawn out proofs and/or derivations of the key results, and concentrated primarily on trying to interpret the meaning of the math, and how it could be implemented to illustrate some applications of the theory of HMMs to simple problems in speech recognition, and pointed out how the techniques could be (and have been) applied to more advanced speech recognition.

M. Karnjanadecha and Stephen A. Zahorian [1999] highlighted that word-based HMM recognizers have been claimed to give poorer recognition performance compared to phoneme-based recognizers. Experimental results reported in this paper have shown that if features are computed properly, whole word recognizers *can*, in fact, surpass the best reported results for phoneme based recognizers. Whole word HMM based recognizers are easier to implement, at least for isolated word recognition. Results show that the proposed signal modeling techniques are straightforward, efficient and robust.

III. PROPOSED WORK

A wide variety of approaches have been proposed to recognize isolated words, based on standard statistical-pattern-recognition techniques. In this study about Hidden Markov Model Techniques and Dynamic Time Wrap for Isolated Word Speech Recognition are studied and analysis.

SYMMETRICAL DTW ALGORITHM

1. Algorithm [E, cost, n, t]
2. E is the set of edges in G. G has n vertices, cost[u,v] is the cost of edge(u,v). t is the set of edges in the minimum cost spanning tree. The final cost is required.
3. {
4. Construct a heap out of the edge cost using heapify.
5. for i=1 to n do parent [i]=-1.
6. I=0, mincost=0.0.
7. while ((i<n-1) and (heap not empty)) do
8. {
9. delete a minimum cost edge (u,v) from the heap.
10. and reheapify using adjust.
11. j=find (u); k=find(v).
12. if (j!= k) then
13. {
14. i=i+1.
15. i[I,1]=u. t[I,2]= v.

16. mincost = mincost + cost[u,v].
17. union(j,k).
18. }
19. }
20. if (i!= n-1) then write (“no spanning tree”).
21. else return mincost.
22. }

HIDDEN MARKOV MODEL ALGORITHM

1. Algorithm [v, cost, dist, n]
2. dist[j], $1 \leq j \leq n$ is set to the length of the shortest path from vertex v to vertex j in a digraph. G with n vertices, dist[v] is set to zero. G is represented by its cost adjacency matrix, cost[1:n, 1:n].
3. {
4. for (i=1 to n) do
5. initialize S.
6. S[i]= false, dist [i]= cost [v,i].
7. }
8. S[v] = true, dist [v]= 0.0.
9. for num=2 to n-1 do
10. {
11. Detrimine n-1 pats from v
12. choose u from among those vertices not.
13. in S such that dist[u] is minimum.
14. s[u] = true.
15. for (each w adjacent to u with S[w]= false) do
16. update distances.
17. if (dist[w]> dist[u] + cost [u,w]) then
18. dist [w]= dist [u] + cost [u,w].
19. }
20. }

Difference between HMM and DTW

Sr. No.	Parameters	Hidden Markov Model	Dynamic Time Wrap
1.	Complexity	The Hidden Markov model Algorithm is more complex in comparison to dynamic time wrap. The complexity of Hidden Markov model is: $O(E \log E)$ In HMM one node is traverse exactly once. No repetition is allowed.	The complexity of dynamic time wrap is $O((n+ E) \log n)$ In DTW there may be no of paths to traverse the data. There may be chance to repetition of nodes to traverse the data.
2.	Structure	The structure of the Hidden Markov model is more complex. The HMM is used for unicasting. So the no of comparisons in HMM is less. But in unicasting the comparisons are so complex and to find the shortest path at minimum cost is very difficult.	The structure of Dynamic time wrap is very easy to understand. In DTW the broad casting we used to traverse the data. In DTW the no of comparisons are more but there may be no of paths for traversing the data.
3.	Security	The HMM is more secure in comparison to DTW, because HMM use unicasting to transfer the data. So only the source and destination know about the data so the data transmission is secure.	In DTW the security of data is not possible. Because it use broad casting. Any node can be access the data, whether it is required by them or not. So it is not secure.
4.	Reliability	The HMM is not reliable. Because there are only one source and one destination. If due to any reason the system get fail then there will be no chance to recover the data	The DTW is more reliable in comparison to HMM. Because it use broad casting and there may be no of destinations, so if any one node become fail, the other system get the data and the data will be secure. And in case of corruption of any system the data can be recover.
5.	Backup	In HMM there should be no back up of data	In DTW the back up of data should be kept on other system.

6.	Data Transmission	It uses the unicasting technology for data transmission.	It uses the broad casting technology for data transmission.
7.	Economical	The HMM is economical in comparison to DTW. Because in HMM the components and S/w will be installed on only single system for data transmission, so the no of installation are not required. It is economical.	In DTW numbers of nodes are used for transmission. So the s/w and the components are installed on all the system, Which is very costly.

IV. RESULT AND DISCUSSION

Both Dynamic Time Warp approach and Hidden Markov approach for isolated word recognition of small vocabulary are implemented. In DTW approach, time warping technique is combined with linear predictive coding analysis. In HMM approach, well known techniques of vector quantization and hidden Markov modeling are combined with a linear predictive coding analysis. This is done in the framework of a standard statistical pattern recognition model. To provide some basis of comparison for the performance of the DTW recognizer with HMM recognizer, the same speech data was tested on both of the recognizers. The recognition accuracies of the DTW recognizer and HMM recognizers are given in the table:

WORD	DTW ACCURACY (%)	HMM ACCURACY (%)
Hindi	96.0	92.0
English	96.0	76.0
Physics	100	96.0
Chemistry	100	100
Computer	80	96.0
Average	94.4	92.0

Fig. 1 Performance of DTW Isolated Word Recognition System

Direct comparisons of the results of HMM based recognizer with the results of the DTW based recognizer with single template for each vocabulary word shows that the HMM based recognizer performs only a little worse than the DTW based recognizer. The results show that when DTW recognizer has incorrectly identified the word, most of the time the HMM recognizer has correctly identified (see the results of utterances English and Computer). The systems were trained with one utterance per word from a single speaker. These results are for a typical speaker, hence there is so much variation in recognition accuracies of different words. Still an inference can be drawn about the overall behavior of these approaches.

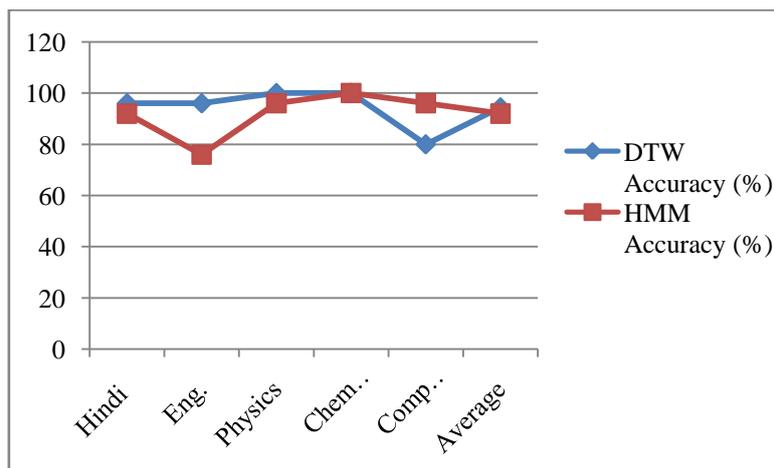


Fig. 2 Graphical Representation of Accuracy of DTW vs. HMM

V. CONCLUSION

In this research we try to find the better technique for speech reorganization. The fact that performance of Hidden Markov Model recognizer is somewhat poorer than the Dynamic Time Wrap based recognizer appears to be primarily because of the insufficiency of the Hidden Markov Model training data. Having seen one consequence of this inadequacy in that to use the constraints on value whose value fell below the threshold value. With the increase in the size of the codebook, the accuracy of the HMM based recognizer improves. The performance of the Hidden Markov Model recognizer also depends on the number of states of the model. It is necessary that number of states should be such

that they can model the word. The time and space complexity of the Hidden Markov Model approach is less than the Dynamic Time Wrap approach because during Hidden Markov Model testing to compute the probability of each model to produce that observed sequence. In Dynamic Time Wrap testing, the distance of the input pattern from every reference pattern is computed, which is computationally more expensive.

The results have shown that the accuracy of Dynamic time Wrap technique is better in comparison to the Hidden Markov Model for speech recognition. In some cases the accuracy of Hidden Markov Model are better in comparison to the Dynamic time Wrap technique for speech recognition. The errors made by the two recognizers are largely disjoint. Hence there exist the potential of using some fairly standard techniques to combine the two recognizers to give better accuracy.

REFERENCES

- [1] D. Raj Reddy, *Speech Recognition by Machine: A Review*, Proceedings of the IEEE, Vol. 64, No. 4, April 1976, pp 501-531.
- [2] Bishnu S. Atal, and Lawrence R. Rabiner, *A Pattern Recognition Approach to Voiced-Unvoiced Classification with Application to Speech Recognition*, IEEE Transaction on Acoustics, Speech, and Signal Processing, Vol. ASSP 24, No. 3, June 1976, pp 201-212.
- [3] Lawrence R. Rabiner, S.E. Levinson, and M. M. Sondhi, *On the Application of Vector Quantization and Hidden Markov Models to Speaker-Independent, Isolated Word Recognition*, AT&T Technical Journal, Vol. 62, No. 4, April 1983, pp. 1075-1105
- [4] M. Karnjanadecha and Stephen A. Zahorian, *Signal Modeling for Isolated Word Recognition*, IEEE International Conference on Acoustics, Speech, and Signal Processing, 1999, Vol. 1, pp. 293-296