

A NEW HMM TOPOLOGY FOR SHAPE RECOGNITION

Nafiz ARICA, Fatos. T. YARMAN-VURAL

Department of Computer Engineering, METU-TURKEY

{vural, arica}@ceng.metu.edu.tr

ABSTRACT

This study deals with the shape recognition problem using the Hidden Markov Model (HMM). In many pattern recognition applications, selection of the size and topology of the HMM is mostly done by heuristics or using trial and error methods. It is well known that as the number of states and the non-zero state transition increases, the complexity of the HMM training and recognition algorithms increases exponentially. On the other hand, many studies indicate that increasing the size and non-zero state transition does not always yield better recognition rate. Therefore, designing the HMM topology and estimating the number of states for a specific problem is still an unsolved problem and requires initial investigation on the test data.

This study addresses a specific class of recognition problems based on the boundary of shapes. The paper investigates the affect of the HMM topology on the recognition rate. A new topology, called circular HMM, is proposed and tested on the handwritten character recognition problem. The proposed topology is both ergodic and temporal. It eliminates the starting and ending states with the circular state transitions. The experiments indicate excellent performance compared to the classical temporal and ergodic HMM models.

1. INTRODUCTION

Hidden Markov Model (HMM) is a widely used powerful tool for many pattern recognition and image analysis problems. There is a tremendous amount of variations of HMM applications, which input various feature sets into various HMM sizes and topologies [1]. Efficient iterative algorithms are available for estimating the model parameters of observation probability and state transition matrices. However, all of the approaches presume a model size and topology [2]-[6]. Unfortunately, there are no effective methods for estimating the optimal number of states and/or the nonzero state transitions for a specific feature set. Many pattern recognition applications indicate

that the number of states should be somehow proportional to the length of the observation sequence and to the code book size, for the HMM. However, higher number of non-zero transitions does not always, provide better recognition rates, but result in extra computational cost.

Ergodic topologies enable the revisits of each state with probability one in finite intervals, by allowing non-zero state transition paths between any two states. However, they do not impose a temporal order. Therefore, when the observation sequence has a temporal order, ergodic models do not fully utilize the temporal information of the data.

On the other hand, the temporal topologies do not allow the revisits to the previous states by constraining the state transition probabilities, $a_{ij}=0$ for $j \geq i+k$, where k is a small integer compared to the total number of the states. This constraint yields a sparse state transition matrix, where the nonzero entries lie only in the few upper diagonals. For this reason, in most of the pattern recognition applications, it is accustomed to use, so called, left-right model. This model eliminates estimating the initial state probabilities because it has a single starting and terminating state.

Experimental results of many studies indicate that left-right topologies are more appropriate to reach the maximum recognition rates in many applications, such as speech and optical character recognition. However, when the feature set consists of the quantized values of a closed boundary, it is very difficult to identify consistent starting and ending points on a boundary of the object to represent the observation sequence. Therefore, in the recognition problem based on the object boundaries, the available HMM models yield very low recognition rates if the feature sets do not have a geometrically meaningful starting and terminating points.

Although, there is not an exact criterion, it is generally accepted that the number of states is taken to be proportional to the length of the observation sequences and/or the number of distinct observation [6]. This fact brings another complicated problem for size invariance:

In order to make a consistent platform for comparison of the HMM probabilities, the sizes of the boundaries are to be normalized for generating fixed length observation sequences for the patterns.

In this study, a new topology, called circular HMM, is proposed. This topology is a simple modification of left-to-right HMM model, where the initial and terminal states are connected through the state transition probabilities. This connection eliminates the need to define a starting point of a closed boundary, in the recognition problem. The proposed HMM topology is both temporal and ergodic. Therefore, the states can be revisited in finite time intervals. This structure enables one to decide on the optimal state order by simple experiments on the training data and requires no size normalization.

The circular HMM is tested on the optical character recognition problem based on the boundary features. Although the computational complexity is the same, the circular HMM has many superiorities compared to the left-right model. First of all, the circular HMM does not require to increase the number of states as the size of the boundary increases. Therefore, it is size invariant. Secondly, circular HMM does not require as many non-zero state transition probabilities as the left-right models. Therefore, the computational complexity of the circular HMM is less than the left-right models or other more complicated topologies for the same recognition rates.

In Section 2, the circular HMM and its mathematical representation is introduced. In Section 3, an application of circular HMM to optical character recognition problem is presented. Finally, Section 4 concludes the paper and gives the experimental results.

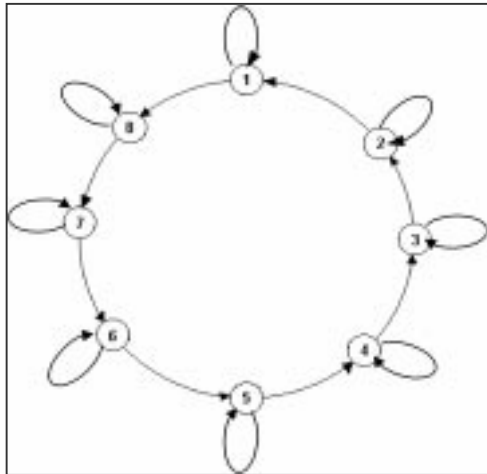


Figure 1. Circular HMM for S=8 and N=1

2. THE CIRCULAR HMM

Suppose that a shape can be characterized by its discrete set of boundary points drawn from a finite alphabet or from quantized vectors of a code-book. Suppose, also, that the boundary string is the observable output of a parametric random process. Let, $O = (O_t, O_{t+1}, O_{t+2}, \dots, O_{t+T-1})$ represents the closed boundary of length T, over an alphabet $V = \{v_1, \dots, v_k, \dots, v_M\}$, with $\forall t, O_t = O_{t+T}$. Our goal is to define a discrete density Hidden Markov Model, which represents each boundary class and labels an unknown boundary.

The circular HMM for each boundary class $l = 1, \dots, c$ is represented by a three tuple $\lambda_l = \{A_l, B_l, S\}$. The state transition probability matrix, $A_l = [a_{ij}]$ and the observation probability sequence of observing the code k in i^{th} state for $1 \leq i, j \leq S$, $B_l = \{b_i(k)\}$ satisfies the following conditions:

- 1) $j = i+n, \quad n=0,1,\dots,N$
- 2) $a_{ij} = a_{i+S,j+S}$,
- 3) $b_i(k) = b_{i+S}(k)$ and
- 4) $N \ll S$,

where S represents the number of states and N represents the maximum number of difference between i and j. Notice that the state transition probability matrix, where each entry, a_{ij} , represents the probability of moving from state i to j is still very sparse ($N \ll S$) as in the left-right HMM. For example for N=1, the State transition matrix has the following form:

$$A_l = \begin{pmatrix} a_{11} & a_{12} & 0 & 0 & 0 & \dots & 0 \\ 0 & a_{22} & a_{23} & 0 & 0 & \dots & 0 \\ 0 & 0 & a_{33} & a_{34} & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{S1} & 0 & 0 & \dots & 0 & \dots & a_{SS} \end{pmatrix}$$

The probability of observing a specific boundary sequence by a HMM is obtained as the sum of the probabilities of

observing this sequence for every possible state sequence of the HMM, i.e.:

$$P(O|\lambda_l) = \sum_{all Q} P(O, Q|\lambda_l) ,$$

where Q is the hidden state sequence, which generates the given observation sequence O and λ_l is the HMM model for l^{th} boundary class. Adjusting the A_l and B_l parameters of a HMM model, we may obtain high $P(O|\lambda_l)$ probability values for observations from the true class and low probabilities for false ones.

A popular algorithm for the parameter adjustment is the Baum-Welch method [1] which is an iterative update and improvement of HMM parameters according to the training set of the coded patterns. In the recognition stage, the probability of observing the coded patterns by every HMM is calculated. Then, the observed string is labeled with the class which maximizes the probability $P(O|\lambda_l)$. Computation of $P(O|\lambda_l)$ for each λ_l requires an iterative process, called forward-backward algorithm [6].

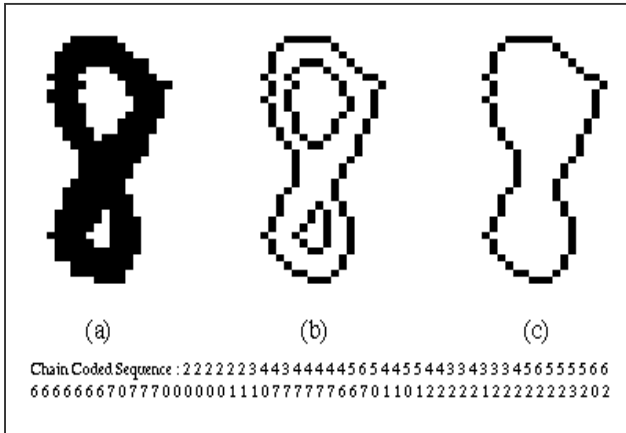


Figure 2. Coding of the handwriting using freeman's chain codes (a) binary image, (b) boundary extracted image and (c) outer contours .

3. OPTICAL CHARACTER RECOGNITION (OCR) WITH THE CIRCULAR HMM

Although the proposed HMM topology is applicable to any shape recognition problem, in this study, it is tested on the handwritten character recognition problem, since this problem is well defined and investigated for a long time. There are standard handwriting databases available

to cast consistent comparison platforms. Hundreds of research articles and patents are available in the literature and dozens of commercially products are available in the market. In spite of the intensive effort the Optical Character recognition of free style hand writing problem is not fully solved yet. Rather than developing an optical character recognizer, the goal of this study is to show the power of the circular HMM in the shape recognition problems, using a HMM based shape recognizers.

First, the boundaries of characters are extracted from the binarized document image. Then, they are coded by Freeman's chain code, as indicated in Figure 2. The cursive handwriting is segmented by the algorithm proposed in [3]. Since our goal is not to design the best optical character recognizer, only the outer boundaries for each character are used for recognition. However, inclusion of the inner boundaries definitely yields better results in the OCR problem. However, it makes no difference in comparing the proposed HMM topology to the other topologies, studied in the literature. The coded boundaries are used as feature vectors of a discrete density HMM model, with varying sizes and topologies. The effect of changing the size and non-zero state transitions on the recognition rates are investigated.

4. EXPERIMENTS AND CONCLUSION

The experiments are performed on the NIST-SD 7 (Special Database for Handwritten Numbers). C programming language is used on a UNIX workstation environment. In the preprocessing stage, the binary image is smoothed by an averaging filter and the outer contour of each number is coded by the Freeman's chain code. This coding scheme yields observation sequences of the HMM with varying length, depending on the size and type of the number digits. Evidentially, some numbers (like; 1's, 2's) have relatively shorter observation sequences then the others (like; 4's or 9's). It is observed that the length of the observation sequence varies between 70-120 codes for number digit in a frame of 32x32 pixels.

First, the experiments are performed on the discrete density left-to-right HMM of various sizes and topologies. 100 samples from each class are used for training. Then the remaining samples of NIST-SD 7 database are fed to the recognizer. The optimal topology, which makes the recognition rate maximum is investigated by trial and error. For this purpose, the number of states (S) of the HMM is increased gradually and for each S the number of non-zero state transition probabilities (N) are increased to reach the full rank of the state transition probability matrix. It is observed from Table 1 that for each S , the

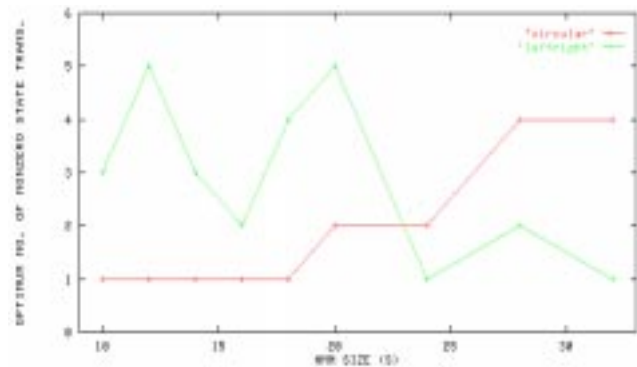
optimum number of the non-zero state transition is different. For example; for $S=10$ states, the optimum number of non-zero state transition is $N=4$ which gives 84.2% recognition rate. For $N=5$ the recognition rate decreases to 79.8%. Similar trends are observed as we increase the HMM size (See: Table 1). This result indicates the sensitivity of the HMM recognizer to the HMM topology. Therefore, HMM topology depends on the number of states for the left-to-right model. As we increase the number of states, the recognition rates increase. After a certain size, there is no significant improvement on the recognition rates. For very large sizes ($N>32$) of HMM the recognition rates start to gradually decrease. The length of the observation sequence is normalized to a fixed size for this case. The starting and terminating points of the boundaries are manually selected. It is observed that if the starting and terminating point of the observation sequence are not selected carefully or the observation sequences are not normalized to a fixed size, the recognition rates get as low as 60%. Table 1 indicates that the recognition rates are achieved in between 79-92%.

The experiments performed on the circular HMM did not pay any attention to estimate the initial points on the boundary. The observation sequences are not normalized to a fixed size. Therefore, it requires less preprocessing power compared to the left-to-right HMM. For relatively small HMM sizes ($N=10, 12, 14, 16, 18$), the highest recognition rates are always for $S=1$, indicating the best nonzero state transition number for each size. As we increase the size S , the optimum topology which makes the recognition rates maximum requires larger S , indicating a meaningful relation between the HMM size and number of nonzero state transition (See; Figure:3). This result indicates the stableness of the circular HMM to the size variations. Nevertheless, keeping $S=1$ and increasing HMM size gives the most efficient method for identifying the optimum state size S . Because, this approach steadily increases the recognition rates with the least amount of computational complexity. For $N=1$, the recognition rates of the circular HMM varies between 86-92% as the state size increases from $S=10$ to $S=32$. Figure: 4 indicates the maximum recognition rates at each HMM size for the circular HMM and the left-to-right HMM. As it is observed from this figure, circular HMM has superior performance for all sizes except for $S=24$. Note also that, the circular HMM requires less computational complexity compared to the classical HMM topologies because the optimal nonzero state transition for

circular HMM is always less then the optimal nonzero state transition of the classical models.

5. REFERENCES

- [1] S. Connel, " A Comparison of Hidden Markov Model Features for the Recognition of Cursive Handwriting", Computer Science Department, Michigan State University, MS Thesis (1996).
- [2] Atici, F.T. Yarman-Vural, " A Heuristic Method for Arabic Character Recognition, Journ. of Signal Processing, vol. 62, issue.1, 1997.
- [3] N. Arica, F. T. Yarman-Vural " One Dimensional Representation of two dimensional Information for HMM Based Handwritten, Proc. of the IEEE-Int. Conf. on Image Processing, Chicago, 1998.
- [4] M.A. Özdil, F.T. Yarman,Vural, Optical Character Recognition Without Segmentation, Proc. of the Int. Conf. on Document Analysis and Recognition Ulm-Germany, 483-487, (1997).
- [5] N. Arica, F. T. Yarman-Vural, " A New Scheme for Off-Line Handwritten Connected Digit Recognition" Proc. of the IAPR- Int.Conf. on Pattern Recognition, 1998.
- [6] B.H.Juang and L.R.Rabiner, (1990), "The Segmental K-



means Algorithm for Estimating Parameters of Hidden Markov Models", IEEE Trans. on A.S.S.P., Vol:38, No:9.

Figure 3. Relationship between the state size and optimum topology.

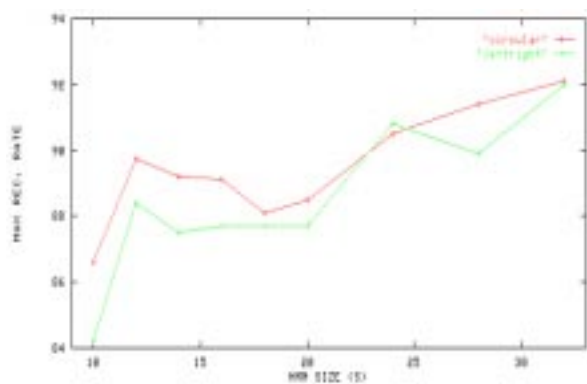


Figure 4. Performance of the left-to right and circular HMM for the best topology.

	3	83.1	87.4
	4	83.2	87.7
	5	87.7	87.1
24	1	90.8	89.8
	2	89.5	90.6
	3	88.1	89.3
	4	88.6	90.6
	5	90.5	90.6
28	1	85.8	90.7
	2	89.9	90.3
	3	89.5	90.2
	4	87.3	91.4
	5	89.8	91.0
32	1	92.0	92.1
	2	90.8	91.8
	3	84.6	91.7
	4	86.2	92.4
	5	88.5	90.6

TABLE 1

No. of States	No. of Nonzero State Trans.	Rec. Rates in <u>Left-to-right</u> HMM	Rec. Rates in <u>Circular</u> HMM
10	1	82.7	86.6
	2	83.1	80.1
	3	84.2	77.5
	4	82.3	75.6
	5	79.8	77.1
12	1	88.0	89.7
	2	87.6	88.7
	3	84.3	77.4
	4	85.0	80.4
	5	88.4	77.1
14	1	84.5	89.2
	2	85.7	88.1
	3	87.5	83.6
	4	85.3	80.6
	5	86.4	81.8
16	1	87.4	89.1
	2	87.7	86.7
	3	87.3	85.6
	4	87.4	83.4
	5	87.3	84.2
18	1	85.3	88.1
	2	85.4	87.6
	3	86.2	85.0
	4	87.7	84.9
	5	87.5	84.8
20	1	87.2	86.5
	2	87.2	88.5