

A Shape Descriptor Based on Circular Hidden Markov Model

Nafiz Arica and Fatos T. Yarman Vural
Computer Engineering Department,
Middle East Technical University, Ankara / Turkey
{nafiz,vural}@ceng.metu.edu.tr

Abstract

Given the shape information of an object, can we find visually meaningful "n" objects in an image database, which is ranked from the most similar to the nth similar one? The answer to this question depends on the complexity of the images in the database and the complexity of the objects in the query.

This study represents a robust shape descriptor, which compares a given object to the objects in an image database and identifies "n" shapes, ranked from the most similar to the least similar one, in the database. The intended shape descriptor is based on the circular Hidden Markov Model (HMM) proposed by the authors of this study. Circular HMM is both ergodic and temporal. It is insensitive to size changes. Since it has no starting and terminating state, it is insensitive to the starting point of the shape boundary. The experiments, performed on 100 test shapes, indicate excellent result.

1. Introduction

The wide spread use of Internet and accumulation of vast amount of images in databases in many disciplines, necessitates developing database queries based on the image descriptors. Existing image file formats have very little emphasis on the content information of the image. Because of the great demand, the current research in image analysis has shifted towards content-based representation of images and developing a content-based image file format for storage and retrieval purposes. The MPEG-7 activities are concentrated on finding a set of descriptors for color, texture and/or shape. Nowadays, researchers are focused on developing algorithms, which can successfully extract color, texture and shape information in a given set of images. Although a lot of progress is achieved in specific applications, the problem of representing the content information for generic and large image databases remains unsolved.

In this study, we attack the content-based image representation problem from the shape side. Hidden Markov Model is used to define a shape descriptor. It is well known that Hidden Markov Model (HMM) is a

widely used powerful tool for many image analysis problems. There is a tremendous amount of variations of HMM applications, which input various feature sets into various HMM sizes and topologies. Efficient iterative algorithms are available for estimating the model parameters of observation probability and state transition matrices. However, all the approaches presume a model size and topology [1]-[3]. Unfortunately, there are no effective methods for estimating the optimal number of states and/or the nonzero state transitions for a specific feature set. All of the practiced topologies have their advantages and disadvantages. The vast amount of HMM topologies, used in various application domains, can be investigated in two categories: Ergodic and temporal topologies.

Ergodic topologies enable the revisits of each state with probability one in finite intervals, by allowing nonzero state transition paths between any two states. However, they do not impose a sequential or temporal order. Therefore, when the observation vector is temporal or an ordered sequence, ergodic models do not fully utilize the sequential 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 diagonal entries. Therefore, in most of the pattern description 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 content description 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. The periodic nature of the boundary requires a Hidden Markov topology which is both temporal (has a sequential order) and ergodic to allow the revisits of a state as the boundary returns to the starting point and repeats itself.

In this study, a new topology, called circular HMM, is used as a shape descriptor [4]. 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 description problem. The circular HMM is both temporal and ergodic. Therefore, the states can be revisited in finite time intervals while keeping the sequential information in the string which represents the shape. 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, as a shape descriptor, has many superiorities compared to the classical topologies in the literature. First of all, the circular HMM does not require increasing 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 classical topologies. Therefore, the computational complexity of the circular HMM is relatively less than the other topologies for the same recognition rates.

In Section 2, the theoretical basis of circular HMM, as a shape descriptor, is described. In section 3, the proposed shape descriptor is tested on a data set which is created from 10 basic test images by using deformation operations such as pruning, clipping, skew, stretching and squishing. Section 4 concludes the paper and directs the future studies.

2. A Shape Descriptor: 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 and rank the boundaries in an object database according to their similarities to the object under consideration.

The circular HMM for each boundary $l = 1, \dots, c$, in an image database, 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 very sparse ($N \ll S$) as in the left-right HMM. For example for $N=2$, the State transition matrix has the following form:

$$A_l = \begin{pmatrix} a_{11} & a_{12} & a_{13} & 0 & 0 & 0 & \dots & 0 \\ 0 & a_{22} & a_{23} & a_{23} & 0 & 0 & \dots & 0 \\ 0 & 0 & a_{33} & a_{34} & a_{34} & 0 & \dots & 0 \\ \vdots & & & & & & & \\ a_{S1} & a_{S2} & 0 & \dots & \dots & \dots & 0 & a_{SS} \end{pmatrix}$$

The similarity of a given object, represented by the observation sequence O , to an object in the database, can be measured by the conditional probability of observing O , given the object model λ_l in the database. This probability is obtained as the sum of the probabilities of observing the sequence O , for every possible state sequence of the HMM, i.e.:

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

where Q is the hidden state sequence, which generates the given observation sequence O and λ_l is the HMM model for l^{th} object in the database. The model parameters A_l and B_l define a HMM model for every shape in an image database. We expect to obtain high probability values $P(O|I_l)$, for similar shapes and low probability values for distinct ones.

A popular algorithm for the parameter estimation is the Baum-Welch method [3], which is an iterative update and improvement of HMM parameters according to the

training set of the coded shapes. In order to measure the similarity between the modeled shape and an observation sequence, the probability of observing the coded shape by every HMM is calculated. Then, the maximum probability $P(O/\lambda_l)$ indicates the most similar shape while the minimum probability indicates the least similar one. Computation of $P(O/\lambda_l)$ for each λ_l requires an iterative process, called forward-backward algorithm [3].

In this study, the object boundaries are characterized by 8-directional Freeman's chain code. Therefore, the observation sequence, $O = (O_t, O_{t+1}, O_{t+2}, \dots, O_{t+T-1})$ represents the closed boundary of length T , over a set of integers, $V = \{0, 1 \dots 7\}$, with $\forall t, O_t = O_{t+T}$.

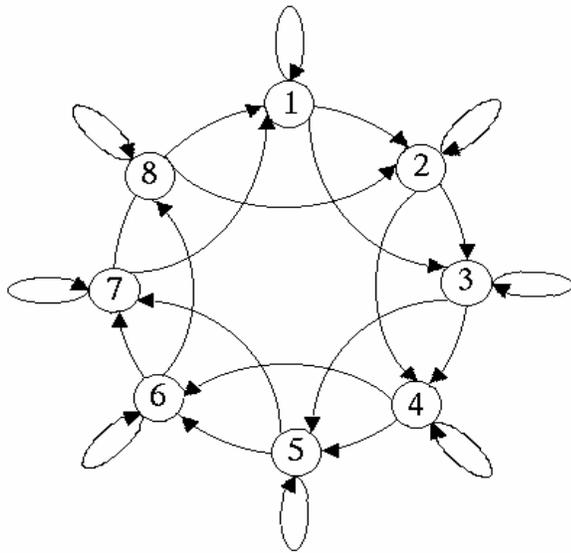


Figure 1. Circular HMM for $S=8, N=2$.

3. Experiments

In order to form a test set, first, 10 distinct images are selected from a picture alphabet (House, plane, bird, spade, hand, telephone, goose, dog and cow). Second, each of the original images is deformed by stretching, squishing, skewing, pruning and clipping. 10 similar images are generated by deforming each distinct image. The experiments are performed on these 100 test images, (see: Figure 2), in C programming language on a workstation environment.

The boundaries of objects are extracted from the binarized images. Then, they are coded by 8-directional Freeman's chain code. The length of the coded sequence changes between $T = 120-540$. The coded boundaries are used as feature vectors of a discrete density Circular HMM model. The model size and nonzero state transitions are obtained by trial and error, as $N=2$ with $S=30$ states. It is observed that increasing the model size,

S up to 50 states, has no effect on the similarity rank. However, increasing the number of non-zero state transition, N more than 5, change the ranking slightly for this data set. Each of the 100 images is modeled by the circular HMM.

Two sets of experiments are performed to test the robustness of the proposed shape descriptor for deformation of shape and size invariance. In order to test the robustness of the circular HMM to shape deformation, for each sample image in the dataset, $P(O/I_l)$ is calculated and ranked from the highest to the lowest probability values. In Figure 3, the first column indicates the object under consideration. Each row indicates the objects, ranked from 1 to 9 according to the probability values between the most similar and least similar ones. It is clearly seen that the proposed shape descriptor is quite insensitive to the deformations. The results of the ranking are very consistent with the human visual system.

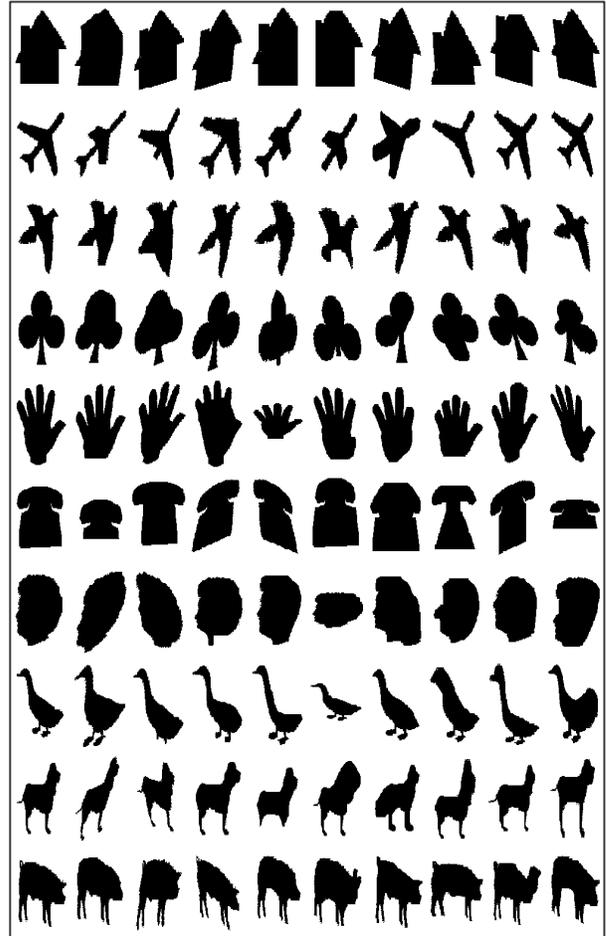


Figure 2. Test Data modeled by Circular HMM.

The robustness of the descriptor to the size invariance is tested by varying the size of the objects in the database. For this purpose, one of the samples in the database is scaled up to 2, 3 and 4 times larger than the original size. During this process the rest of the objects are kept in the same size. The scaled object is, then, compared to the rest of the objects in the data set. This task is achieved by computing and ranking $P(O/\lambda_l)$, for each enlarged object O_e and the objects modeled by $\lambda_l, l=1, \dots, 10$, in the data set. Figure 4 indicates test samples of the plane image and the ranked similarity when the test objects are enlarged 2, 3 and 4 times. Experiments indicate very satisfactory results for the rest of the objects in the database. The circular HMM easily finds the visually similar objects in every case even if the length of the object boundary is the several multiples of the rest of the boundaries in the database.

Object	Most Similar	----->	Least Similar
			
			
			
			
			
			
			
			
			
			

Figure 3. Some test objects and the ranks from the most similar object to the least similar one.

4. Conclusion

This study presents a robust shape descriptor for identifying the similar objects in an image database. The proposed descriptor is based on HMM with circular topology. It catches the similar shapes in an image database, successfully. It pays no attention to estimate the initial points on the shape boundary. It is invariant to shape size. Therefore, no pre-processing is required for normalizing the shape sequence to a fixed size.

The circular nature of the HMM topology makes the model topology (N and S) quite stable for shape description. In other words, changing the model size and non-zero state transition in a certain range has no effect on the similarity rank. On the contrary of the classical topologies, the model topology is quite insensitive to length of the observation sequence.

In conclusion, the structure of the circular HMM, makes it a very robust descriptor for shape-based content representation.

Object	Most Similar	----->	Least Similar
			
2 times enlarged			
3 times enlarged			
4 times enlarged			

Figure 4. A test object enlarged 2, 3, and 4 times and ranked from the most similar object to the least similar one.

References

- [1] A. Atici, F.T. Yarman-Vural, " A Heuristic Method for Arabic Character Recognition, J. of Signal Processing, vol. 62, issue.1, pp.87-99, 1997.
- [2] 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, pp. 986-991 1998.

- [3] 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.
- [4] N. Arica and F. T. Yarman-Vural, "A New HMM Topology for Shape Recognition", Proc. of the IEEE-EURASIP Workshop on Nonlinear Signal and Image Processing, Vol II, pp. 756-761, 1999.