Cai, Jinhai and Ee, Dominic and Smith, Robert (2007) Image Retrieval Using Circular Hidden Markov Models with a Garbage State. In *Proceedings Image and Vision Computing New Zealand 2007*, pages pp. 115-120, Hamilton, New Zealand.

This is the author-manuscript version of this work - accessed from http://eprints.qut.edu.au

Copyright 2007 (please consult authors)

Image Retrieval Using Circular Hidden Markov Models with a Garbage State

Jinhai Cai, Dominic Ee, and Robert Smith

Faculty of Information Technology Queensland University of Technology QLD 4001, Brisbane, Australia Email: j.cai,r.smith@qut.edu.au,~m.ee@student.qut.edu.au

Abstract

Shape-based image and video retrieval is an active research topic in multimedia information retrieval. It is well known that there are significant variations in shapes of the same category extracted from images and videos. In this paper, we propose to use circular hidden Markov models for shape recognition and image retrieval. In our approach, we use a garbage state to explicitly deal with shape mismatch caused by shape deformation and occlusion. We will propose a modified circular hidden Markov model (HMM) for shape-based image retrieval and then use circular HMMs with a garbage state to further improve the performance.

To evaluate the proposed algorithms, we have conducted experiments using the database of the MPEG-7 Core Experiments Shape-1, Part B. The experiments show that our approaches are robust to shape deformations such as shape variations and occlusion. The performance of our approaches is comparable to that of the state-of-the-art shape-based image retrieval systems in terms of accuracy and speed.

Keywords: Shape-based image retrieval, Hidden Markov model, Garbage state, Shape deformation

1 Introduction

Shape recognition is a vital part in content-based image retrieval (CBIR) using high-level semantics. This is because shape recognition can be used to bridge the gap between low-level features and highlevel semantics. Most previous work on contentbased image retrieval used colours, textures, histograms and shapes [1]. In this paper, we will propose use constrained circular hidden Markov models for shape-based image retrieval.

Shapes of objects can be extracted from images and videos via segmentation and motion analysis. The ability of recognizing objects regardless of their position, size, rotation and reflection is very important to many real world applications. Contour-based features of object shapes have been widely used to achieve this ability. The main advantages of using contour-based features for pattern recognition are: (1) It is extremely efficient to use contours for coding; (2) It is easy to obtain object contours from files in MPEG-7 which is a standard format for meta-data; (3) Single closed shapes can also be extracted from videos using motion detection. Figure 1 gives an example of extracting shapes from motion analysis. Due to these advantages, we adopt contour-based features for shape



Figure 1: 2D shapes from motion analysis.

recognition and image retrieval. Shape recognition algorithms using contour-based features can be divided into two main categories: global featurebased [2, 3] and local feature-based [4, 5, 6]. Fourier descriptors have been widely used in pattern recognition as the absolute values of the descriptors are invariant to rotation and reflection which is a desirable property in some applications as hand-tool recognition [3]. As Fourier descriptors use global shape features, they are sensitive to deformations caused by noises, occlusion and etc. In order to solve this problem, most recently developed methods use local features to represent shapes. Among them, curvature scale space (CSS) shape descriptors [7, 8] and polygonal shape descriptors [9] have demonstrated their successes in many applications. A CSS image provides a multi-scale representation of the curvature zero crossings of a closed contour. However, the CSS representation has two drawbacks. The CSS is inadequate to represent convex segments of contours [8] and some similar shapes may produce quite different CSS images. Polygonal shape descriptors approximate 2D shapes by polygons. In such a system features are generated by polygonal approximations and the recognition decision is made by using classification trees. However, it is difficult to automatically obtain the classification trees and the recognition based on semantics was sensitive to shape variations and noise.

Hidden Markov models (HMMs) are powerful in coping with large variations in shapes. Many HMMbased methods have to find a reference point for each image in order to obtain well trained models or templates. To find a reference point in an image is a not easy task, significant deformation makes this task even more difficult. Fortunately, circular HMMs are available for 2D shape recognition. In circular HMMs, any state can be the initial state and any state can be the ending state, therefore circular HMMs can use any point as a starting point. However, there is no constraint between starting states and ending states in conventional circular HMMs, consequentially they have limitations on modelling shapes at starting points/ending points.

The motivation of this paper is to overcome the limitation by imposing some constraints on circular HMMs, therefore we are able to further improve their performance on 2D shape classification.

2 Hidden Markov model

The contours of 2D shapes can be represented by sequential feature vectors, which are regarded as the equivalence of time-varying signals. In a short segment, the signal can be considered as a stationary process with minor fluctuations. Therefore, it is suitable to use an individual state of an HMM to model the steady statistical information in a short segment. Significant changes of properties over segments can be modelled by state transitions.

Generally, an HMM with N states can be characterized by: (1) the state transition probability matrix, $\mathbf{A} = \{a_{ij}\}$, where a_{ij} is the probability of making the transition from state *i* to state *j*; (2) the output probability matrix, $\mathbf{B} = \{b_j(v)\}$, where *v* can be a discrete observation symbol or continuous observation vector or even a hybrid observation vector; (3) the initial state distribution, $\pi = \{\pi_i\}$. Now an HMM can be represented compactly as:

$$\lambda_w = (\mathbf{A}, \mathbf{B}, \pi), \tag{1}$$

where λ_w is one model in the collection λ . It is well known that **A** determines the HMM topology. If there is no constraints on **A**, the HMM is ergodic. If we want a circular HMM, we simply impose the following constraint on **A**

$$a_{i,j} = 0,$$
(2)
((N + j - i) mod N) > d > 1 & 0 \le i, j < N,

where d is the maximum state gap. Figure 2 gives an example of a circular HMM with N states and d = 2. Circular HMMs are robust to variations as



Figure 2: A circular HMM with N states and d = 2).

they are statistic models. Circular HMMs are also suitable to represent closed contours due to their topology. We will use Circular HMMs to model shapes for image retrieval.

3 Constrained circular HMM

3.1 The limitations of conventional circular HMMs

For a left-right HMM, the initial state can only be the one from the initial state set $\{s_i | \pi_i > 0\}$. In some applications, constraints are imposed on the final states as well to further improve the performance of these systems [10]. While, there is no constraint on the initial state and the final state for a circular HMM, i.e. any state of the circular HMM can be the initial state and the final state. As a result, circular HMMs have natural immunity to the rotation of shapes. However one of the side effects is that circular HMMs have difficulties in distinguishing some shapes which human can do easily. Figure 3 and 4 give some examples of shapes that are confusing to circular HMMs.

To highlight the problem, let's assume that a circular HMM is well trained for the shape shown in Figure 3a. Without losing generality, we assume that the initial state is $s_0 = 0$ and each edge is modeled by one state. So the polygonal shape in Figure 3a is modeled by a circular HMM with 14 states. Now we use the shape in Figure 3b to

evaluate this HMM, a quite high probability would be produced by this circular HMM. The most likely state sequence would be $\{0, 1, \dots, 10, 11, 0, 1\}$. In the same way, if we use a circular HMM trained by shapes similar to the one shown in Figure 3b, the HMM would produce a quite high probability when the shape in Figure 3a is used for evaluation. This is because that the shapes in Figure 3a and 3b are made up using a common component shown in Figure 3c and there is no constraint on the relation between initial states and final states. Figure 4 shows other interesting shapes, where their relationship is the whole-part relationship. If a circular HMM is trained by shapes similar to Figure 4b and tested by the shape in Figure 4a, the output probability will be high, but not vice versa. These examples have revealed the limitations of circular HMMs.



Figure 3: Confusing shapes to circular HMMs: same components.



Figure 4: Confusing shapes to circular HMMs: Whole-Part

3.2 The constraint on the initial and final states

The solution to this problem is to impose a constraint on the initial state and the final state of a given circular HMM. Let take a circular HMM with 14 states and d = 2 as an example. If the initial state is s_0 , we force the final state of the HMM to be s_{12} , s_{13} or s_0 . In this way, the constrained circular HMM trained by the shape in Figure 3a will produce a low probability when the shape in Figure 3b is evaluated.

To define the constraint on states for a given HMM in a general case, let the initial state be s_j and the final state be s_i . The constraint on the initial state and the final state can be formally defined as follows.

$$a_{ij} \neq 0, \qquad 0 \le i, j < N. \tag{3}$$

3.3 The implementation of the constrained circular HMMM

In order to impose the constraint on the initial and final states, we have to know the initial and the final states or to find the best state sequence. In this paper, we use the Viterbi algorithm due to following reasons:

- The state sequence of HMM is not observable. However, the best state sequence can be found by the Viterbi algorithm.
- The Viterbi algorithm makes it possible to replace the multiplication in probability computation with summations. Consequently, the scaling procedure, which is necessary in the Baum-Welch algorithm, can be avoided. This results in better computational performance.

3.3.1 The solution with known initial state

From the standard Viterbi algorithm, you can find that it is impossible to know the initial state for a given final state without using the state-path backtracking. The easiest solution is to consider each state as an initial state in a loop. The constraint can be imposed at the step of termination. Therefore we have the modified Viterbi algorithm for our constrained circular HMM (**CCHMM1**) as:

Loop:
$$s_0 = k$$
, $0 \le k \le N$

Step 1: Initialization
$$(t = 0)$$

 $\alpha_0(k) = b_k(O_0),$

Step 2: Recursion $\alpha_t(j) = \max_i [\alpha_{t-1}(i)a_{ij}b_j(O_t)],$

Step 3: Constraining

$$p^*(\mathbf{O}|\lambda_w, s_0 = k) = \max_s[\alpha_{T-1}(s)a_{sk}],$$

End of the loop: Termination $p^*(\mathbf{O}|\lambda_w) = \max_{s_0}[p^*(\mathbf{O}|\lambda_w, s_0)].$

As this algorithm does not need the state-path backtracking, we can implement a similar algorithm based on the Baum-Welch algorithm. However, the complexity of this algorithm is in the order $O(N^2T)$ instead of O(NT).

3.3.2 The solution with the state-path backtracking

Another solution is to find the best state sequence for a given final state using the state-path backtracking. In order to do this, we use a variable $\Phi_t(i)$ to track state transition. In our algorithm, the state-path backtracking will be employed after recursion so that the constraint can be imposed before the highest probability is worked out. Our modified Viterbi algorithm for our constrained circular HMM (**CCHMM2**)is:

Step 1: Initialization
$$(t = 0)$$

 $\alpha_0(i) = b_i(O_1),$
 $\Phi_0(i) = 0 \quad 0 \le i \le N - 1.$

- Step 2: Recursion $\begin{aligned} \alpha_t(j) &= \max_i [\alpha_{t-1}(i)a_{ij}b_j(O_t)], \\ \Phi_t(j) &= \arg\max_i [\alpha_{t-1}(i)a_{ij}] & 1 \le t < T, \end{aligned}$
- Step 3: State-path backtracking $s_t = \Phi_{t+1}(s_{t+1}), \qquad (0 \le t < T-1)$
- Step 4: Constraining $\alpha_{T-1}(s_{T-1}) = \alpha_{T-1}(s_{T-1})a_{s_{T-1}s_0},$
- Step 5: Termination $p^*(\mathbf{O}|\lambda_w) = \max_{s_{T-1}}[\alpha_{T-1}(s_{T-1})].$

The computational complexity of this algorithm is in the order of O(NT). However, the significant reduction of the computational complexity comes at the cost of optimization. It is easy to find that **CCHMM1** in Section 3.3.1 keeps all $N \times N$ paths from all possible initial states to all possible final states. Therefore, $p^*(\mathbf{O}|\lambda_w)$ is the probability from the best state-path after applying the constraint. In this algorithm, however, the best state-path might be blocked by some "promising" paths during the recursion. This is because that these "promising" paths might be false after applying the constraint in eqn 3. As a consequence, $p^*(\mathbf{O}|\lambda_w)$ calculated by this algorithm might be not the probability from the best state-path. We will discuss this issue in the Section 5.

4 Circular hidden Markov models with a garbage state

For image retrieval using HMM, we need to obtain an HMM using only one query image. Therefore we have to train an HMM using one image. To deal with this problem, we use the simple architecture of the circular HMM and we use a small number of states for each HMM. We also force the minimum duration of a state to be 3 to reduce the chance of ill-conditions, which is the variance of a state parameter tends to be zero. However, some states have only 3 points of features to estimate the Gaussian functions, it is inevitable that the estimated parameters of these Gaussian functions are not accuracy and some values are close to zero. As a result, a little mismatch can cause significant penalty. In order to alleviate this problem, we set a minimum deviation for each parameter, where the minimum deviation is set to the value at the bottom 30%. As speed is very important to image retrieval, we use CCHMM2 in this experiment.

HMMs are robust to minor shape variation but sensitive to occlusion and major deformation. In order deal with this problem, we introduce two constants: the garbage state score (GSC) and the minimum state score (MSC). The GSC is the average score, $p(\mathbf{O})$, of using a single state to model a random shape. The MSC is much smaller than GSC. We set MSC = GSC/3. The garbage states can occur at any position but only one sequential segment on the whole contour and its minimum duration is 4 so that individual abnormal pointers would not be modelled by the garbage state instead of MSC. The architecture of the constrained circular HMM as shown in Figure 5.



Figure 5: Circular HMM with d = 1.

5 Experiments

We conducted our experiments using the shape database for the MPEG-7 core experiment CE-Shape-1, Part B [11]. In our experiments, each contour is re-sampled with only 128 points and the length of each contour is normalized to a constant. Then each contour is smoothened by passing a low-passing filter. Constrained circular HMMs are insensitive to starting points, but they are sensitive to reflection. Therefore, we input two contours to HMMs for a given shape, one original and one with reflection. The first experiment is to verify the improvements of the constrained circular HMMs over standard circular HMMs on 2D pattern recognition. The second experiment is to evaluate the performance of constrained circular HMMs with a garbage model on shape-based image retrieval.

5.1 Constrained circular HMMs for 2D pattern recognition

There are 1400 shape samples of 70 classes in the database, and each shape class has only 20 samples. Therefore, we used 20 fold cross validation to carry out our experiments, i.e. leaving one out for testing. In our first experiment, we selected a sub-database including 7 shape classes so that we can compare the performance of the proposed algorithms to that of others [12] using the same sub-database. The selected 7 classes are Bone, Fork, Fountain, Glass, Hammer, Heart and Key. Some shapes of these 7 classes are shown in Figure 6.

In the experiment, we used three different algorithms described in Section 3.3 to train HMMs and used different testing algorithms to evaluate the performance of these trained HMMs on recognizing 7 classes. The number of states in one HMM varies from 8 to 40 according to the overall curvature in the training set of the class. The state probability density function used in this experiment is the Gaussian mixture density, which has only two components. The results are given in Table 1.



Figure 6: Shape samples from 7 classes

 Table 1: The performance on sub-database with 7 classes

Training method	Testing method	Accuracy
CCHMM1	CCHMM1	100%
CCHMM1	CCHMM2	100%
CCHMM2	CCHMM2	100%
CCHMM2	No constraint	97.9% (137/140)
No constraint	No constraint	96.4% (135/140)

The results have shown that both CCHMM1 & CCHMM2 can achieve the accuracy of 100%, while the best result reported in [12] is 98.8%. As the computational complexity of CCHMM2 is $O(N \times T)$, it is suitable for many real-time applications such as image retrieval.

As the shapes in Figure 6 are quite distinctive, it is not difficult to achieve good performance on such a database. In order to test our algorithms on recognizing 2D shapes with significant deformations, we used all shapes from the database in the second experiment. The training and testing procedure used in here is the same as that in the first experiment. The results are given in Table 2.

Table 2: The performance on the MPEG-7 CE-Shape-1, Part B

T	raining method	Testing method	Accuracy
	CCHMM1	CCHMM1	91.6%(1283/1400)
	CCHMM2	CCHMM2	90.8%(1271/1400)
	No constraint	No constraint	85.1% (1192/1400)

The results confirm that the constraint on the initial and final states can improve the performance of 2D shape recognition. It is very interesting to compare the performance of CCHMM1 & CCHMM2. CCHMM1 is slightly better than CCHMM2 in terms of accuracy but N times slower than CCHMM2 in terms of speed. After all, CCHMM2 can achieve very good performance, which is very close to that of CCHMM1, with the same computational complexity of the standard Viterbi algorithm in big-O notation. Comparing to elastic matching algorithms, which widely applied in image retrieval are in the order of $O(T^2)$, CCHMM2 is much faster as the number of states in an HMM is usually much smaller than the number of points in a contour.

5.2 Constrained circular HMMs for image retrieval

Speed is very important to image retrieval, so we use CCHMM2 in this experiment. We used MPEG-7 core experiment CE-Shape-1 Part B [11] as the database in this experiment. This is mainly because that we want to compare the performance of our approach with that of many other algorithms using the same database. We use each image in the database in turn as a query and give a list of 40 relevant images whose scores are in the top 40 among 1400 images. The Table 3 gives the retrieval rates of the proposed method, Constrained Circular HMM with a garbage state (CCHMM-G), and others, where results of P298, P320, P517, P567, P687 and DAG are from [4] and the result of WARP is from [13].

Table	3:	The	perfor	rma	nce	on	image	re	etriev	val
	Tr	aining	; metho	bd	Re	etrie	val rate	9		
i		COTT	DIC			=0	0.04			

framing method	itetiievai rate
CCHMM-G	72.0%
P298	76.5%
P320	75.4%
P517	70.3%
P567	67.8%
P687	70.2%
DAG	60.0%
WARP	59.5%

It is very interesting to investigate the influence of occlusion on image retrieval, as the proposed HMMs (CCHMM-G) have the garbage state that is specially designed to alleviate the problems caused by deformation and occlusion. During the test, we use other shapes to create partial occlusion to the query image. Figure 7 shows a screenshot of using a hand-drawn image as a query and Table 4 gives the performance of the proposed method under different occlusion conditions.



Figure 7: Using hand-drawn image as a query

If we use CCHMM2 without the garbage state, we can get a slightly better performance (72.4% retrieval rate) than that of CCHMM-G, however, the performance degrades rapidly with the increase of occlusion percentage. When the occlusion is 20%,

 Table 4: The performance of image retrieval using

 CCHMM-G

Amount of occlusion	Retrieval rate
0%	72.0%
10%	67.9%
20%	66.9%
30%	65.7%
40%	64.2%
50%	51.4%

the performance of CCHMM2 for image retrieval is 52.1%. The performance of CCHMM-G is very promising, it is expected that we can improve it by sampling more pointers on each contour and using more features instead of 4 only in our experiments.

6 Conclusion

In this paper, we have analyzed the properties of the standard circular HMMs and have pointed out the limitation of the circular HMMs. We have proposed to impose the constraint on the initial and final states of hidden Markov models to improve the performance of HMMs. We have also developed two algorithms to implement our proposal. The proposed algorithms have been tested on the database of the MPEG-7 Core Experiments Shape-1, Part B for 2D shape recognition and shape-based image retrieval. The experimental results show that both proposed algorithms can achieve better performance on 2D shape recognition than that of the standard circular HMM in terms of accuracy. The CCHMM2 with a garbage state (CCHMM-G) has been also evaluated on shape-based image retrieval because CCHMM-G is much faster than elastic matching algorithms. The results on image retrieval are promising and that proves the CCHMM2 and CCHMM-G have much potential due to its accuracy and speed.

References

- D. S. Zhang and G. Lu, "Content-based shape retrieval using different shape descriptors: A comparative study," in *Proceedings of IEEE International Conference on Multimedia and Expo*, Tokyo, Japan, August 2001, pp. 317– 320.
- [2] E. Persoon and K. Fu, "Shape discrimination using fourier descriptors," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 8, no. 3, pp. 388–397, May 1986.
- [3] J. Cai and Z. Q. Liu, "Hidden markov models with spectral features for 2d shape recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 23, no. 12, pp. 1454–1458, December 2001.

- [4] R. L. L. J. Latecki and U. Eckhardt, "Shape descriptors for non-rigid shapes with a single closed contour," in *Proc. of IEEE Conf. on Computer Vision and Pattern Recognition*, June 2000, pp. 424–429.
- [5] J. Cai and Z. Q. Liu, "Integration of structural and statistical information for unconstrained handwritten numeral recognition," *IEEE Trans. on Pattern Analysis and Machine Intelligence*, vol. 21, no. 3, pp. 263–270, March 1999.
- [6] N. Arica and F. Yarman-Vural, "A shape descriptor based on circular hidden markov model," in *Proceedings of IEEE International Conference on Pattern Recognition*, 2000, pp. 924–927.
- [7] F. Mokhtarian and M.Bober, Curvature Scale Space Representation: Theory, Applications, and MPEG-7 Standardization, ser. Computational Imaging and Vision. Springer, 2003, vol. 25.
- [8] S. Kopf, T. Haenselmann, and W. Effelsberg, "Enhancing Curvature Scale Space Features for Robust Shape Classification," in *Proceed*ings of IEEE International Conference on Multimedia and Expo (ICME'2005), Amsterdam, The Netherlands, July 2005.
- [9] A. Kaup and J. Heuer, "Polygonal Shape Descriptors - An Efficient Solution for Image Retrieval and Object Localization," in *Pro*ceedings of 34th Asilomar Conference on Signals, Systems, and Computers, Pacific Grove , CA, October 2000, pp. 59–64.
- [10] Z. Liu, J. Cai, and R. Buse, Handwriting Recognition: Soft Computing and Probabilistic Approaches, ser. Fuzziness and Soft Computing. Springer, 2003, vol. 133.
- [11] "Shape data for the MPEG-7 core experiment ce-shape-1, PART B," http://www.cis.temple.edu/latecki/TestData/ mpeg7shapeB.tar.gz.
- [12] V. M. M. Bicego and M. Figueiredo, "Similarity-based classification of sequences using hidden markov models," *Pattern Recognition*, vol. 37, no. 12, pp. 2281–2291, May 2004.
- [13] I. Bartolini, P. Ciaccia, and M. Patella, "Warp: Accurate retrieval of shapes using phase of fourier descriptors and time warping distance," *IEEE Trans. on Pattern Analysis* and Machine Intelligence, vol. 27, pp. 142– 147, 2005.