PARAMETERIZATION METHODOLOGY FOR 2D SHAPE CLASSIFICATION BY HIDDEN MARKOV MODELS

Miguel A. Ferrer, Jesús B. Alonso, Sebastien David, Carlos M. Travieso

Departamento de Señales y Comunicaciones, Universidad de Las Palmas de Gran Canaria,
Campus de Tafira 35017 Las Palmas G.C., Spain (Europe)
Pone: +34 928 451 269, fax: +34 928 451 243, email mferrer@dsc.ulpgc.es, web: www.gpds.ulpgc.es

ABSTRACT

In computer vision, two-dimensional shape classification is a complex and well known topic, often basic for three-dimensional object recognition. Among different classification methods, this paper is focus on those that describe the 2D shape by means of a sequence of d-dimensional vectors which feeds a left to right hidden Markov model (HMM) recogniser. We propose a methodology for featuring the 2D shape with a sequence of vectors that take advantage of the HMM ability to spot the times when the infrequent vectors of the input sequence of vectors occur. This propierty is deduced by the repetition of the same HMM state during the moments in which the infrequent vectors is repeated. These HMM states are called by us synchronism states. The synchronization between the HMM and the input sequence of vectors can be improved thanks to adding an index component to the vectors. We show the recognition rate improvement of our proposal on selected applications.

1. INTRODUCTION

Object recognition, shape modelling, and shape classification constitute active research areas in computer vision. Its applications has been found in various areas such as industrial part identification, target identification, character recognition, and medical diagnosis. Moreover, these issues are receiving a growing attention due to the advent of visual databases and the related necessity to retrieve information not only by using textual queries, but also on the basis of the image context.

Shape classification requires a shape description method. Shape description refers to the methods that result in a numeric descriptor or shape descriptor vector from a given shape. The goal of the description is to uniquely characterize the shape using its shape descriptor vector. The required properties of a shape description are invariance to translation, scale and rotation. This is required because these three transformation do not change the shape of the object.

Many techniques are based on the description of two-dimensional (2D) aspects of the objects, and a wide literature can be found about 2-D shape classification and planar object recognition. Early techniques are available in Pavlidis [1] and Loncaric [2]. In this paper we will focus on boundary transform techniques which consist of transforming the 2D shape into a sequence of vectors. For example, a 2D shape can be represented by the sequence of radius (one dimensional vectors) from the shape centre of gravity to each pixel of the contour. Note that time series modelling tools can be used to describe 2D shape from its sequence transform. In conventional time series analysis, a major objective is to forecast, and the adequacy of the model fitted to a series may be judged by its forecast performance [3]. Taking into account that forecast with linear models is not useful here because it smoothes the corners of the shape, which have high information content according modern theories of visual perception [4].

Within this context, the use of hidden Markov models [5] for shape classification when the shape is described as a time series seems appropriate. Hidden Markov models represent a widespread approach to the modelling of sequences as they attempt to capture the underlying structure of a set symbol strings.

The use of HMM for shape recognition has not been widely addressed. Only a few work have been found to have some similarities with our approach. In the first, He and Kundu [6] utilize HMMs to model shape contours through autoregressive (AR) coefficients. The uses of circular HMM for shape recognition improving scaling and deformation robustness is proposed at [7]. A particular HMM topology providing rotation invariance is the Pseudo 2-D hidden Markov models (P2DHMM) proposed in [8]. No particular original solution for the HMM design are proposed, perhaps because there are not a clear knowledge about what is the meaning of a *HMM states* on shape recognition by hidden Markov models.

The goal of this work is to deepen in the meaning of a HMM *state* on shape recognition and to use this knowledge to propose an efficient parameterization method.

The rest of the paper is organized as follows. In Section 2 the HMM-based 2D shape recogniser is described. Section 3 report our proposal on the conditions to fulfill by the sequence of vectors that feature the 2D shape in order to obtain a HMM-based high discriminative recognizer. Section 4 is devoted to synthetic experiment for illustrating the technique proposed at section 3. Finally, section 5 gives the experimental results and then comes the conclusion.

2. HIDDEN MARKOV MODEL FOR 2D SHAPE RECOGNITION

2.1 Hidden Markov Models

A Hidden Markov Model (HMM) is basically a stochastic finite state automaton, formally defined by the following elements [5]: a set $S=\{S_1, \ldots S_i, \ldots, S_j, \ldots, S_N\}$ of status, a *state transition probability distribution matrix* $A=\{a_{i,j}\}_{1\leq i,j\leq N}$ representing the probability to go from state S_i to state S_j, a set $V=\{v_1, v_2, \ldots, v_k, \ldots, v_M,\}$of observation symbols where v_k use to be a *d*-dimensional vector (in the case of discrete HMM), an *observation symbol probability distribution* or *emission matrix* $B=\{b_j(v_k)\,|_{1\leq k\leq M}\}_{1\leq j\leq N}$ indicating the probability of emission of symbol v_k when system state is S_j, and an *intial state probability distribution* $\pi=\{\pi_i\}_{1\leq i\leq N}$, representing probabilities of initial states. For convenience, we denote an HMM as a triplet $\lambda=\{A,B,\pi\}$, which determine uniquely the model.

Modelling a sequence of observation symbols, as will be our case, is usual to use a so called *left to right HMM*, which has only partial state transition matrix such that $a_{i,j}=0$ $j<i$, $j>i+step$, where *step* is a constant usually equal to 1 or 2.

There are three main problems involved with HMM use, which are described next.

First problem: Given the *HMM* $\lambda = \{A, B, \pi\}$, and the observation symbol sequence $O=O_1, O_2, \ldots, O_t \ldots O_T$ with $O_t \in V$, we want to compute $P(O\,|\,\lambda)$, i.e., the probability that the observation sequence O is generated by the model λ. This is usually solved using the so called *forward-backward procedure*.

Second problem: Given the model $\lambda = \{A, B, \pi\}$ and a observation sequence $O=O_1, \ldots, O_T$, we want to determine the status sequence $Q = \{q_1, q_2, \ldots, q_t \ldots, q_T\}$ $q_t \in S$ such that $P(Q\,|\,O,\lambda)$ is maximum. In other words, we want to compute the most probably state sequence. There are several possible optimality criteria. For example, one possible optimality criterion is to choose the states q_t which are individually most likely. To implement this solution, is defined the variable gamma $\gamma_t(i) = P(q_t = S_i|O,\lambda)$ which represents the probability of being in state S_i at time t. Another criterion is to find the single best state sequence (path), based on dynamic programming methods, and is called the Viterbi Algorithm.

Third problem: Given a set of L observation symbol sequences $\{O_l\}_l=\{O_{1,l}, O_{2,l}, \ldots, O_{t,l}, \ldots, O_{T(l),l}\}_{1\leq l\leq L}$ with $O_{t,l} \in V$, we want to determine $\lambda = \{A, B, \pi\}$ such that $P(\{O_l\}_l\,|\,\lambda)$ is maximized: this is the problem of training an HMM. The best-known method to perform this operation is the so called Baum-Welch reestimation technique. It is an iterative procedure based on Expectation-Maximization (EM) algorithm, and it tries to maximize the loglikelihood $P(O\,|\,\lambda)$ of the model with respect to the data.

2.2 HMM-based 2D shape recogniser

Consider using HMMs to build a 2D shape recogniser. Assume we have a set of C classes of 2D shapes and that each class is to be modelled by a distinct HMM. Further assume that for each class c, $1\leq c\leq C$, we have a training set of L_c samples. The value of L_c must be sufficient great in order to know the intraclass and interclass variability. Each 2D shape sample is transformed to a sequence of vectors that describe the 2D shape. The sequence of vectors is mapped to observation symbol sequence by means of a vector quantiser (VQ) [5]. The VQ has as many codevectors as possible symbols, in fact the codevectors are the symbols, which are obtained by means of applying the K–means algorithm to the $\sum_{c=1}^{C} L_c$ sequences of vector. So, an infrequent vector of sequence of vectors that describe a 2D shape originate an infrequent symbol of observation symbol sequence.

In order to do 2D shape recognition, we must perform the following: 1) we must build an HMM for each class c of 2D shapes, i.e., we must estimate the model parameters $\lambda_c=\{A_c,B_c,\pi_c\}$ that optimise the likelihood of the training set observation symbol sequences for class c, and 2) for each unknown 2D shape sample which is to be recognized, we calculate its observation symbol sequence O and it will be assigned to the class that maximizes $P(O\,|\,\lambda_c)$, $c=1,2,\ldots,C$.

3. PARAMETERIZATION METHODOLOGY

Since an efficient parameterization method has a strong dependence of the classifier chosen, it is crucial to deepen in the operation that carries out the HMM proposed.

Suppose a observation string $O=O_1, O_2, \ldots, O_{t-1}, O_t \ldots O_T$ with $O_t \in V$ and that we are in the instant $t-1$ in the state S_i. The probability of remain in the state S_i at the next instant t depends on both the state transition probabilities $\{a_{ij}\}_{1\leq j\leq N}$ and the observation symbol probabilities $\{b_j(O_t)\}_{1\leq j\leq N}$. If $b_i(O_t)>>b_j(O_t)$ $1\leq j\leq N$ with $j\neq i$, the HMM will tend to remain in the same state despite of a low state transition probability a_{ii}. On the other hand, if $b_i(O_t)\approx b_j(O_t)$ the possibility of transition will be governed by A matrix.

That is to say, the observation sequence will remain in a state S_i while a symbol with high probability in the state S_i and low probability in the others is repeated in the sequence. On the other hand, it will jump out when a symbol of equal probability in various states occurs. The first case indicates that such a symbol is an infrequent symbol and that it usually happens about the instant t whereas the second case indicates that the symbol is very frequent and occurs in different t. It seems reasonable to think that the first sort of symbols is the one that endows the discriminative power of the HMM.

Therefore, we can talk about a speed propagation way of the sequence through the HMM visible by means of the *gamma* matrix. When the observable symbol sequence O spreads following the main propagation way, the HMM emission probability $P(O\,|\,\lambda)$ will be high. In this case we can say that the HMM is synchronized with the input sequence. By synchronism we understand that the input sequence arrives in the state S_i at the instant t in which one of its infrequent symbols is very probable. This involves that the sequence remains in such a state as long as the symbol observed does not change. When the symbol

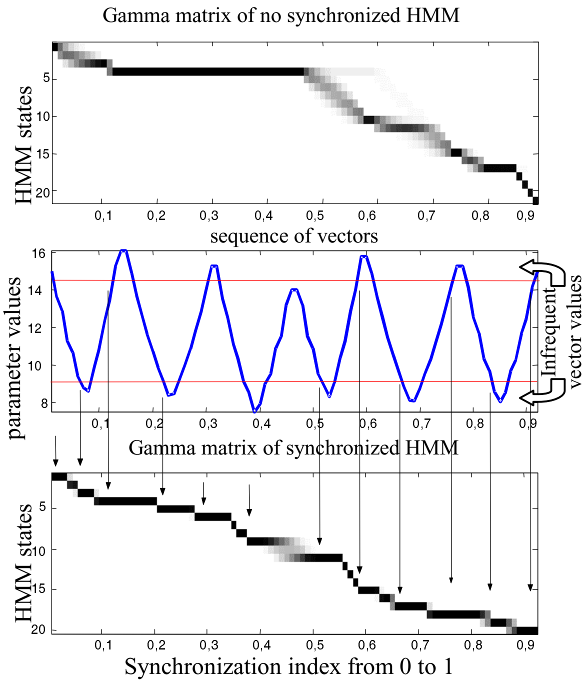

Figure 1. effect of synchronization on gamma matrix.

changes, the HMM will search again the stability trying to match another synchronism state. Hence, in the case of synchronism, the *gamma* matrix presents a figure with stair appearance. In other case, the *gamma* matrix presents a diffused structure as can be seen at fig. 1.

Therefore, if we look for high discriminative HMM-based 2D shape recognizer, each HMM $\lambda_c=\{A_c,B_c,\pi_c\}$ would have to have either different infrequent symbols $v_k \in V$ or be given at different states. This condition can be verified analysing B matrix. If the HMM do not match this condition, a new 2D shape description would be advisable.

To improve the HMM ability of being synchronized with the observation symbol sequence, we propose to add a synchronism information to the sequence of vectors. This can be provided, for instance, by means of increasing the dimension of each vector of the sequence of vectors with a index. An example, with Matlab code, of adding a synchronization index to sequence of vectors is the next:

```
% seq is a matrix whose rows contain the vectors
synchronized_seq = [seq linspace(0,1,size(seq,1))'];
```

Using this synchronization index, we have obtained improvements as much with 2D shape and with speech. Although this last application, and all those in which the vector of the sequence of vectors has a great dimension, require to emphasize the importance of the synchronization index when the likelihoods are claculated.

4 SYNTHETIC EXPERIMENT

To verify the previous proposal, we propose the next experiment: 4 classes of objects with different contours are created. The class 1 objects are 5-sided almost regular polygons with vertexes of variable position. The vertexes are located inside a ball of radio r centered in the vertex of the original regular polygon. The position of the vertexes

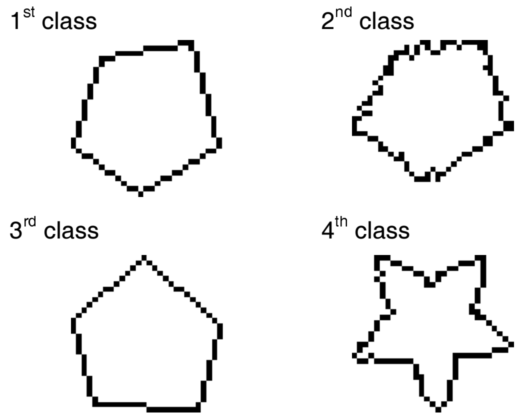

Figure 2. Each class example of proposed experiment

presents a uniform distribution inside the ball. The class 2 are almost regular polygons as those of the class 1 with vertexes connected by lines with Gaussian noise added. The class 3 is made of polygons as those of the class 1 rotated such angle that the positions of their vertexes do not coincide with those of the class 1. Finally, the class 4 include stars of 5 tips with variable position of its external and internal vertexes. An example of each class is represented in the figure 2. The size of the images used in our experiments is of 30x30 pixels, and the radius ball 4.

The sequence of vectors that represents the 2D shape is the radius from the center of gravity to each pixel of the contour. So, the infrequent vectors are those radius from the center to contour corners. The average length of signal sequence is 85 components. The initial point is the pixel with angle 0. The initial point does not provide rotation invariance but it is considered appropriate for this paper purposes.

Following the above statements, the class 1 and 2 should be the only one to be confused by a synchronized HMM because they are the singles that share the same infrequent vectors at the same instants while the remainder should be well discriminated among themselves.

To train the discrete HMM recognizer, 2000 samples of each class were generated, 1000 of them were used to train and the 1000 remainders for test. The HMMs have 32 symbols by state, and 20 states were chosen. The four HMM have been trained with the iterative algorithm of Baum-Welch. The initialization method is the equal occupation of states [9] and the stop criterion is the fixed threshold equal to 0.001. The maximum number of iterations is fixed to 30. The above-mentioned experiment was repeated with the synchronized sequence of vectors. The confusion matrix obtained at both experiments is showed at table I and confirms our above proposal.

TABLE I. CONFUSSION MATRIX OF HMM RECOGNIZER FOR SYNTHETIC EXPERIMENT

Without synchronization index						With synchronization index					
		1	2	3	4			1	2	3	4
classes	1	940	60	0	0	classes	1	780	220	0	0
	2	100	900	0	0		2	160	840	0	0
	3	0	80	920	0		3	0	0	1000	0
	4	0	0	0	1000		4	0	0	0	1000
	Recognition rate: 94%						Recognition rate: 90.5%				

Synchronized sequences of vectors reduces the recognition rate because the synthetic experiment presents low variability and because the synchronization technique increases the confusion among the 1st and 2nd class. In real applications with higher variability the synchronization technique increase significantly the recognition rate (see Table II and III).

5. EXPERIMENTS

In order to verify the previous affirmations, experiments with handwritten numerals and letter will be carried out. Highlight that the objective is not to propose a digits recognizer in the state of the art but to verify the HMM way of operation described.

5.1 Application to manuscript digits

For hand printed digits, NIST special database 19 has been used. It consists of 409467 manuscript digits. Partitions *hsf_0* to *hsf_3* (229639 digits) were used for the training, and two tests were done: the first with *hsf_4* (58645 digits) partition, and the second with *hsf_6* and *hsf_7* partitions (121183 digits). The envelopes of above-mentioned digits are extracted and the sequence of vectors is built with the sequence of radius from the mass center of the envelope up to each outline pixel. As it can be easily verified through the probability density function of vectors that compose the sequence of vectors, the radius is not a good 2D shape descriptor for HMM-based recognized because the infrequent radius values are the same for the most of the classes and the great variability of the instants when occurs make almost impossible the synchronization task, despite the fact of using the synchronization component.

In order to avoid the drawback of infrequent values coincidence we have applied the difference operator to the sequence of vectors. This transformation changes (of different way per class) the infrequent vectors and their instants increasing the discriminative HMM ability. The HMM are a left-to-right model of 40 states and 64 symbols per state trained as described in the synthetic case.

Table II shows the recognition rate of the three experiments done: first working with radius as vectors, second working with polar coordinates as vector (radiuses plus angle as synchronism index), and third working with the radiuses differentiate plus angles as synchronism index.

TABLE II HMM SYNCHRONOUS RATE WITH DIGITS NIST DB19

radius sequence	*hsf_4*	*hsf_6y7*
no synchronized	23.89%	25.91%
Synchronized	67.97%	77.64%
differentiate & synchronized	81.38%	89.14%

5.2 Results with handprinted letter

For hand printed letters, NIST special database 19 has been used. It consists of 220304 manuscript upper and 190998 lower case manuscript letters. Partitions *hsf_0* to *hsf_3* (177792 upper and 155215 lower case manuscript letters) were used for the training. We have done two test, first with, *hsf_4* partition, and second with *hsf_6* and *hsf_7* partitions. The HMMs used are discrete, with a codebook of 64 code vectors and 40 states. The left to right HMMs have been trained as described above. Results showed in table III confirm above statements.

TABLE III HMM SYNCHRONOUS RATE WITH LETTERS NIST DB19

	upper		lower	
radius sequence	*hsf_4*	*hsf_6y7*	*hsf_4*	*hsf_6y7*
no synchronized	8.52%	8.77%	11.46%	12.50%
synchronized	62.65%	66.35%	61.45%	64.37%
differentiate and syn.	71.97%	73.85%	67.45%	70.55%

In order to obtain the same effect with continuous HMM, it is necessary to initialize them carefully. In fact, to obtain results similar to those of Table II and III, we have initiated the continuous HMM with the discrete HMM.

6. CONCLUSION

For high discriminative HMM based recognizer applied to 2D shape recognition problem, the vectors of the sequence of vectors that describe the 2D shape should be chosen such that the infrequent vector or their occurring instant can be distinguished. Looking for 2D shape descriptors that match above condition, and ways to improve the synchronism between the input sequence of vectors and the HMM are new research line proposed by this paper.

Acknowledges: This work has been funded by Spanish government TIC2000-0586.

REFERENCES

[1] T.Pavlidis, "Algorithms for shape analysis of contour and waveforms", in *IEEE transactions on pattern analysis and machine intelligence*, vol. 2, no. 8, pp. 301-312, July 1980.

[2] Loncaric S., "A survey of shape analysis techniques", in *EURASIP Pattern Recognition*, vol.31, no.8, pp. 983-1001, 1998

[3] Kartikeyan B., Sarkar A., "Shape description by time series", *in IEEE transactions on pattern analisys and machine intelligence*, vol. 11, no. 9, pp. 977-984, September 1989

[4] Attneave, F., Arnoult, M. D., "The quantitative study of shape and pattern recognition",in *Psychological Bulletin*, vol.53, pp. 452-471, 1956. Also in *Pattern Recognition*, Editor L. Uhr, pp. 123-141, John Wiley, 1966.

[5] Rabiner L.R., "A tutorial on hidden Markov models and selected applications in speech recognition", in *Proceedings of IEEE*, vol. 77, no.2, pp. 257-286, 1989.

[6] He Y., Kundu A., "2-D shape classification using hidden Markov model", in *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, no. 11, pp. 1172-1184., November 1991

[7] Arica N., Vural F.T.Y., "A shape descriptor based on circular hidden harkov model",in *IEEE 15th international conference on pattern recognition*, vol. 1, pp. 924-927, 2000.

[8] Kuo S., Agazzi O.E., "Keyword spotting in poorly printed documents using pseudo 2-D hidden Markov models", in *IEEE Transactions on pattern analysis and machine intelligence*, vol. 16, no. 8, pp. 842-848, August 1994.

[9] Ferrer M.A., Alonso I., Travieso C.M., " Influence of initialization and stop criteria on HMM based recognizers", in *IEE Electronic Letters*, vol 36, no. 13, pp. 1165-1166, June 2000.