Introduction of fuzzy logic in the Hidden Markov Models

M. Kelarestaghi
LI/E3I – Université de Tours 64, Av. Jean Portalis 37200 Tours, France
kelarestaghi@univ-tours.fr

M. Slimane
LI/E3I – Université de Tours 64, Av. Jean Portalis 37200 Tours, France
slimane@univ-tours.fr

N. Vincent
LI/E3I – Université de Tours 64, Av. Jean Portalis 37200 Tours, France
vincent@univ-tours.fr

Abstract

In this paper, we present a fuzzy logic integration to the Hidden Markov Models (HMM). We have replaced the basic arithmetic operators by some adequate fuzzy operators. Using fuzzy operators permits us to relax the additivity constraint of probability measures. So we will need only the monotonicity with respect to set inclusion. Some results show the interest of our approach.

Keywords: Hidden Markov Model, Hybridising, Fuzzy logic, Fuzzy operators, Image processing, Learning, Pattern recognition, Face recognition, Signal Processing.

1 Introduction

The HMMs have shown their interest in many fields such as: Image Analysis [9], Speech recognition [7], Handwritten recognition [4] or Signal processing [8]. The properties of HMM are based on a strong mathematic frame issued from the probability theory that imposes the additivity constraint. In our study this constraint will be replaced by monotonicity that is much less restrictive. For that we have chosen to replace traditional calculations of the HMM by fuzzy operators. In addition to the relaxation they bring us other advantages such as the possibility to treat the imprecise fuzzy data and the attenuation of the rescaling problem that often occurs in computing HMM, so we can solve it more easily.

After recalling the HMM basis, we show the use of fuzzy operators makes it possible to process fuzzy data and to manage uncertainty, which is inherent in the HMM usage. We present the fuzzification of two algorithms. At last we compare the obtained results by our algorithm with those obtained on the one hand by Mohamed's method [6] and on the other hand with those obtained by a net approach.

2 Hidden Markov Models

Conventionally, we note by the triplet (A,B, Π) an HMM λ. A and B are the stochastic matrices and we have: $\sum_{i=1}^{n} \pi_i = 1$ .Here O, O = (Ot) t represents an observation of length T.

We distinguish three types of problems while using the HMM :

- Evaluation of an HMM. It is estimated by calculating the $P(O|\lambda)$ (Forward algorithm [7])
- The optimal sequence research (Viterbi algorithm [10])
- Learning the best HMM associated with a given learning set (Baum-welch algorithm [1]).

For example the three steps of the Forward algorithm can be stated by :

1. Initialisation :

$$\alpha_1(i) \leftarrow \pi_i * b_i(O1) \quad \forall i \in 1..n$$

2. Induction :

$$\alpha_{t+1}(j) = \Sigma_{i=1 \text{ to } n} (\alpha_t(i) * a_{ij}) * b_j(O_{t+1}) \quad \forall t \in 1..T\text{-}1 , \forall j \in 1..n \quad (1)$$

3. Termination :

$$P(O|\lambda) \leftarrow \Sigma_{i=1 \text{ to } n} \alpha_T (i)$$

In fact the method that we develop here has been generalized also to the other algorithms of HMM in order to obtain the results presented in the last section.

3 Fuzzy Approach

3.1 Fuzzy Operators

The fuzzy subset theory was presented by Zadeh [11] to treat imprecise and uncertain data. It defines the membership function with respect to the fuzzy subsets to measure the likeness between the data and these subsets. The fuzzy data are treated by means of fuzzy operators. These operators are the basis of the generalisation of the fuzzy logic in different domains such as control [5], approximate reasoning [2], data fusion [3]. Generally we can design a fuzzy system, based on a non-fuzzy one by generalising the calculus to the fuzzy data or also, based on a completely fuzzy way of reasoning. In the next paragraph we present the generalisation of a HMM to the fuzzy case.

3.2 Application to the Forward algorithm

A fuzzy modelling can be introduced at different levels, particularly at the data level, at reasoning mode level or directly using a fuzzy Markov model itself whose elements would be fuzzy. For the data level integration, it is possible to apply the classical extension principles of the net functions to fuzzy sets. We didn't choose this solution. It seems wiser to search a fuzzy version of the Hidden Markov Models (fHMM) and to use the fuzzy operators that lead to the possibility to use all kinds of data. So the elements of the matrices A, B, Π are considered as the fuzzy subsets. We limit ourselves to the subsets with triangular belonging functions, in fact, their base width can be seen as an uncertainty measure. We have transformed the classical formulas used in the different algorithms, and for example the formula (1) becomes formula (2):

$$\alpha_{t+1}(j) = \text{MIN} \{\text{MAX} \{ \text{MIN} \{\alpha_t(i) , a_{ij} \}\} , b_j (O_{t+1})\} \quad \forall t \in 1..T\text{-}1 , \forall j \in 1..n \quad (2)$$

The MIN and the MAX operators are not the usual ones. They do not even are extensions of the classical ones. They are defined on fuzzy sets as following. If two fuzzy subsets have any overlap we will use the normal operators, but for the subsets without overlapping parts the normal operators will always give the results equal to zero. To be able to afford this problem we will use the mean of the two subsets. In one hand we solve our problem and on the other hand we could avoid the rescaling problem specially for the MIN operator.
The MIN and MAX operators are defined in such a way that, in each calculus step, we find a triangular subset to model α . It is a very important point since it permits us to avoid the rescaling phase. In fact the fuzzy information is indicated in a function the support of which does not become too wide. Other aggregation or MEAN operators have been tried in place of product. All figures must be centered like Figure 1. The number and caption of the figure must always appear below the figure.

4 Results

Experiments were performed on a face image data base(400 images). We have compared the results of HMM algorithm and fHMM algorithm with also the results obtained from Mohamed's [5] algorithm. Our image data base contains the images of 100 persons each one in 4 different positions. For each person a learning is performed by using three positions and we want to recognize this person in his fourth image. The obtained results show a good performance for fHMM specially for the cases in which the noise level of the image to be recognized is important. Figure 1 gives a brief comparison between three methods: firstly the learning using the Mohamed and Gader method and secondly the classical method using the normal HMM and thirdly our method using the fuzzified HMM (fHMM).

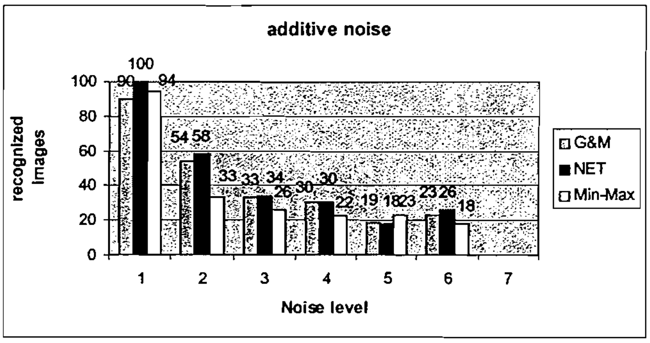

Figure1: comparative results.

In this figure we study the recognition rate of different methods for the noisy signals. We observe that for all the methods the recognition rate decreases when we increase the injected noise in signal. But at the same time we observe a good behavior of our method, for the signals with a big noise injection degree.

5 Conclusion and perspectives

We have presented the integration of fuzzy logic to the Hidden Markov Models. This integration consists in replacing the basic arithmetic operations by the adequate fuzzy operators. Using the fuzzy operators permits us in one hand to manage the imprecision concerning the data, and in the other hand it relaxes the additivity constraint, necessary for the HMM, towards the monotonicity one, much less restrictive. The new algorithms could be applied directly on fuzzy data. The performed tests for noisy images recognition, showed encouraging results particularly for the cases with an important noise. After us it should be interesting to try to model the HMM by the fuzzy rules and so we could translate the HMM to a more adequate fHMM.

References

[1] L. E. Baum, J. A. Eagon, (1967)"An inequality with application to statistical estimation for probabilistic functions of Markov processes and to a model of ecology", Bull American Society, Vol, 73, pp. 360-363, 1967.

[2] D. Dubois, H. Prade, (1991)" Fuzzy sets in approximate reasoning Part 1 : Inference with possibility distributions," Fuzzy Sets and Systems 40, 143-202, 1991.

[3] C. Dujet, N. Vincent, (1998)« Data fusion Modelling human behaviour,» International Journal of intelligent systems, vol. 13, 27-39, 1998.

[4] M. Gilloux, (1997)"Real-time handwritten word recognition within large lexicon", in Progress in Handwriting Recognition, A.C. Downton and S. Impedovo Eds, World Scientific, pp. 45-50, 1997.

[5] E. H. Mamdani, S. Assilian,(1975) "An experiment in linguistic synthesis with a fuzzy logic controller," International Journal of Man-Machine studies 7, 1-13, 1975.

[6] M. A. Mohamed, P. Gader, (2000)"Generalized hidden markov models-part I: Theoretical frameworks," IEEE Transaction on fuzzy systems, Vol. 8, No. 1, pp. 67-81, 2000.

[7] L. R. Rabiner, (1989)"A tutorial on hidden Markov models and selected application in speech recognition", Pro. of IEEE, vo77, p. 257-286,1989.

[8] M, Saerens, "Hidden Markov models assuming a continuous time dynamic emission of acoustic vectors", Proceedings of Eurospeech ,1993.

[9] M. Slimane, G. Venturini, J.-P. Asselin De Beauville, T. Brouard, A. Brandeau, (1996)«Optimizing HMM with a genetic algorithm,», Artificial Evolution, Lecture notes in computer science, Springer Verlag, Vol. 1063, pp ; 384-396, 1996.

[10] A. J. Viterbi, (1967)"Error bounds for convolutional codes and asymptotically optimum encoding algorithms", IEEE Transactions on information Theory, Vol ; 13, pp. 260-269, 1967.

[11] L. A. Zadeh, (1965)"Fuzzy sets," Information and Control 8, 338-353, 1965.