HIDDEN MARKOV MODELS

Summary

Introduction

Hidden Markov Model (HMM) is a very common technique in recognition and classification, especially in the case of sequential data processes such as speech, music, text. The simplicity, the well-known algorithms and the proven efficiency of this approach in many research problems led to numerous HMM-based researches in music information retrieval.

Markov Processes

Markov process is a model of memoryless random process. It is used in many domains such as finance, game theory or telecommunications where it can be used to predict future values of a share or to size a phone network. The process reach a state (e.g. number of lines occupied in a phone network) at each time step (e.g. beginning/end of a phone call), and the probability distribution of the future states that could be reached depends only on present state of the system.

A Markov process is defined by two parameters. The first one is the initial distribution vector π that gives the probability of each state to be the first one. The second one is the transition probability distribution that quantifies the probability to reach a certain state conditionally to the present one. In the case of processes with discrete time and state-space, this last parameter is modelled as the transition matrix P, where P_{ij} gives the probability to reach the state j while coming from the state i.

Hidden Markov Models

HMMs are an extension of Markov processes. In this case, rather than knowing the actual state sequence of the system, we only have observations related to the states (e.g. temperature for weather previsions). These observations are bound statistically to the states by the confusion matrix that gives the probability of one state of the system to have generated a particular observation.

There are three common problem classes in HMM approach. The first one is the evaluation: given a sequence of observations, the goal is to compute the likeliness of this sequence to be produced by a given HMM. It is a way to compare HMMs to find the one that best fits the observations. The second one is the decoding: given a sequence of observations, the goal is to compute the more likely sequence of states that could have generated it in the HMM. Thus, we “recognize” the original information of the underlying Markov process. The third one is the learning: given sequences of observations, the goal is to find the optimal HMM where these sequences are the more likely to be produced. These three problems are related to efficient algorithms since they were optimized by using dynamic programming. The two algorithms widely used in recognition are the Viterbi algorithm (decoding) and the Forward-backward algorithm (learning).

HMMs were derived to increase their performance in some applications. These extensions are based on using more complex observation mechanisms. Examples of these are Layered HMM, Hierarchical HMM or Probability Distribution HMM.

The basic recognition scheme used in research is the following. First, we define parameters of the HMM adapted to the recognition type. Then, the HMM probability distributions (confusion and transition matrices) are optimized by solving the learning problem. At last, the Viterbi algorithm solved the decoding problem linked with the recognition. The first step is the most important since the two others are quite standard. Indeed, the characteristics of the HMM is strongly related to the process efficiency, but the importance of the different HMM parameters has to be determined by experiments.

Hidden Markov Model Applications in Music Information Retrieval

Many applications of HMMs in music information retrieval were experimented in recent research. Each of them implements the recognition scheme exposed before.

Classification of European folk music by geographic origin was experimented (Chai and Vercoe 2001). The idea is to detect structure similarity between pieces from the observation of pitch sequences. Different HMM characteristics are compared in the article. The experiment results legitimate this approach as a way to draw a characterization of similarities between same origin pieces.

Another field of research where HMM are used is chord and key recognition (Burgoyne and Saul 2005; Sheh and Ellis 2003). These articles are identifying probability distribution characteristics of features from different analyses of the input audio signal used as observations. The model is used on the limited problem of detecting key/chord segment knowing the key/chord sequence of the piece or on the problem of full recognition.

HMM were also used for score following problems (Orio and Déchelle 2001). The presented model use a more complex state-space description with two levels to deal with note contour (lower level) and score abstraction (higher level). The goal is to achieve score following with an input signal despite errors made by the performer and degradation of signal quality.

The problem of melody spotting is also an opportunity of using this approach (Durey and Clements 2001). HMMs are used to recognize a piece from a melody query of a user (e.g. humming). The paper compared different observation spaces and deal with the question of error measuring.

At last, HMM-based method was studied for optical music recognition of early music prints (Pugin 2006). It showed actually that this approach could reach very satisfying recognition rates without staff removal by using an adapted description for the observation of features.

All of these articles are showing a good efficiency of HMM approach. Indeed, the sequential aspect of music makes it fit with this kind of model. The main problem is the observation space description since pitch detection in real audio streams could be quite inaccurate, especially in polyphonic and multi-timbral music. Another interesting issue raised in the articles is error rate measurement since errors produce sometimes a false but plausible response that can be used in chord detection or melody spotting.

Other Applications

HMMs are also widely used in different domains. Speech and phone recognitions were the first applications developed in research. It has been extended to recognition in language or literature, and more widely in domains where sequential events occurred. Classification models based on HMM are using the idea that sequence correlation could lead to a characterization of the different classes. This approach is used in music or literature classification. HMMs are also used in the biomedical domain, on problems related to DNA and proteins. Other punctual applications are signal processing or fault-tolerance modelling.

Conclusion

HMM is a powerful model for recognition and classification in the domain of sequential events. Researchers have tested quite successfully this approach in common problems of music information retrieval. The results of current experiments are promising but further work has to be done to improve HMM efficiency, particularly on representation of music and signal features.

References

Burgoyne, J. A., and L. Saul. 2005. Learning harmonic relationships in digital audio with Dirichlet-based hidden Markov models. *Proceedings of the International Symposium on Music Information Retrieval*:438-443.

Chai, W., and B. Vercoe. 2001. Folk music classification using hidden Markov models. *Proceedings of the International Conference on Artificial Intelligence*.

Durey, Adriane S., and M. Clements. 2001. Melody spotting using hidden Markov models. *Proceedings of the International Symposium on Music Information Retrieval*.

Lee, K., and H. Hon. 1989. Speaker-independent phone recognition using hidden Markov models. *IEEE Transactions on Acoustics. Speech, and Signal Processing* 37 (11):1641-8.

Orio, N., and F. Déchelle. 2001. Score following using spectral analysis and hidden Markov models. *Proceedings of the International Computer Music Conference*:125-9.

Pugin, L. 2006. Optical music recognition of early typographic prints using hidden Markov models. *Proceedings of the International Symposium on Music Information Retrieval*:53-6.

Rabiner, L. R. 1989. A tutorial on hidden Markov-models and selected applications in speech recognition. *Proceedings of the IEEE* 77 (2):257-286.

Sheh, A., and D. P. W. Ellis. 2003. Chord segmentation and recognition using EM-trained hidden Markov models. *Proceedings of the International Symposium on Music Information Retrieval*.

Links

Roger Boyle, "Hidden Markov Models", University of Leeds,
http://sitesweb/www.comp.leeds.ac.uk/roger/HiddenMarkovModels/html_dev/main.html

Warakagoda, Narada D., "A Hybrid ANN-HMM ASR system with NN based adaptive preprocessing", Norges Tekniske Høgskole,
http://sitesweb/jedlik.phy.bme.hu/~gerjanos/HMM/hoved.html