Data mining techniques for time series processing

Application to industrial sound classification

Javier Romero[1], Amalia Luque[1*], Alejandro Carrasco[1]

[1]Escuela Politécnica Superior, Universidad de Sevilla, Spain
javier@romeroyromero.es,{amalialuque,acarrasco}@us.es

Abstract. Sound classification is a realm where signal processing techniques are usually deployed. But a sound is, in fact, a particular case of a time series or, even more general, sequential data. On the other hand, data mining techniques have been widely faced to large volumes of data, offering a large set of tools for classifying purposes. In this thesis, an approach to use data mining techniques to sound classification is explored. Furthermore, industrial concerns (noise immunity, standardization, processing time, hardware requirements) are taken into account. The results will be applied to anuran species identification from their sounds.

Keywords: sound classification · data mining· time series processing

1 Thesis data

(a) Author: Javier Romero Lemos
(b) Thesis title: Data mining techniques for time series processing. Application to industrial sound classification.
(Técnicas de minería de datos en el proceso de secuencias temporales. Aplicaciones a la clasificación industrial de sonidos)
(c) Supervisors: Amalia Luque, Alejandro Carrasco
(d) Department (University): Informatics Doctoral Program, University of Seville
(e) Starting date: May, 2015

2 Summary

In many industrial projects, as well as in extensive areas of science and technology, we must cope with large volumes of data having a time series structure. Their processing, in many cases, is accomplished by problem oriented techniques only appropriate for a narrow type of applications.

On the other hand, the increasing storage capacity and processing power is leading to a scenario where large volumes of information have to be considered. This situation

adfa, p. 1, 2011.

has motivated the appearance of techniques that, under the generic name of data mining, can address many problems using statistical and artificial intelligence approaches.

The aim of the thesis is to study the applicability of these data mining techniques to the case where data have a time series structure, which entails significant constraints. Their results will be employed to the problem of industrial sound classification.

The thesis' conclusions will be applied to automatic identification of anuran species. Anuran sexual calls are highly influenced by temperature. So the existence of some species in a certain geographical area becomes a good indicator of climate change. For this purpose, biologists are recording huge amounts of animal sounds which have to be classified.

The industrial features of the classification process is indicated by its noise and disturbance immunity, standardized sound description, real-time processing capabilities, and low power and low cost concerns.

3 Methodology and action plan

The thesis will be developed following a predetermined outline. Research work and writing labors will be jointly executed. In the next paragraphs thesis' content is detailed, explaining the proposed methodology and the action plan.

1. INTRODUCTION

This chapter describes the problem to be solved, set thesis goals, and defines the document structure. It includes a description of the basic biological problem (detection of climate change through analysis of anuran populations) and the characteristics of the sounds to be processed. It should also include a specification of the "industrial" processing features: noise and disturbance immunity, standardized sound description, real-time processing capabilities, and low power and low cost concerns. The chapter will be organized according to the following structure:

(a) Problem description
(b) Thesis goals
(c) Thesis outline

2. CLASICAL APPROACHES IN AUDIO SIGNAL PROCESSING

Sound processing classical techniques are described in this section, combining time domain and frequency domain audio signal processing. Also sound homomorphic (cepstral) description is explored, introducing the MFCC (Mel Frequency Cepstral Coefficients) and its application to sound classification. Finally, it describes Linear Predictive Coding (LPC) techniques and its relevance for audio signal description. Throughout this section several anuran sounds are considered. The chapter primary

intention is to describe the state of the art in audio signal processing classical techniques. The chapter will be organized according to the following structure:

(a) Time domain processing
(b) Frequency domain processing
(c) Homomorphic processing
(d) Linear Predictive Coding

3. MPEG-7 SOUND DESCRIPTION

Sound parametric description is commonly required for classification. However the state of the art regarding sound descriptors shows a great heterogeneity, thus using standard features would greatly enhance the industrial requirement of classification processes. To cope with heterogeneity problem, MPEG-7 standard is used through the thesis. In this chapter the MPEG-7 standard scope and objectives is described, and how it could be applied to sounds description. The meaning of every standard parameter and the algorithms for their computation are also described. Finally the HMM (Hidden Markov Models) are presented as they are the sound classification technique suggested by the MPEG-7 standard. Throughout this section several anuran sounds are described using the MPEG-7 parameters. The chapter will be organized according to the following structure:

(a) MPEG-7 standard scope and objectives
(b) MPEG-7 parameters definition and computation
(c) MPEG-7 sound classification algorithm

4. CLASSIFICATION TECHNIQUES

Data mining classification techniques are considered and described along this section. Previously extracted MPEG-7 features are now exploited for sound recognition purposes, applying these data mining methods. For the sake of simplicity, the sequential characteristic of sounds are not considered in this chapter, regarding the MPEG-7 parameters vectors as a non-sequential set of points to be classified. Through the chapter, the following algorithms and their application to anuran sounds will be considered:

(a) Minimum distance
(b) Maximum likelihood
(c) Decision trees
(d) k-nearest neighbor
(e) SVM
(f) Logistic regression
(g) Neural networks
(h) Discriminant function
(i) Bayes classifiers

5. TIME SERIES CLASSIFICATION

In previous sections, sounds are represented as a non-structured set of points. Every point is described as an MPEG-7 n-length vector in an $\mathbb{R}^n$ space. However, introducing timing information could enhance classifier's performance. So it should be better considering sounds as vector time series, that is, as a sequence of MPEG-7 n-length vectors. This section will carry out the task of describing data mining techniques for time series classification and their application to anuran sounds. The chapter will be organized according to the following structure:

(a) Sequence oriented parameters
(b) Sliding windows
(c) Recursive sliding windows
(d) Hidden Markov Models
(e) Time series

6. DERIVATE TIME SERIES

Applying the algorithms previously described, usually emerge derivative (secondary) time series which contain the classifying decision for every sound frame. This piece of information could be a simple class name (or number) but usually carries on a more detailed data (a decision confidence assessment). Categorizing not original but derivative time series, and their application to anuran sounds, is explored through this section with the goal of better classification performances. Also implementation concerns are considered in order to ensure the industrial accomplishments in sound classification. The chapter will be organized according to the following structure:

(a) Counting-based classification
(b) Similarity-based classification
(c) Parametric classification
(d) Industrial implementation issues

7. CONCLUSIONS

This section will summarize the main tasks developed through the development of the thesis besides emphasizing its main results. The principal thesis contributions should be clearly described besides the advances in scientific and technological knowledge. Potential practical applications would also be highlighted.

8. REFERENCES

A complete list of the references used through the thesis will be provided in this section. Some of the most relevant are the following:

(a) Problem description

1. Llusia, D., Márquez, R., Beltrán, J. F., Benitez, M., & Do Amaral, J. P. (2013). Calling behaviour under climate change: geographical and seasonal variation of calling temperatures in ectotherms. *Global change biology*, *19*(9), 2655-2674.
2. Aide, T. M., Corrada-Bravo, C., Campos-Cerqueira, M., Milan, C., Vega, G., & Alvarez, R. (2013). Real-time bioacoustics monitoring and automated species identification. *PeerJ*, *1*, e103.

(b) Sound processing classical techniques

3. Fulop, S. (2011). *Speech spectrum analysis*. Springer Science & Business Media.

(c) MPEG-7 sound description

4. ISO (2001). *ISO 15938-4:2001 (MPEG-7: Multimedia Content Description Interface), Part 4: Audio*. ISO
5. Kim, H. G., & Sikora, T. (2004, June). How efficient is MPEG-7 for general sound recognition?. In *Audio Engineering Society Conference: 25th International Conference: Metadata for Audio*. Audio Engineering Society.

(d) Data mining classification techniques

6. Härdle, W. K., & Simar, L. (2012). *Applied multivariate statistical analysis*. Springer Science & Business Media.
7. Hastie, T., Tibshirani, R., & Friedman, J. (2005). The elements of statistical learning: data mining, inference and prediction. Springer-Verlag.

(e) Time series classification

8. Esling, P., & Agon, C. (2012). Time-series data mining. *ACM Computing Surveys (CSUR)*, *45*(1), 12.
9. Fu, T. C. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24(1), 164-181.
10. Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257-286.
11. Box, G. E., Jenkins, G. M., & Reinsel, G. C. (2011). *Time series analysis: forecasting and control* (Vol. 734). John Wiley & Sons.

4 Relevance

This thesis would make significant contributions to scientific knowledge and be fully applicable to real problems. The expected main results are the following:

- MPEG-7 parameters properly describe sounds for classification purposes.
- Non-sequential sounds description is adequate for a good classification result.
- Data mining techniques are proven to be useful for sounds classification purposes.
- Sequential sound description significantly enhances classification performances.
- Derivative time series processing improves overall classification results.
- Industrial automatic classification is a feasible task for real anuran audio databases.