SMS IDENTIFICATION USING PPM, PSYCHOPHYSIOLOGICAL CONCEPTS, AND MELODIC AND RYHMIC ELEMENTS

First author
School
Department

Second author
Company
Address

ABSTRACT

The aim of this work is to propose a method for MIREX 2012 in SYMBOLIC MELODIC SIMILARITY task. The method apply the Prediction by Partial Matching (PPM), and also propose data modeling inspired on psychophysiological aspects. Rhythmic and melodic elements are combined, instead of using only melody or rhythm alone. The models consider the perception of pitch changing and note durations articulations, and them the models are used to classify melodies.

1. INTRODUCTION

The analysis of information by computers has been increasingly improved since the establishment of Information Theory. Retrieving information from music using concepts based on this theory is a challenge faced by researchers that have to consider aspects of both music representation and musical information content. However, thinking about music and some ways the brain obtains musical information might help to develop algorithms to perform better musical analysis. When listening to a melody, the physiology of auditory system discriminates each pitch sequentially and then sends a signal with this information to the brain until the stimulus comes to an end [5]. This means that each musical note is listened to and treated by the auditive cortex in the context of its predecessor [8]. In this work we combine notes and its predecessors with concepts of Information Theory and then propose a new way of analysis based on sound perception to create a classifier.

The development of Information Theory allowed data processing in different forms that are relevant to many areas of science. Among them, Music Information Retrieval is benefited by some algorithms brought by this theory. The state of art of data compression known from Information Theory is the Prediction by Partial Matching (PPM) method [13], and here we correlate the similarities between this method and the perception of musical pitches by the brain. The PPM method acts similarly the auditory perception considering that the main computation is based on a sequence of symbols, and the context of each symbol (its neighbours) is used to calculate the probability of its occurrence.

A similar point of view is described on related works that associate concepts from different areas. Dowling [3,4] discussed the advantages of psychological and physiological concepts approach, especially in melody memorization by intervals or scale contours. Similar concepts have also been used by Londei et al. [10] for the computation of the difference of pitches between sequential notes. Suyoto et al. [14] considered the *Longest Commom Subsequence* (LCS) algorithm to find similarities among musical pieces, showing that these sequences definitely are useful for information retrieving in the musical context.

Although there are several works with significant results for musical analysis through melodies, our main goal is to propose a new way of using the sequential analysis with pattern recognition. Here, we propose to consider melodic and rhythmic structures together and apply the PPM method on data modeled using psychophysiological aspects. Our method of analysis is based on pitch intervals and note duration ratios retrieved from symbolic data representation that permits an easy evaluation of each of these aspects in a sequence of notes interpreted from MIDI files. The MIDI format importance arises from its wide use and its firm position as an industry standard, so you can expect a lot of databases and softwares using this format of codification.

The use of the PPM method on Music Information Retrieval (MIR) is an old recommendation for this area. Lartilot et al. [9] made an evaluation of different statistical models for the modeling of musical style and suggest the use of PPM as a possible statistical technique to be considered. Afterwards, Pearce et al. [11] tested some techniques for PPM application on monophonic music, evaluating advantages of N-gram Markov models in MIR. Hazan et al. [6] evaluated PPM on audio signals, and showed another range of PPM use by pitch and onset analysis.

2. MUSICAL PATTERNS ANALYSIS

Recognize patterns by organizing and interpreting sensory information is a natural way of learning new things. Considering that every individual develops sensory perception before meaningful perception, one can say that the aspects of perception may aid the interpretation of human habits. Keeping this in mind, the study of music perception together with some peculiarities from Gestalt laws has lots of subjects to be investigated [7].

Symbolic data formats include pitches and note duration that can represent the melody, the rhythmic elements of a melody, and also other elements with relevant information for pattern recognition. The perception of those elements have its analysis and foundations based on areas such as psychology and physiology. It also gives some cues for analysis of musical patterns considering note intervals and durations in the same way as they are perceived [15].

2.1 Using PPM with symbolic data

The PPM method is an advanced compression method. It was developed by J. Cleary and I. Witten, and is based on an encoder that calculates and maintains a statistical model of the context symbols. When the encoder reads the next symbol from input, a probability is assigned to the new symbol. This statistical model looks at the number of times each symbol was encountered in the past in a given context, and assigns a conditional probability based on that, thus creating a context-based tree. The context-based tree works with the context of each item, and so depending on the selected length for the context, the model is created based on the number of times each symbol was followed by each context [13].

There are some similarities between our application of PPM and the process of sound perception by the humans. PPM context-based tree is created during a sequence interpretation. The notes from a melody are also perceived sequentially by cochlea cells. The perception is based on pitch articulation (variation between the cell stimulated) and pitch duration (time spent during a cell stimulation). The primary auditory cortex is tonotopically organized, the secondary process melodic and rhythmic patterns, and tertiary integrates everything. In the same way, we read a MIDI file sequentially and interpret the melody from pitch articulation and duration. Those informations are sent to PPM to create the context-based tree and PPM acts as the secondary auditory cortex. We also try to simulate the tertiary auditory cortex by combining melodic and rhythmic elements in different ways.

The PPM method has different versions, but the most similar with human perception is the PPM-C, as this version consider a fixed context and use a escape to characterize something new. Fixed context represents the way humans group melodic patterns from short sequences, and the escape can be compared with the surprise of a new pattern perceived. PPM-C version has been also suggested on Lartillot et al. [9] and Pearce et al. [11] researches about music information retrieval. At new work, Pearce et al. [12] also shows improvements applying PPM* instead of PPM-C used on previous work, but this version of PPM consider undefined context lenght, and for that reason the PPM* will not be analyzed here.

2.2 Melodic similarity analysis in symbolic data

Aiming symbolic melodic similarity identification, we analyze melody similarity with MIDI files matching the most frequent sequences of pitch intervals or note duration ratios between files. As an example we can see a matching on two parts of different violin scores composed by J. S. Bach. In Figure 1, pitch intervals are represented as the difference between the MIDI representation of each note. The rhythmic elements of the melody derived from note duration is also shown on the same figure. These approaches find similarities, even though they have different note and beat representations.

On the example in Figure 1, the melodic sequence analyzed at BWV 1002 is comprised by the notes B4, G4, E4, D5, C5#, B4, A4#, and at BWV 1004 is comprised by D4, A3#, G3, F4, E4, D4, C4#. The notes are not the same and neither is their note contexts, but in terms of intervals counted by the semitone (e.g., E5, F5 = +1 and D5, C5 = -2) the description of both melodies is the same (-4, -3, +10, -1, -2, -1). For these notes, the relative durations at BWV 1002 are $\frac{3}{32}$, $\frac{1}{32}$, $\frac{3}{32}$, $\frac{1}{32}$, $\frac{3}{32}$, $\frac{1}{32}$, $\frac{3}{32}$, and at BWV 1004 are $\frac{3}{16}$, $\frac{1}{16}$, $\frac{3}{16}$, $\frac{1}{16}$, $\frac{3}{16}$, $\frac{1}{16}$, $\frac{3}{16}$, but both duration ratios are $\frac{1}{3}$, 3, $\frac{1}{3}$, 3, $\frac{1}{3}$, 3 . It is noticeable that some sequences can match through this method, and than we can analyze the patterns occurrences.

3. EVALUATION METHOD DESCRIPTION

The main idea of our evaluation process consisted on grouping some melodic sequences to create models with the PPM method. The MIDI files used to create the models were normalized to ensure the same number of symbols avoiding disproportionate models. The normalization is done before the models creation. The purpose of this normalization is to guarantee that all models have the same quantity of intervals. The main consequences of this normalization might happen when some melodic patterns happens only in the end of a sequence, and they will be ignored here.

Some considerations with the alphabet had to been taken into account for the interpretation of each MIDI file . The PPM method uses symbols from an alphabet, and at this work the type of symbol chosen was one unsigned byte, so each interval has to be represented as a number between 0 and 255. Based on the fact that MIDI note numbers have a range between 0 and 127, the pitch intervals from melodic sequences can generate numbers from -127 to 127, so there is no problem to convert it to a byte range by summing each result with 127.

On the other hand, duration ratios from rhythmic sequences were represented using a special description due to the ratios being normally represented by floating numbers or fractions. In this case, floating numbers equal or greater than 1 were truncated after the first decimal digit and multiplied by 10. Floating numbers smaller than 1 were truncated after the second decimal digit, multiplied by 100, and then summed with 127. Below is shown an example of these truncations used to generate symbols for the PPM from Figure 1 results:

$$3.0 \Rightarrow 3 * 10 = 30$$

$$\frac{1}{3} = 0.3333.. \Rightarrow 0.33 * 100 + 127 = 160$$

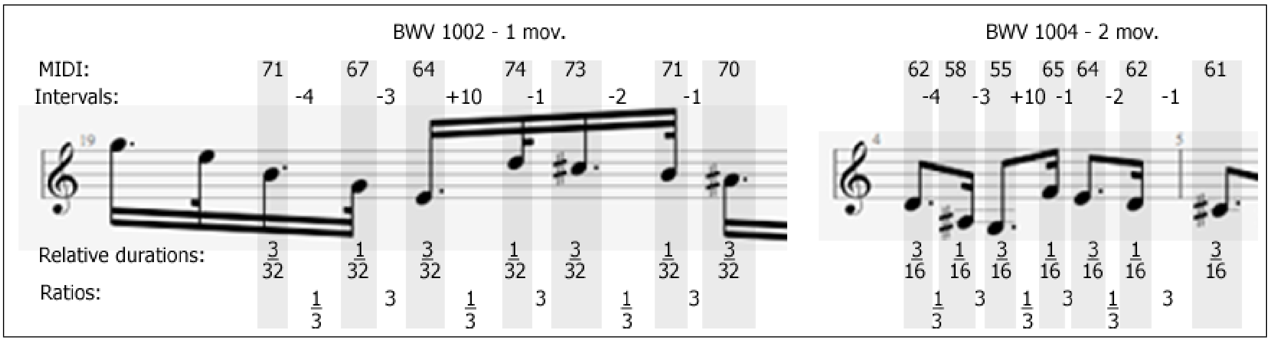

Figure 1. Representation of melodic similarities on melodies.

With this alphabet other considerations, we now describe the proposed methods to classify melodies using PPM. On our evaluation methods, the strategy are meant to use pitch intervals and note duration ratios, grouping them afterward.

For *i* models and a file F, we now define mathematically the evaluation method proposed At description, we used "melody model" and "rhythm model" to characterize when a PPM model is created using pitch information or note duration ratios respectively.

PPM-DJ Calculation of the mean of the ranking position of each model for a file classification on both melody and rhythm model:

$$PPM\text{-}DJ(F) = argmax[mean_i(\#m_i(F), \#r_i(F)), \{M_1, ..., M_i\}]$$

$m_i(F)$ is the function that compress the file F on melody model M_i using PPM and returns the compression rate. $r_i(F)$ is the function that compress the file F on rhythm model M_i using PPM and returns the compression rate. $\#m_i(F)$ and $\#r_i(F)$ give us the ranking position of model M_i during the classification of file F on the melody model and rhythm model, respectively, using PPM.

PPM-DJ tests the ranking position of the composer on both melody an rhythm models. While two or more files can compress the same file with compression rates quite close by, the ranking position will reconsider this closeness by one unit of difference.

There are also some details about the PPM model that need consideration. The PPM method uses context lengths to determine the maximum length of the partial sequences that will be examined. If the context length selected is 3, there will be considered contexts with lengths 0 (no context at all), 1, 2, and 3. In musical terms, the context length represents how many preceding notes are considered during the evaluation of each note.

Thinking about pitch intervals, if we consider a context length 3 then we will evaluate 5 notes and 4 kinds of intervals. For example, for the simple melodic sequence C, D, E, F, G the determined pitch intervals will be 2, 2, 1, and 2. Depending on the PPM method version, it is possible to consider arbitrary contexts. In this work, we are using the PPM-C version because it was the version with the best results in an specific work from Pearce et al. [11] through an analysis with monophonic music. Here we considered contexts with length 1 aiming to simulate a recognition based on motifs and short sequences of notes.

4. CONCLUSIONS

Methods used for musical pattern recognition with symbolic data have been improved each year, getting better results with many different algorithms and methods. Here we proposed to use a method not only by its acknowledgment of being useful for pattern recognition, but also having in mind the similarity between this method and other concepts from psychology and physiology regarding human music perception. Although several research has also been carried in this area, specially the symbolic domain [1,2,11] , our improvements are based on some considerations of the methods discussed in the literature with some modifications related to multiple view points [2].

In this work, we used pitch intervals and durations inspired by the perception of melody sounds by auditory cortex. The union of melody model and rhythm model in a single validation, as we did on PPM-DJ, may improve the results significantly. Although we know that are some good results by analyzing melody model isolated,melodic and rhythmic elements combination may show that there are some relationship between them.

5. REFERENCES

[1] Emilios Cambouropoulos, Tim Crawford, and Costas S. Iliopoulos. Pattern processing in melodic sequences: Challenges, caveats and prospects. In *Proceedings of the AISB'99 Convention (Artificial Intelligence and Simulation of Behaviour*, pages 42–47, 1999.

[2] Darrell Conklin Department, Darrell Conklin, and Ian H. Witten. Multiple viewpoint systems for music prediction. *Journal of New Music Research*, 24:51–73, 1995.

[3] W Jay Dowling. Scale and contour: Two components of a theory of memory for melodies. *Psychological Review*, 1978.

[4] W Jay Dowling. Context effects on melody recognition: Scale-step versus interval representations. *Music Perception*, 3(3):281– 296, 1986.

[5] Arthur C. Guyton and John E. Hall. *Textbook of Medical Physiology*. Saunders, 12 edition, 2010.

[6] Amaury Hazan, Ricard Marxer, Paul Brossier, Hendrik Purwins, Perfecto Herrera, and Xavier Serra. What/when causal expectation modelling applied to audio signals. *Connect. Sci*, 21(2-3):119–143, June 2009.

[7] Mari Riess Jones, Richard R Fay, Arthur N Popper, and Bruno L Giordano. Music perception (springer handbook of auditory research). *Journal of the Acoustical Society of America*, 123(2):935–945, 2011.

[8] Robert Jourdain. *Music, The Brain, And Ecstasy: How Music Captures Our Imagination*. Harper Perennial, 1998.

[9] O. Lartillot, S. Dubnov, G. Assayag, and G. Bejerano. Automatic modelling of musical style. In *Proceedings of the 2001 International Computer Music Conference*, pages 447–454. San Francisco: ICMA, 2001.

[10] A. Londei, V. Loreto, and M. O. Belardinelli. Musical style and authorship categorization by informative compressors. In *Proceedings of the 5th Triennial ESCOM Conference*, pages 200–203, 2003.

[11] Marcus Pearce and Geraint Wiggins. An empirical comparison of the performance of ppm variants on a prediction task with monophonic music. *Artificial Intelligence and Creativity in Arts and Science Symposium*, 2003.

[12] Marcus T Pearce and Geraint A Wiggins. Evaluating cognitive models of musical composition. In A Cardoso and G AEditors Wiggins, editors, *Proceedings of the 4th International Joint Workshop on Computational Creativity*, pages 73–80. Goldsmiths, University of London, 2007.

[13] David Salomon and Giovanni Motta. *Handbook of Data Compression (5. ed.)*. Springer, 2010.

[14] Iman S. H. Suyoto, Alexandra L. Uitdenbogerd, and Falk Scholer. Effective retrieval of polyphonic audio with polyphonic symbolic queries. In *Proceedings of the 9th ACM SIGMM International Workshop on Multimedia Information Retrieval*, pages 105–114, 2007.

[15] Thomas J.I. Tighe and Jay W. Dowling. *Psychology and Music: the understanding of melody and rhythm*. Lawrence Erlbaum Associates, New Jersey, Hillsdale, 1993.