Knowledge-Based Systems 71 (2014) 419–434

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Automatic melody composition based on a probabilistic model of music style and harmonic rules

Carles Roig, Lorenzo J. Tardón *, Isabel Barbancho, Ana M. Barbancho

Universidad de Málaga, ATIC Research Group, Andalucía Tech, ETSI Telecomunicación, Campus de Teatinos s/n, E29071 Málaga, Spain

ARTICLE INFO

Article history:
Received 22 March 2014
Received in revised form 10 August 2014
Accepted 22 August 2014
Available online 28 August 2014

Keywords:
Automatic melody composition
Music style
Music analysis
Probabilistic model
Machine learning

ABSTRACT

The aim of the present work is to perform a step towards the design of specific algorithms and methods for automatic music generation. A novel probabilistic model for the characterization of music learned from music samples is designed. This model makes use of automatically extracted music parameters, namely tempo, time signature, rhythmic patterns and pitch contours, to characterize music. Specifically, learned rhythmic patterns and pitch contours are employed to characterize music styles. Then, a novel autonomous music compositor that generates new melodies using the model developed will be presented. The methods proposed in this paper take into consideration different aspects related to the traditional way in which music is composed by humans such as harmony evolution and structure repetitions and apply them together with the probabilistic reutilization of rhythm patterns and pitch contours learned beforehand to compose music pieces.

1. Introduction

The main objective in the field of music information retrieval is to provide algorithms and methods for the analysis and description of musical pieces to computational systems [1]. The advances achieved allow computers to perform different tasks like automatic music transcription [2], or unattended genre classification, among others [3]. Furthermore, the computational model of the human experience in the musical field and the way in which humans process this information are topics of great interest for psychology and musicology [4].

In this context, the automatic generation of musical content is the topic considered in this paper. Often, music is defined as 'organized sound' [5], this organization requires order, structure and a logical composition style learned by training [6]. Thus, a novel algorithm designed for learning composition rules and patterns will be shown in this paper. Specifically, three different musical aspects will be considered: time (tempo and time signature estimation [1]), rhythm (rhythmic pattern learning) and intonation (musical contour detection).

Different methods for music style classification have already been described [7,8]. However, unlike cited works, in our case, the classification is not performed to train a classifier but to build a data model upon the selected descriptors of a particular music style. It is not the characterization of music styles what is ultimately pursued in this work but the utilization of the descriptions selected for the unattended generation of new melodies according to those styles.

In other words, in this work, style-based composition parameters will be learned from the low level features for style classification. Thus, a main novelty that will be presented will be based on the analysis and post-processing of low level musical features to determine the style of the music that will be generated. The procedure developed improves the performance and adds a different level of complexity with respect to other approaches. An ad hoc database will be built from the melodic and rhythmic patterns found for different musical styles. Using this information, the proposed composer system will be able to generate new melodies according to the characterized style [9] in a way that will be similar to the one followed by a human composer.

As mentioned above, a set of descriptors must be analysed in order to model the style of the melodies. Concerning the temporal descriptors, namely tempo and time signature, previous related works can be found. The work presented by Uhle and Herre in [1] is focused on onset estimation by means of spectral analysis and a subsequent study to estimate the length of the bar. On the other hand, in [10], histograms are employed to find the mostly repeated inter onset value in order to estimate the tempo without performing a spectral analysis. Also, there are recent methods for structural analysis, as the one described in [11], based on a time-span tree for the detection of structural similarity; this task can be addressed by making use of the auto-similarity matrix [12]. This

* Corresponding author.
E-mail addresses: carles@ic.uma.es (C. Roig), lorenzo@ic.uma.es (L.J. Tardón), ibp@ic.uma.es (I. Barbancho), abp@ic.uma.es (A.M. Barbancho).

http://dx.doi.org/10.1016/j.knosys.2014.08.018
0950-7051/

method can be employed for both the extraction of music patterns and tempo estimation.

The low level temporal estimation process can be related to certain stages of previous MIDI-to-score works like the one described in [13], which is focused on the complex study of the musical object in the MIDI representation through the analysis of the accent and beat induction. Unlike this work, the goal of the temporal estimation scheme that will be presented in this manuscript is the characterization of simple rhythmic patterns (melodies without rhythmic ornaments).

The scheme designed in this work proposes an innovative approach for tempo estimation based on the inter onset interval (IOI) histogram. Our approach is inspired by [1,10]. However, some improvements are proposed in order to attain better accuracy and lower computational cost. Specifically, the proposed system uses a multi-resolution method that reduces the number of cost function evaluations and achieves better accuracy.

Regarding musical composition based on pattern reallocation and variations, which is the key point of the present work, other methods are described in the bibliography. In [14–16], Markov models are used for the modeling and composition process. The use of genetic algorithms, such as in Biles' GenJam system [17], has also been considered in the automatic music composition context. Biles' system and ours are based on learning musical sequences (rhythm and pitch) plus a mutation. However the system we propose does not require a fitness evaluation stage since every learned pattern can be adapted to fit the composition. Furthermore, GenJam implements pattern mutation capability trough genetic algorithms which increases the flexibility and variability of the content. However, this feature cannot be applied in our approach since our scheme pursues style-replication.

Methods based on probabilistic approaches such as Cope's Experiments in Musical Intelligence (EMI) [18,19] also focus on the creation of an automatic music composition framework. Both EMI framework and the proposed system are based on probabilistic models (for pattern selection) and the training data stored in the database comes from musical features extracted from MIDI files (rhythm, pitch, dynamics,...). However EMI performs beat-to-beat pattern matching while our approach makes use of complete measures. Although both schemes use actual motives, the use of longer patterns, as in our approach, provides a clearer style replication. Moreover, unlike Cope's approach, that generates non-tonal compositions, the proposed system defines a harmonic and rhythmic structure to compose tonal music.

Another music generation method based on probabilistic structures is Inmamusys [20]. This method follows a global scheme similar to the proposed system. Both schemes use previously learned patterns for generating new compositions by replicating and reallocating them. However, the learning stages and features are different, the composition rules are distinct and, also, our system performs a post-processing process to allow all motives in the database fit in a composition. The solution to this problem adopted in Inmamusys is to restrict the combinations to a subset of motives that had been previously tagged as compatible.

Fractals, fuzzy logic and experts systems have also been used for the composition of melodies [21]. However, unlike the proposed system, the melodies generated by the scheme described in [21] are not related to any learned style and the composition scheme is radically different.

Summing up, the main contribution of the composition stage of the proposed system is the reduction of the importance of fitness when reallocating musical motives. Biles solves this issue by using genetic algorithms that force the mutation of the patterns to fit [17]. In our proposal, the reallocation of randomly selected patterns plus a global melody contour arrangement stage allow all the elements to fit together. The global music arrangement stage, which is crucial in our work, is based on music theory rules.

Additionally, a set of automatic processes for the extraction of musical features from MIDI files has been designed to learn the musical patterns required by the composition scheme. Thus, we have designed a key-independent, unattended and style-based pattern learning system plus a composition process that replicate the style of the training data assuming that the musical style is related to rhythmic patterns and melodic contours. Although harmony and chord progressions are also closely related to music style definition, the chord analysis was postponed for further development and it is placed out of the scope of this manuscript.

This paper is organized in five sections. After the introduction, in Section 2, a description of the analysis module will be presented. In Section 3, the automatic composer will be explained and a complete specification of all its subsystems will be given. In Section 4, the tests done to evaluate the system and the results obtained will be drawn. Finally, the last section will present the conclusions coming out from this work.

2. Pattern analyser

The selected approach for the design of the generation scheme is based on the music theory method called *obstinato* [9], as described in Sections 1 and 3. This method considers the composition of music on the basis of the repetition of patterns so that the repetitions themselves constitute the melodic structure. In order to learn the patterns that will be used to compose new melodies, it is necessary to analyse the available samples to identify and extract the required elements for the composition, namely: rhythmic patterns, pitch contours, harmonic progressions and tempo information (needed for the proper analysis of the data). Similar developments are proposed in [22].

Thus, a database of musical parameters is designed to model musical styles (as in [23]) and organize and store the data that will be used for the composition of new melodies. Since the main objective is to develop a valid music model for the automatic creation of composition with style replication, it is necessary to discover which parameters can be used for the proper modeling of musical styles. We decided to base our approach on the probabilistic analysis of musical elements such as rhythm, pitch and harmony. Fig. 1 presents a scheme of the analysis system.

Tempo can be extracted easily by analysing MIDI meta-data messages [24]. However, this piece of information can be missing (undefined) or incorrect. In order to develop a generally usable robust scheme, an algorithm to estimate the tempo and time signature has been designed. Note that tempo information is critical in order to correctly identify rhythmic figures.

Rhythm, pitch motives (melody) and harmony (which is considered at the generation stage) are main pieces of information necessary for music generation [25]. This information will be coded and stored in the database.

The database will be divided into three levels hierarchically organized containing (1) time signatures, (2) rhythm patterns, and (3) pitch contours (see Fig. 2). The symbolic notation used in this database to store the duration of the rhythmic figures corresponds to the ratio between the lengths of the notes and the length of a whole note. In this way, figure identification is tempo independent and the complexity is reduced. So, the whole notes will be denoted as 1, the half notes as 0.5, and so on. Recall that the specific duration (in s) of the figures is important only when the waveform is created. Melodic contours will be converted to the equivalent MIDI note number [24].

As an example, the rhythm instance associated with a 2/4 measure with C–D–E notes being a quarter note and two eighth