MUSICAL STYLE AND AUTHORSHIP CATEGORIZATION BY INFORMATIVE COMPRESSORS

Alessandro Londei[1], Vittorio Loreto[2], Marta Olivetti Belardinelli[3,1]

[1] ECONA: Interuniversity Center for Research in Cognitive Processing in Natural and Artificial Systems, Italy

[2] Department of Physics, University of Rome"La Sapienza", Italy

[3] Department of Psychology, "La Sapienza" University of Rome, Italy

ABSTRACT

Musical style and authorship detection is a cognitive human ability, depending on the degree of musical acculturation, experience and emotional sensibility of the listener.

This work is aimed to verify how informative entropy of a musical sequence may represent a formal descriptive parameter for musical categorization by evaluating the differential entropy distribution over the usual stylistic classes as artistic currents, historical periods and musicians' style.

Recently, a novel parameter based on the compressibility of an informative sequence was introduced. The best compression rate of a data sequence is related to the self-similarity of the sequence and then to its complexity. Typical compression algorithms were applied to MIDI files of musical pieces written by several authors belonging to different stylistic currents. Besides, the same technique was applied to a specific class of musical stimuli used in psychological analysis of perception of tonality and salience. Finally, a distribution analysis of the compression parameters was performed in order to reveal a stylistic clusterization of the musical pieces.

We observed a clusterization of compression parameters around some representative values, by assuming that musical pieces belonging to a specific author or artistic current reveal a similar amount of complexity. Clusters revealed a structured distribution depending on the similarity of artistic currents, providing a natural artistic distance among musical styles. Furthermore, musical pieces not included in the previous cluster formation were correctly categorized and attributed to the related author in most of cases.

Formal description provided by entropy analysis may represent a powerful tool for the determination of a uniform class of musical stimuli useful in psychological cognitive tasks. Moreover, it may be related to the ability of a human listener to extract musical descriptive components.

1. INTRODUCTION

Music can be considered as a non-verbal language equipped with a stylistic and cultural-depending grammar. Every musical piece written by a particular author reflects, at a low level, the established composition techniques determined by the specific historical or cultural environment [1]. At a higher level, the atmosphere a musical piece may give rise to a human listener, like a sort of semantic level, can be related to the stylistic current which the author belonged to.

Although a trained listener or a musician is able to classify an unknown music into a well-defined musical current while determining eventual influences from specific authors and characterizing the piece in terms of specific parameters as tonality, rhythm or ethnic influence, an artificial algorithmic method to perform the same task has not been yet identified. Even some simple concepts as tonality represent a hard identification problem for an automatic approach.

Anyway, since music may be considered as an evolving temporal linguistic signal, we tried to extract an invariant parameter related to the complexity of the musical piece. According to the information theory, complexity of a temporal flow of data emitted by a source can be identified by a particular parameter called entropy. Such a parameter expresses the amount of surprise the source emitting the symbol sequence can reserve to us. Of course this parameter gives a description of the amount of redundancy of a symbolic sequence. In music the concept of redundancy is straightforward connected to rhythmic homogeneity or to self-similarity between different parts of the musical piece, as thematic recurrence and counterpoint. An important parameter characterizing intervallic and rhythmic redundancy in music is represented by salience. As shown in paragraph 4, salience plays a crucial role in detection of homogeneous musical pieces.

In a recent paper [2], a new method for analyzing and characterizing textual sequences based on entropy analysis has been introduced. By this technique it is possible to assign a specific piece of literature to the correct author by evaluating the relative entropy between the actual text and a class of assigned reference textual pieces. The evaluation of the relative entropy is performed by means of the popular compressing algorithms used for zipping files.

In order to face the problem of stylistic and authorship assignment, we extracted the relative entropy from a set of musical pieces originally coded in MIDI format and then converted them into a sequence of textual characters by preserving the pitch, the duration and the temporal position of each note. As it will be shown in the next sections, this method gives interesting results at the scope of assigning of the correct authorship and stylistic current of music.

ISBN 3-931852-67-9
ISSN 1617-6847

2. THEORY

Compressing an information sequence means to find an optimal coding for translating the information in a way that the final coded sequence results shorter than the original one. As Shannon shown [3], there exists a limitation to the maximal compressibility of information, which is given by entropy. An interesting definition of entropy was given by Chaitin-Kolmogorov: the entropy of a string of characters is the length (in bits) of the smallest program which produces as output the string. Nevertheless there are algorithms explicitly conceived to approach this theoretical limit like file compressors or zippers. The operation performed by this kind of algorithms is based on the identification of equal sequences of characters contained in the file to compress. The more a particular sequence is found into the file data, the most such sequence can be coded by a shorter coding. Files containing highly random data cannot be efficiently compressed while redundant sequences, as natural languages, show a greater level of compressibility.

The compression algorithms provide a powerful tool for the measure of the entropy. In particular, in this paper we exploit this kind of algorithms to measure the remoteness between pairs of musical sequences. The notion of relative entropy can be easily grasped by the following example. Let us consider two sources *A* and *B* emitting sequences of 0 and 1. *A* emits a 0 with probability *p* and 1 with probability (1-*p*) while *B* emits a 0 with probability *q* and 1 with probability (1-*q*). The compression algorithm applied to the sequence emitted by *A* will be able to encode the sequence almost optimally: i.e. coding a 0 with $-\log_2 p$ bits and a 1 with $-\log_2(1-p)$ bits. These values correspond to the amount of information brought by each character. Anyway this optimal coding will not be the optimal one for the sequence emitted by *B*. In particular the entropy per character of the sequence emitted by *B* in the coding optimal for *A* will be $-q\log_2 p-(1-q)\log_2(1-p)$ while the entropy per character of the sequence emitted by *B* in its optimal codig is $-q\log_2 q-(1-q)\log_2(1-q)$. The number of bits per character wasted to encode the sequence emitted by *B* with the coding optimal for *A* is the relative entropy of *A* and *B*, $S_{AB}= -q\log_2(p/q)-(1-q)\log_2(1-p)/(1-q)$.

In order to define the relative entropy between two sources *A* and *B* we extract a long sequence *A* from the source *A* and a long sequence *B* as well as a small sequence *b* from the source *B*. We create a new sequence *A*+*b* simply appending *b* after *A*. The sequence *A*+*b* is now zipped and the measure of the length of *b* in the coding optimized for *A* will be $\Delta_{Ab}=L_{A+b}-L_A$ where L_X is the length of the zipped file *X*. The relative entropy S_{AB} will be estimated by

$$S_{AB}=(\Delta_{Ab}-\Delta_{Bb})/|b| \qquad (1)$$

where $|b|$ is the number of characters of the sequence *b* and $\Delta_{Bb}/|b|=(L_{B+b}-L_B)/|b|$ is an estimate of the entropy of the source *B*. Translated in a linguistic framework, if *A* and *B* are texts written in different languages, Δ_{Ab} is a measure of the difficulty for a generic person of mother tongue *A* to understand the text written in the language *B*. The results of experiments on textual files turn out to be very robust with respect to large variation of the size of the file *b* (typically 1-15 Kbytes for a typical size of file *A* of 32-64 Kbytes).

Since $S_{AB}\neq S_{BA}$ it is not possible to consider the relative entropy as a distance measure. A suitable parameter, based on relative entropy, which can describe the distance between two data sequences, may be defined as:

$$D_{AB}=(\Delta_{Ab}-\Delta_{Bb})/\Delta_{Bb}+(\Delta_{Ba}-\Delta_{Aa})/\Delta_{Aa} \qquad (2)$$

Now D_{AB} is positive, symmetric and fulfills the triangular inequality.

3. MUSIC DATA ANALYSIS

The application of relative entropy analysis to musical sequences requires a suitable encoding of data in order to represent the musical piece as a sequence of characters. The final text encoding has to preserve the pitch, the duration and the temporal position of every note.

3.1. From MIDI to text

MIDI (Musical Instrument Digital Interface) data is a very efficient method of representing musical performance information. Musical information stored into a MIDI file is composed of note events and dynamic events distributed on several tracks. For our purpose, we realized a program that extracts information about note pitch, duration and temporal position by neglecting the remaining dynamic and structural information. Then data were organized into a text file as follows. Each row of the text file represents a temporal slot of fixed duration (0.1 sec). The first note in a row is coded by an ASCII character whose value is given by the difference between the note and the previous one plus 127. Hence each note is independent from the musical key (differential encoding). Since a piano keyboard is composed of 88 keys, the value 255 was used to describe the lack of notes (pause). A note whose temporal length is greater than a single slot will be encoded by a single character when it starts to play followed by several 127's.

3.2. From text to Relative Entropy

After the MIDI-to-text encoding was performed, we dealt with a certain amount of textual files, each one describing a particular musical piece under analysis. Then, we applied the algorithm shown in paragraph 2 to every couple of files obtaining as an output a list of relative entropies. Besides we calculated the distance described in (2) in order to characterize the musical pieces into a topographic space.

3.3. From Relative Entropy to Points

After the previous calculations, we described the musical pieces as points whose mutual distances are given by D_{AB} (2). Since such a distance measure fulfills the triangular inequality, it is possible to find a spatial representation of the points into a Euclidian space having a sufficiently large dimensionality. For *N* points, the chosen dimensionality was *N* in order to assure a correct spatial embedding. This can be accomplished by a neural iterative algorithm that at every step minimizes the difference between the calculated distance between points and the distances given by

ISBN 3-931852-67-9
ISSN 1617-6847

(2). For normalized distances we accepted point representations whose error was lesser than 10^{-3}.

3.4. From Points to Topographic Space

The problem which arises when one deals with high-dimensional data is the difficulty to represent the data on a 2-dimensional surface. In our case, we need to give a suitable representation in which the local relations between points are preserved. A powerful tool to approach this problem is represented by Kohonen's Self-Organizing Map [4]. An interesting feature of such neural typology is the ability to give a 2D visualization of high-dimensional patterns by preserving their statistical distributions.

After the translation of entropic distances to multidimensional points was done, we applied the transformed data to Kohonen's network and analyzed the distribution of points over the neural surface to evaluate eventual clusterizations of points belonging to the same class of music.

4. EXPERIMENTS

At present, the application of analysis of relative entropy was conducted on two particular classes of musical examples. In the first experiment, 17 pieces from XIX and XX century were analyzed. The purpose was to assign every piece to its correct class (determined by authors) and to evaluate the relations between couples of classes (i.e. closeness of stylistic homogeneous classes). In the second experiment, 48 musical pieces, characterized by 2 descriptive musical parameters, were analyzed. This experiment is very important for psychological purposes because the same stimuli are used for several experimental designs on human listeners [5].

4.1. Experiment 1

The musical pieces used in the first experiment were:

- The 2nd and 3rd movement from "Mondscheinsonate" op. 27 for piano by L. van Beethoven.
- The 2nd, 3rd and 4th movement from "Sonate 7" op. 10 for piano by L. van Beethoven.
- Movements 1, 4, 5 and 6 from "Dolly" op. 56 for piano by G. Fauré.
- 4 Ragtime for piano by S. Joplin: "Elite Syncopations", "Maple Leaf Rag", "Frog Legs Rag" and "Quality Rag".
- 4 extracts from "The Rite of Spring" (reduction for piano) by I. Stravinsky.

In order to assign every piece to a specific author, we sorted the relative entropies associated to every couple of files. The minimal relative entropy describes the best musical piece to be associated to the piece under analysis in terms of optimal encoding. In Table 1 it is shown the percentage of correct assignment for each class of authors.

Author	Correct assignments (%)
L. van Beethoven	100
G. Fauré	25
S. Joplin	75
I. Stravinsky	75

Table 1: Correct assignments (%) from Relative Entropy calculated on 17 musical pieces chosen from classical repertoire.

The low efficiency in assignment of Fauré works may be explained by the stylistic similarity of "Dolly" with the other modern authors' pieces. Beethoven pieces, that are stylistically more easily identifiable among all groups, give the main difference between the analyzed classes. By training a Kohonen network on the set of 17-dimensional points extracted from entropy distances (2), the distributions of pieces over the 2-dimensional neural surface (for supervised learning) are shown if Figure 1.

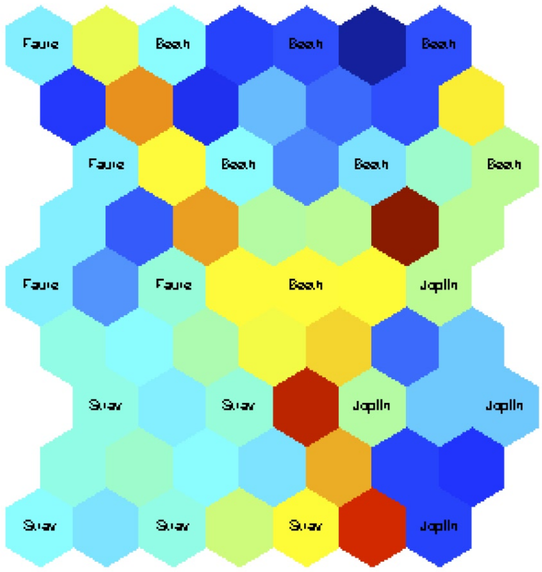

Figure 1: Distribution of musical pieces over a 2-dimensional neural surface of a 5-by-4 Kohonen map (supervised learning). The works of the same author are clusterized together and differences between stylistic currents depend on the distances among the classes.

Musical pieces written by the same author are grouped together, indicating a stylistic homogeneity among files of the same class. Kohonen network does not preserve the real distance between patterns but the distribution of the classes reveals a topographic relation between stylistic currents. For instance, Beethoven is closer to Fauré and Joplin than to Stravinsky.

ISBN 3-931852-67-9
ISSN 1617-6847

4.2. Experiment 2

In this experiment 48 short musical pieces (stimuli) grouped in 4 classes were analyzed. The structural parameters characterizing each class are tonality and salience. Salience is a musical parameter operatively defined as the redundancy of interval or rhythmic parameters of the musical piece [5]. The 4 classes are defined as Tonal/Salient, Non-tonal-Salient, Tonal/Non-salient, Non-Tonal/Non-salient. An Italian musician wrote, under specific stylistic indications, all the 48 stimuli.

The percentage of correct assignment from relative entropy is shown in Table 2.

Classes	Correct assignments (%)
Tonal/Salient	75
Non-Tonal/Salient	41.6
Tonal/Non-Salient	50
Non-Tonal/Non-Salient	41.6
Tonal	95.8
Non-Tonal	50
Salient	75
Non-Salient	58.3

Table 2: Correct assignments (%) from Relative Entropy calculated on 48 musical stimuli characterized by tonality and salience.

The method reveals a good efficiency in assignment of tonal pieces to the correct class. This could be motivated by a non-optimal textual coding of musical pieces for salience description

The Kohonen neural surface obtained after training the network with the 48-dimensional extracted points is shown in Figure 2.

The stimuli show a good ability to be organized in well-defined classes. The error in assigning a stimulus to the correct class is about 2%.

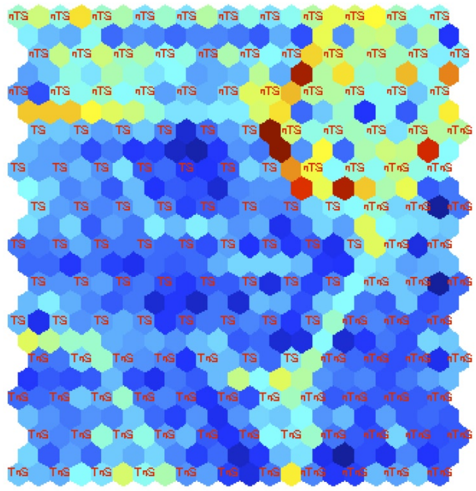

Figure 2: Distribution of 48 musical stimuli over a 2-dimensional neural surface of a 13-by-11 Kohonen map (supervised learning).

5. CONCLUSIONS

In the present study, we proposed the application of relative entropy extraction from musical data and showed some preliminary results on well-characterized musical pieces. Relative entropy seems to represent a good parameter to identify music and to detect similarities among musical pieces. Since entropy plays a crucial role in modern reductionism [6], it is not surprising that entropic analysis of complex data as music is able to reveal hidden common parameters shared by different patterns. In psychology of music, relative entropy method could also represent an efficient approach to characterize music and to associate a well-defined descriptive parameter to musical stimuli.

6. REFERENCES

1. Patel, A.D. (1998). Syntactic Processing in Language and Music: Different Cognitive Operations, Similar Neural Resources?. *Music Perception, Fall, 16(1).*
2. Benedetto, D., Caglioti, E. and Loreto, V. (2002). Language Trees and Zipping. *Physical Review Letters, 88.*
3. Shannon, C.E. (1948). A Mathematical Theory of Communication. *The Bell System Technical Journal,* 27.
4. Kohonen, T. (1995). *Self-Organizing Maps*. Springer, Berlin.
5. Olivetti Belardinelli, M., Rossi-Arnaud, C., Pitti, G. and Vecchio, S.M. (2000). Looking for the Anchor Points for Musical Memory. *Proceedings of the 6th International Conference on Music Perception*. Keele University.
6. Deco, G. and Obradovic, D. (1996). *An Information-Theoretic Approach to Neural Computing*. Springer, Berlin.

ISBN 3-931852-67-9
ISSN 1617-6847