RECOGNISING THE COMPOSITION STYLE BY THEMES SELECTED FROM THE "SONATA" REPERTOIRE: A COMPARISON BETWEEN THE MATHEMATICAL MODELING AND THE EXPERIMENTAL PSYCHOLOGICAL RESULTS

Anna Damiani[1], Pietro Di Lorenzo[2], Giuseppe Di Maio[2] , Marta Olivetti Belardinelli[3]

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

[2] Dipartimento di Matematica, Seconda Università degli Studi di Napoli, Caserta, Italy

[3] Dipartimento di Psicologia, Università di Roma "La Sapienza", Roma, Italy

ABSTRACT

A lot of studies propose the possibility to build up a mathematical model for establishing the distance between musical fragments. The main part of these works deal with melodic similarity using mathematical functions. Some psychological studies suggest the existence of an identification process of salient cues: each subject could pick up distinctive musical features during listening. This work compares the results of a mathematical model of musical similarity to the psychological experimental findings. We select sixty musical thematic fragments (10 and 12 seconds) from piano solo Sonatas composed by Mozart, Beethoven, and Clementi. Each musical excerpt is codified into numerical arrays, paying attention to encode all significant parameters of each sequence in to multivariate framework. We compute the distance between musical sequences using the cross-correlation function. The psychological experiment includes three tasks: 1) familiarization with the sequences; 2) classification of the same musical excerpts in two appropriate categories (X or Y) without any information about the author's name or the pieces composition style; 3) evaluation of their "goodness of representation" using a numerical rank from 1 to 10. Sixty subjects, naïve, i.e. without specific musical education, listen to three series of forty sequences (i.e. twenty from one composer and twenty from another one) mixed in random order. The cluster analysis theory allows us to classify both the mathematical and psychological findings. The preliminary results show that: (i) the naïve listeners classify the musical sequences abstracting salient features in relation to musical style differentiations; (ii) they assess the representativeness' degree to each sequence, by virtue of the differences and similarities detected during the listening. The mathematical model, applied by means of coding criteria selected with *a priori* musicological analysis, provide a similar classification.

1. BACKGROUND

When a listener experiences the form of a complete piece or short musical sequences, he builds up a mental representation which is based on some points of reference (salient cues), that are abstracted during the processing. The cues picked up during the listening are meaningful and distinctive features identified as *invariants* (Deliège & El Ahmadi 1990) at hierarchically different levels of the musical structure.

The salient cues become the starting point around which a comparative evaluation of new data is organised. Furthermore, the collected cues generate classifications and underpin the categorization process of the musical structures (Deliège, 1996).

One of the fundamental notions about the categorisation processes is the well-known concept of *prototype*. It assumes that in all categories there is one token which better represents its category and summarizes the basic features of all the other representatives (Rosch, 1973, 1975). Thus, it seems that the prototype tends to act as "imprint" stored in the memory (Deliège, 2001).

From a mathematical point of view, several papers provide models for musical similarity (Selfridge-Field, 1999; Hofmann-Engl, 2001; Di Lorenzo, 2002). Nevertheless, there are no works concerning the comparison between model simulations and psychological findings. A first paper in this interdisciplinary filed is Damiani (2002). This experimental work is in the same direction too.

2. AIMS

The psychological goal of this work is to measure the capacity of naïve listeners in abstracting salient features with respect to music style, here called *stylistic cues*. We invited the subjects to classify the complete series of sequences within two categories: we furnished no information about the listened music (age, author' names, etc.).

The second and most important goal is to characterize all musical variables of each listened sequences, in order to implement a model of distance among musical fragments.

3. METHODS

3.1 Musical data and psychological method

Each subject listened to thematic fragments selected from piano solo *Sonatas* repertoire by Beethoven, Mozart, and Clementi, authors belonging to the same historical period and who composed music according to the same form (*Forma Sonata*). We selected sixty relevant musical sequences, with the same length in time (ten/twelve seconds). Each subject listened to three series of forty sequences (i.e. twenty by one composer and twenty by

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

another one) lined up in random. All the subjects were naïve i.e. without a specific musical education.

The psychological experiment planned three tasks:

- Familiarization with the material;
- Classification of the fragments and evaluation of their degree of identity and difference;
- Evaluation of the degree of "goodness".

Further details about the experimental procedure are in another specific paper included in ESCOM 5th Conference Proceedings, (see Damiani & Belardinelli, *Recognition of composer's style into musical fragments*)

3.2 Mathematical methods and model

We suggest and apply an operative definition of music that can be more commonly accepted according to the signal theory. Thus, music is defined as a multi-channel and multivariate time signal. For each musical fragment, we look for the maximum number of synchronic independent melodic lines (or horizontal elements with a single pitch for each time) to obtain the number of multivariate channels (or simultaneous arrays). Then, we encode each previously selected channel (with the same time duration) into five one-dimensional arrays of integers. Each array represents the pitches of the notes (according to MIDI), the note duration, the metric-rhythmic accent, the tonal role, and the number of contemporary pitches exceeding the fourth one, respectively. **Code for relative pitch.** Every note corresponds to a numerical bit (from 1 to 12), according to the rule: C=12, C_1=24, C_2=36 etc.. This code is justified on a perceptual basis: people distinguish the intervals in a melody that is the distance between to notes. 0 codifies rests. **Code for note duration.** Note duration is conventionally set at multiple of the elementary time bit. We represent every longer figure with a copy of the same code as much times as it covers up duration. This code is ambiguous because it collects different rhythmic events in the same equivalent code class. Thus, we state in the second variable array a marker of duration. The begin of the sound (or pause) is coded with value 1. 0 codifies the prolongation of the same sound. **Code for metric-rhythmic accent.** Every bit brings up a numerical value to codify the metric-rhythmic accent (strong metric accent, i.e. the first sound in a bar, medium accent for third "movement" in a quaternary time, etc.). Syncope gives rise to a translation of the accent. Then, the syncope note will have a bit increase (+1) while the extremes sounds accent bits will decrease (-1). **Code for tonal role.** We encode it according to consonance perceptual aspects of the scale (in the tonality of the fragment). We classify the notes into five classes: all natural notes (belonging to the tonality of fragment), the minor third degree, all modulating notes (minor seventh, augmented fourth etc.), the minor sixth, every other note (that it is not belonging to the selected tonality). We decided to codify the synchronic aspect into the model according to the following procedure. First, we recognize the maximum number of independent melodic lines into the fragment. Then, we classify all lines in decreasing order according to perceptual melodic pattern importance, assigning to each line a numerical decreasing weight. Finally, we assign an integer number equal to the number of exceeding pitches.

The cross-correlation function (that is the mathematical core of the model) is a well-defined and known statistical function, successfully used in several natural sciences (electric circuits, geophysics etc.). We compute the partial (multivariate) cross-correlation function $R^i_{xy}(\tau)$ between each pair of single array melodic lines: its value for lag time τ=0 is used to define the index (real-valued). Then, we compute the total cross-correlation function $RR_{xy}(\tau)$ as follows:

$$RR_{xy}(\tau) = \sqrt{\sum_{i=1}^{5} w_i [R^i_{xy}(\tau)]^2}\ , \qquad (1)$$

with

$$\sum_{i=1}^{5} w_i = 1\ , \qquad (2)$$

where w_i represents the weight, i.e. the relative importance of melodic pattern ($i \in \{1,2,3,4\}$) and the chord array importance ($i = 5$). Finally, we quantify the relative musical distance between x(t) and y(t) by setting:

$$d_{xy} = d(x(t), y(t)) = \frac{1}{|RR_{xy}(\tau)|} - 1\ . \qquad (3)$$

According to this model, proximal music fragments have distance near 0; vice versa substantially different ones have a greater distance. Using these distances, we build up a numerical square matrix *(d_{xy})* and study its cluster properties.

On the other hand, starting from the experimental psychological results, we compute a "rank index" that represents the human categorization of each set of the fragments (10 sec and 12 sec). We define r_i the rank of each fragment (for all N subjects) with:

$$r_i = a_i \times S_i\ , \qquad (4)$$

where

$$a_i = \frac{1}{N}\sum_{j=1}^{N} a_{ij}\ , \qquad S_i = \frac{1}{N}\sum_{j=1}^{N} S_{ij}\ , \qquad (5)$$

with a_i the frequency of "centred" answers (i.e. the right classified fragments) and S_i the degree of "goodness" of each fragment. Finally, we compare the human and the model classification according to the cluster analysis theory.

4. RESULTS

In the following, we show the preliminary results obtained form the Mozart – Beethoven fragments. The findings confirm the hypothesis: the subjects can classify musical style by differentiations and they abstract different musical patterns. The subjects allot fragments in a well-balanced way. They assign the sequences to the two families (that we identify with Mozart's and Beethoven's categories) on the basis of specific attributes referred

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

to different musical patterns like melodic and rhythmic contours and dynamic changes. Most subjects classify the sequences belonging to well-defined categories. Then, each subject also assesses "the goodness of representation" of each fragment on a scale from 1 to 10, in relation to the two prototype categories, which have been used for the previous categorization tasks.

The rank index, defined by (4), allows us to delineate which fragments are estimated by the subjects as the most representative of the Mozart and Beethoven's composition style. Using rank, it is also possible to recognize two clusters. Figures 1 and 2 represent the cluster analysis based on the rank index obtained from psychological experimental results. There are only three fragments with an unclear classification: B12, B18, B19. Indeed, they are very often classified wrongly.

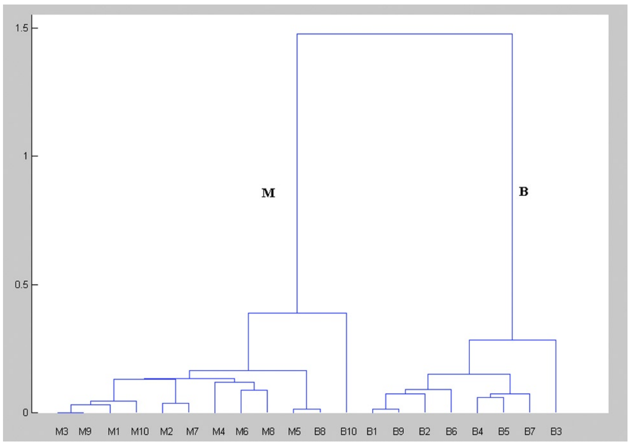

Figure 1: Dendrogram of rank index (10 seconds).

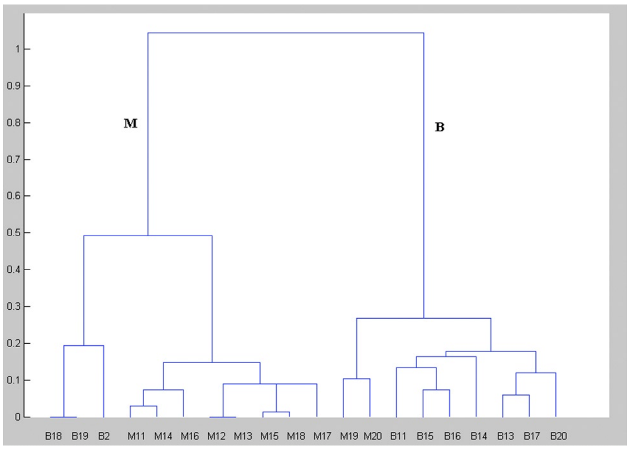

Figure 2: Dendrogram of rank index (12 seconds).

We also analysed distances, computed according to (3), for each set of fragments. The distances take in account a variable set of weights, selected according to a musicological perspective (based on relative importance of the melodic pattern). Figures 3 and 4 show the dendrogram of the cluster analysis computed by distances (with variable weights) between data (10sec and 12sec fragments).

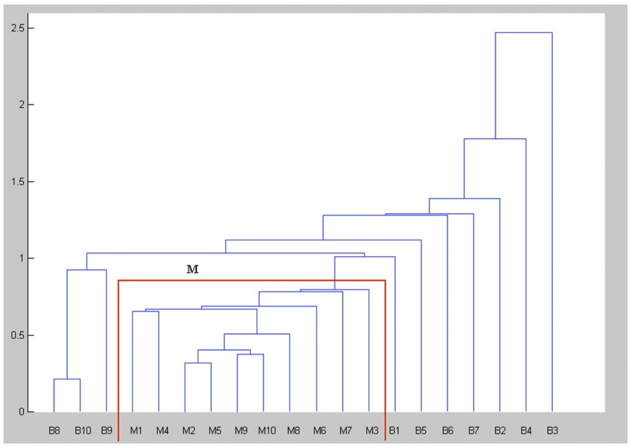

Figure 3: Dendrogram of model results (10 seconds).

There is a good cluster of fragments "M" (i.e. Mozart) that it is a subset of tree of "B" (i.e. Beethoven) family. There are only two fragments (B12 and B19) in the cluster "M": human subjects also classify them in "M" category.

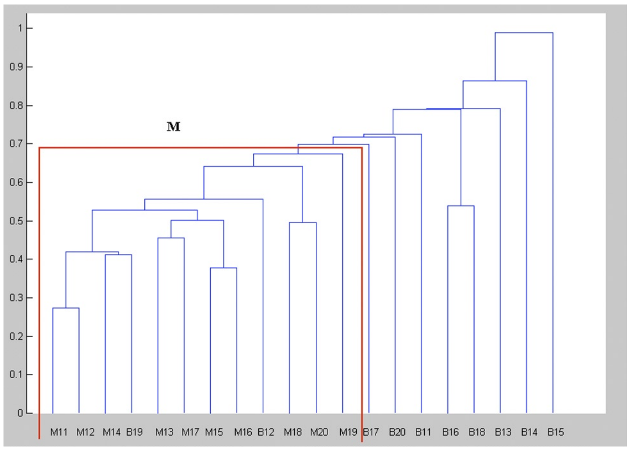

Figure 4: Dendrogram of model results (12sec).

5. PRELIMINARY CONCLUSION

The psychological results are shown in another specific paper in ESCOM 5th Conference (see Damiani & Belardinelli, *Recognition of composer's style into musical fragments*).

The main difference between the human and the model cluster analysis results is the structure of the trees. While human subjects classify fragments in two separated branches, the model provides a more complex structure based on a "nested" pattern, with little distances between the branches. Moreover, Mozart cluster seems to belong to Beethoven cluster. At first glance, the mathematical model of distance based on the cross-correlation (multivariate) function provides a significant fit of the pattern of the experimental psychological data. Moreover the weights, selected with an *a priori* musicological analysis, should be tested on a greater data set and with a more specific psychological experiment.

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

6. REFERENCES

1. Bendat, J. S., Piersol, A. G. (1970). *Random data*. Wiley: New York.

2. H. B. Lincoln (ed.) (1999). *The computer and music*. Cornwell U. Press.

3. Deliège, I. (1989). A perceptual approach to contemporary musical forms. *Contemporary Music Review, 4, 213-230*.

4. Deliège, I. (1996). Cue-Abstraction as a Component of Categorization Processes in Musical Listening. *Psychology of Music, 24 (2) 131-156.*

5. Deliège, I. (2001). Introduction, Similarity Perception – Categorization – Cue Abstraction, *Music Perception, 18 (3), 233-243*

6. Deliège, I. (2001). Prototype Effects in Music Listening: An Empirical Approach to the Notion of Imprint, *Music Perception, 18 (3), 371-407.*

7. Damiani, A. - Di Lorenzo, P. - Di Maio, G. - Olivetti Belardinelli, M. (2002). A Mathematical Model to classify the musical patterns of listened melodic pieces. In *Proceedings of 2nd International Conference "Understanding and creating music", Caserta, November 21-25, 2002.*

8. Di Lorenzo, P. (2002). A mathematical model for a metric index of melodic similarity in *Acta of International Conference "Musical Creativity - 10th anniversary ESCOM", Liège (Belgium).*

9. Hofmann-Engl, L (2001). Towards a cognitive model of melodic similarity. in *Proceedings of the 2nd annual ISMIR, Bloomington, Indiana, 2001*.

10. Murtagh, F. and Heck, A. (1987). *Multivariate data analysis.* D. Reidel Pubb. Comp.: Dordrecht Boston.

11. Rosch, E. (1975) Cognitive reference points. *Journal of experimantal Psychology: General, 194, 192-223.*

12. Rosch, E. (1978). Principles of categorisation. In E. Roach & B.Lloyd (Eds), *Cognition and categorisation* (pp.28-49). Hillsdale N. J., Lawrence Erlbaum.

13. Hewlet, W.B. – Selfidge-Field, E. eds. (1998). *Melodic similarity.* MIT Press: Cambridge MA.

14. Zaripov, R.C. (1970). *Musica con il Calcolatore*. F. Muzzio: Padova.

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