



# Analytical Techniques for the Identification of a Musical Score: The Musical DNA

Michele Della Ventura<sup>(✉)</sup>

Department of Information Technology, Music Academy “Studio Musica”,  
Via Andrea Gritti, 25, 31100 Treviso, Italy  
michele.dellaventura@tin.it

**Abstract.** In the information age, one of the main research field that is being developed is the one related to how to improve the quality of the search engine as regards knowing how to manage the information contained in a document in order to extract its content and interpret it. On the one hand due to the heterogeneity of the information contained on the web (text, image, video, musical scores), and on the other hand to satisfy the user who generally searches for information of a very different type. This paper describes the development and evaluation of an analytical method for the analysis of musical score considered in its symbolic level. The developed method is based on the analysis of the fundamental elements of the musical grammar and takes into account the distance between the sounds (which characterize a melody) and their duration (which makes the melody active and alive). The method has been tested on a set of different musical scores, realizing an algorithm in order to identify a musical score in a database.

**Keywords:** Musical DNA · Information retrieval · Musical score search engine

## 1 Introduction

Internet and the Web are an immense information resource. As such, it represents the first global communication network that allows users to transmit, receive, communicate, and make available information contents. Given the complexity and heterogeneity of the information contained on the web [1], search engines are becoming crucial to allow easy navigation through the data.

Search engines are based on specific algorithms for Information Retrieval (IR). The main objective is to make the right information available to the user based on his requests and expectations [2]. In order to be identified by Information Retrieval (IR) Systems, documents are generally transformed into an adequate representation [3]. Each method of information recovery has a different model based on the type of document [4, 5]: text, image, video.

From these considerations one can immediately infer that while in the case of a linguistic text it is easy to create indexes [6, 7], in the case of a musical language significant difficulties emerge. As far as a musical piece is concerned, the indexes are created exclusively in reference to the title, the name of the author, the tonality and other information of a purely textual and informative kind [8]. Recent research

analyzed the audio files, creating Audio Search Engines, through the audio fingerprint technique that may be used not only to identify an audio file [9] but also to synchronize multiple audio files [10, 11].

However, in the case of a musical text examined at a symbolic level (i.e. the musical score), it becomes difficult to create indexes in the absence of specific indications that often lead to approximate results. There are many scientific researches in the ambit of the musical text that have the objective of searching for more accurate systems in order to determine the identifying elements of a composition, as for instance a melody, a motif, a rhythmic structure [12–14] and so forth, that might be used in order to create an indexation of the same composition.

This document presents a method for the analysis of a musical score considered in its symbolic level to analytically represent its distinctive characters. These refer to the score and are unique: each music is different from another music and each person can interpret these differences by listening and segmenting the continuous sonorous. The representation of these characters through a vector permits to define an objective comparison criterion. The method is based on the mathematical formalization of the distinctive elements of the sound: pitch and duration.

This paper is organized as follows.

Section 2 analysis the concept of “sound” and its characteristics. Section 3 explains the method used to obtain the Fingerprint of a musical score. available experimental results are shown that illustrate the effectiveness of the proposed method. Finally, Sect. 5 concludes this paper with a brief discussion.

## 2 The Concepts of Sound

Melody and rhythm are two fundamental components as far as musical structuring is concerned, two nearly inseparable components: a melody evolves along the rhythm in the absence of which it does not exist [15]: *“melody in itself is weak and quiescent, but when it is joined together with rhythm it becomes alive and active”* [16] (Fig. 1).



**Fig. 1.** Excerpt from the score of Ravel’s “Bolero”. The initial notes of the theme are represented on the first staff without any indication with respect to rhythm; the same notes are represented of the second staff together with the rhythm assigned by the composer.

### 2.1 The Melody

The melody of a musical piece is represented by a number of sounds, each one separated from the next by a number of semitones: the melodic interval.

The various melodic intervals were classified as symbols of the alphabet [17]. The classification of an interval consists in the *denomination* (generic indication) and in the *qualification* (specific indication) [18, 19]. The *denomination* corresponds to the number of degrees that the interval includes, calculated from the lowest one to the highest one; it may be of a 2<sup>nd</sup>, a 3<sup>rd</sup>, 4<sup>th</sup>, 5<sup>th</sup>, and so on.; the *qualification* is deduced from the number of tones and semi-tones that the interval contains; it may be: perfect (G), major (M), *minor* (m), *augmented* (A), *diminished* (d), *more than augmented* (A+), *more than diminished* (d-), *exceeding* (E), *deficient* (def).

A melody is usually represented as a sequence  $S_i$  of  $N$  intervals  $n_x$  indexed on the basis of their order of occurrence  $x$  [17]:

$$S_i = (n_x)_{x \in [0, N-1]}$$

The musical segment may, therefore, be seen as a vector the elements of which are, respectively, the intervals that separate the various sounds from one another. The corresponding value of every interval equals the number of semi-tones between the  $i$ -th note and the preceding one: this value will be respectively positive or negative depending on whether the note is higher or lower than the preceding note (Fig. 2) [17].



**Fig. 2.** Melodic segment and its related vector.

## 2.2 The Rhythm

The rhythm is associated with the duration of the sounds: duration intended as the time interval in which sound becomes perceptible, regardless of whether it is due to a single sign or to several signs joined together by a value connection [17, 20, 21].

If we were to analyze a score, the sound duration will not be expressed in seconds but calculated on the basis of the musical sign (be it sound or rest) with the smallest duration existing in the musical piece [17]. The duration of every single sign will therefore be a (integer) number directly proportional to the smallest duration. In the example shown in Fig. 3, the smallest duration sign is represented by the thirty-second note to which the value 1 is associated (automatically): it follows that the sixteenth note shall have the value 2, the eighth note the value 4, ...



**Fig. 3.** Rhythmic segment and its related vector.

On the base of all the above considerations related to the concept of melody, it is possible to deduce that each melody has a succession of sounds that differentiates it from other melodies (i.e. there are different intervals that separate sounds) and each sound has a specific duration that confers meaning to the whole melody. Figure 4 shows two incipits derived from two different songs that have the same sounds but different rhythm.



**Fig. 4.** A) “Frere Jacques” (French popular song), B) “DO-RE-MI” (from the Musical “The sound of Music”).









3 The DNA of a Musical Score

The aim of this study is to mathematically formalize the features of the sound that allow a listener to recognize a particular composition: the pitch and the duration. These elements are unique for each composition and therefore allow to delineate the “DNA” of the composition.

The difference in pitch between two sounds allows you to define the musical interval; the sequence of intervals within the composition allows the definition of the melody. In order to analyze and represent the succession of the interval within the musical piece it has been used the Markov Process (or Markov Stochastic Process – MSP): the choice was made to describe the passage from one sound to the next sound considering the number of semitones between the two sounds (melodic interval), the trend of the interval (a = ascending or d = descending) and the duration of the two sounds.

Table 1 shows an excerpt of a transition matrix: the first column (and first row) indicates the denomination of the interval (classification), the second column (and the second row) presents the number of semi-tones that make up the interval (qualification), the third column (and the third row) displays the ascending (a) or descending (d) movement between two consecutive sounds, and the fourth column (and forth row) presents the time-space between two consecutive sounds.

**Table 1.** Example of a transitions matrix.

Interval				2° m	2° m	2° m	2° m	...
	Semitones			1	1	1	1	...
		Trend		a	d	a	d	...
			Duration					...
2° m	1	a						...
2° m	1	d						...
2° m	1	a						...
2° m	1	d						...
...	...	...	...	...	...	...	...	...

To determine the DNA of the score (or of the musical research segment) it is necessary to fill in the matrix of the transitions and then for each row it is necessary to add the values present in the matrix: if the result is zero in the DNA column (and in the cell corresponding to the same row) the number zero is written, while if the value is different from zero in the DNA column (and in the cell corresponding to the same row) the number 1 is written.

**Table 2.** DNA of the musical object in Fig. 1.

I				2° m	2° m	2° m	2° m	2° m	2° M	2° M	2° M	2° M	3° m	3° m	3° m	3° m	3° M	3° M	3° M	3° M	...
	S			1	1	1	1	2	2	2	2	2	3	3	3	3	4	4	4	4	...
	T			a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	d	a	...
		D																			DNA
2° m	1	a												1							1
2° m	1	d																			0
2° m	1	a																			0
2° m	1	d																			0
2° M	2	a		1				1												1	1
2° M	2	d																			0
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...

Given a musical segment  $S_2$  of length  $N_I < N_O$ , where  $N_O$  is the length of the music score, it is necessary to define its representative vector following the procedure described above, and compare it with the representative vector of the music score. To identify the DNA of the segment  $S_2$  it is necessary to use a transition matrix with the same columns and rows of the music score: only in this way it is possible to obtain a vector with the same length as the vector of the score.

The comparison takes place by making a bit-to-bit difference, in correspondence with the bits with value 1 of the segment  $S_2$ : if the resulting difference is zero for each bit, the score was identified (Fig. 5) otherwise the choice is made considering the musical score with the greatest number of zeros (which means a greater number of common elements between the musical segment  $S_2$  and the musical scores) (see Fig. 6). In this case the concept of similarity is taken into consideration, as a discriminating factor [22]. The distinction between similarity and identity of two musical segments is very rigorous [23]: similarity is defined as partial identity, that is, two entities (A and B) are similar if they share some properties, but not necessarily all. Therefore, the similarity between A and B depend on their common features: the more common features they share, the more similar they are (Fig. 6). At the same time, the similarity between A and B depend on the differences between them: the more differences they have, the less similar they are.



<b>Frère Jacques</b>																								
DNA	0	1	1	0	0	0	0	0	0	0	1	1	1	0	0	0	0	0	0	0	...			
Vector $S_2$											1					1					...			
Check											0					-1					...			

<b>DO-RE-MI</b>																								
DNA	1	1	0	1	1	0	0	0	0	1	1	1	1	1	1	0	0	1	0	1	...			
Vector $S_2$										1					1						...			
Check										0					0						...			

Fig. 5. Musical score identification.

<b>Musical Score 1</b>																								
DNA	1	1	0	1	1	0	0	0	0	0	1	1	1	1	1	0	0	1	0	1	...			
Vector $S_2$							1				1	1				1					...			
Check							-1				-1	0				0					...			

<b>Musical Score 2</b>																								
DNA	1	1	0	1	1	0	1	0	0	0	1	1	1	1	1	0	0	1	0	1	...			
Vector $S_2$							1				1	1				1					...			
Check							0				-1	0				0					...			

Fig. 6. Musical score identification.

## 4 Application and Analysis: Obtained Results

The aim of this research was to develop an analytical method in order to realize a Search Engine able to identify a musical score in a database.

To evaluate the proposed method, it has been realized an algorithm the structure of which takes in consideration each and every single aspect previously described. The algorithm does not provide any limitations with regard the number of notes that can be wrote for the research and therefore the musical score representative matrix that is automatically dimensioned on the base of the features of the musical score. The algorithm presents some limitations related to the duration of the notes used for the research. This is due to the fact that the solidity of the algorithm had to be checked against complex rhythmic structures (typical of the contemporary music).

### 4.1 Database Preparation

The collection of documents was prepared using musical scores of 18<sup>th</sup>, 19<sup>th</sup> and 20<sup>th</sup> Century of different authors, for a total of 220 musical scores.

All of the musical scores were stored in MIDI format (a symbolic music interchange format), or in the Lilypond notation format (a textual notation for music based

on the set of ASCII characters) [24], or in the more recent MusicXML file (a text-based language that permits to represent common Western musical notation) [24].

All the musical scores were for piano, in order to take polyphony into consideration; grace notes (or musical ornaments) were removed from each score (Fig. 7), because musically they are secondary notes and do not characterize the musical structure.



Fig. 7. Database preparation: removing grace notes.

## 4.2 Musical Score Decoding

In the decoding stage, it is important to underline the fact that in the case of two or more sounds connected together, these are considered as a single sound whose duration is equal to the sum of the durations of the individual sounds [17] (see example in Fig. 8).



Fig. 8. Sound duration.

Each musical score (read from the file) is transformed into a list of numbers (Musical DNA) [17]: each sound is associated with two values, the first one related to the distance with the next sound and the second one related to its duration [25, 26] (Fig. 9).



Fig. 9. Musical DNA.

In case the musical score is composed of several voices placed on different staves (such as the piano score) the list of sounds is composed first of all by the relative numbers of the first staff, then those of the second staff and so on (Fig. 10) [12].

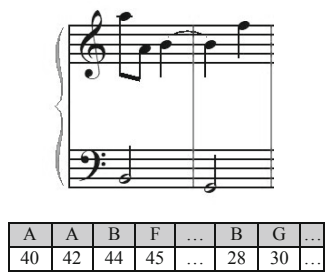


Fig. 10. Sequential representation of the score.

### 4.3 Input Sample Music

After filling in the representative matrix of the musical score, its binary representative vector (Musical DNA) is defined: see example in Table 2.

The search for a score in a database is performed by writing the sounds of a sample musical segment (with their respective durations) within a dialog similar to the dialog of a search engine, with the difference that this is represented by a musical pentagram (Fig. 11).

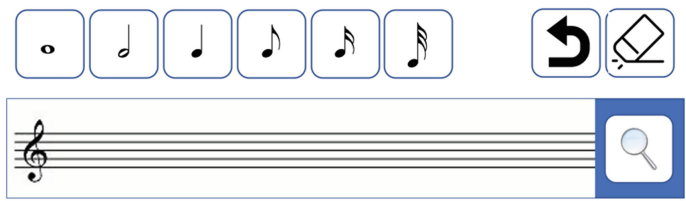


Fig. 11. Music search engine interface.

### 4.4 Obtained Results

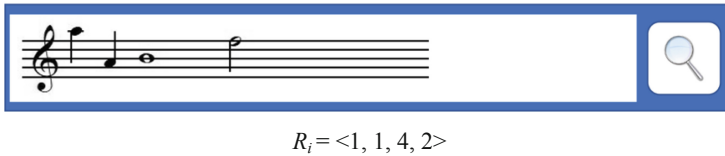
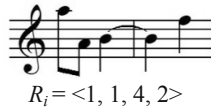
The initial tests were carried out on a set of sample musical segments of two different lengths, 5 and 10 musical notes. For the tests the processing time was not taken into consideration, because it was strictly connected to the type of computer used for the analysis.

It was not important that the durations of the sounds indicated in the sample were the same as the durations present in the scores: the durations could be different under the condition that the mathematical proportion was always respected (Fig. 12).

The results of information retrieval are present in Table 3. The percentage improvement obtained considering 10 notes instead of 5 notes is approximately 5% in the case of musical scores of the 18th Century, 7% in the case of musical scores of the 19th Century and 4% in the case of the musical scores of the 20th Century. This is explained by the fact that: in the 18th Century music, the type of intervals present in a score is not very varied; in 19th Century music the type of intervals increases and



therefore a greater number of notes allows to diversify the research results; finally, in the music of the 20th Century there is a particular element, namely the presence of complex rhythms which reduces the similarity between musical segments.



**Fig. 12.** Proportional duration.

**Table 3.** Comparative performance under different length of the sample musical segment.

Period of the musical scores	Averaged recognition rate (%)	
	5 musical notes	10 musical notes
18 <sup>th</sup> century	87	92
19 <sup>th</sup> century	84	91
20 <sup>th</sup> century	63	67

## 5 Discussion and Conclusions

The Internet and the Web represent an immense and heterogeneous information resource for a vast and heterogeneous public like current Internet users. However, the nature and characteristics of the Internet highlight problems relating to how to search and find information online, in order to satisfy the user's requests.

This study has been able to successfully develop an algorithm (a sort of IR system) that allows the identification of a musical score within a database, through the indication of musical notes. The results show that to obtain a satisfactory result it is necessary to insert a suitable number of notes, capable of analytically describing the identifying characteristics of the musical score.

In addition, the following study areas can be improved in future studies to create a more robust IR:

- Increase in the size of the database size (number of musical scores);
- Representation of irregular musical rhythms (typical of contemporary music).

## References

1. Dong, Y., Chawla, N.V., Swami, A.: metapath2vec: scalable representation learning for heterogeneous networks. In: Proceedings of the KDD, pp. 135–144 (2017)
2. Cao, X., Zheng, Y., Shi, C., Li, J., Wu, B.: Meta-path-based link prediction in schema-rich heterogeneous information network. *Int. J. Data Sci. Anal.* **3**(4), 285–296 (2017). <https://doi.org/10.1007/s41060-017-0046-1>
3. Blummer, B., Kenton, J.M.: Information research and the search process. In: Improving Student Information Search, pp. 11–21. Chandos Publishing (2014)
4. Lewandowski, D.: A framework for evaluating the retrieval effectiveness of search engines. In: Jouis, C., Biskri, I., Ganascia, G., Roux, M. (eds.) *Next Generation Search Engines: Advances Models for Information Retrieval*, pp. S456–S479. IGI Global, Hershey (2012)
5. Lazonder, A.W., Biemans, H.J.A., Wopereis, I.G.J.H.: Differences between novice and experienced users in searching information on the world wide web. *J. Am. Soc. Inf. Sci.* **51**(6), 576–581 (2000)
6. Chang, G., Healey, M.J., McHugh, J.A.M., Wang, J.T.L.: Multimedia search engines. In: Chang, G., Healey, M.J., McHugh, J.A.M., Wang, J.T.L. (eds.) *Mining the World Wide Web*. INRE, vol. 10, pp. 51–63. Springer, Boston (2001). [https://doi.org/10.1007/978-1-4615-1639-2\\_4](https://doi.org/10.1007/978-1-4615-1639-2_4)
7. Yuan, X.J., Zhang, X.J., Avery, J.: Seeking information with an information visualization system: a study of cognitive styles. *Inf. Res.* **16**(4), 499 (2011)
8. Zmudzinski, S., Steinebach, M., Butt, M.: Watermark embedding using audio fingerprinting. In: Shi, Y.Q., Katzenbeisser, S. (eds.) *Transactions on Data Hiding and Multimedia Security VIII*. LNCS, vol. 7228, pp. 63–79. Springer, Heidelberg (2012). [https://doi.org/10.1007/978-3-642-31971-6\\_4](https://doi.org/10.1007/978-3-642-31971-6_4)
9. Miller, M.L., Rodriguez, M.A., Cox, I.J.: Audio fingerprinting: nearest neighbor search in high dimensional binary spaces. *J. VLSI Signal Process. Syst. Signal Image Video Technol.* **41**(3), 285–291 (2005). <https://doi.org/10.1007/s11265-005-4152-2>
10. Oostveen, J., Haitsma, J., Kalker, T.: Algorithms for audio and video fingerprinting. In: Verhaegh, W.F.J., Aarts, E., Korst, J. (eds.) *Algorithms in Ambient Intelligence*. PRBS, vol. 2, pp. 201–219. Springer, Dordrecht (2004). [https://doi.org/10.1007/978-94-017-0703-9\\_11](https://doi.org/10.1007/978-94-017-0703-9_11)
11. Mau, T.N., Inoguchi, Y.: Audio fingerprint hierarchy searching on massively parallel with multi-GPGPUS using K-modes and LSH. In: *Eighth International Conference on Knowledge and Systems Engineering (KSE)*, pp. 49–54. IEEE (2016)
12. Della Ventura, M.: *Musical DNA*. ABEditore, Milano (2018). ISBN 978-88-6551-281-4
13. Neve, G., Orio, N.: A comparison of melodic segmentation techniques for music information retrieval. In: Rauber, A., Christodoulakis, S., Tjoa, A.M. (eds.) *ECDL 2005*. LNCS, vol. 3652, pp. 49–56. Springer, Heidelberg (2005). [https://doi.org/10.1007/11551362\\_5](https://doi.org/10.1007/11551362_5)
14. Rodriguez Lopez, M.E.: Automatic melody segmentation. Ph.D. thesis, Utrecht University (2016)
15. Fraisse, P.: *Psychologie du rythme*. Puf, Paris (1974)
16. Sioros, G., Davies, M.E., Guedes, C.: A generativemodel for the characterization of musical rhythms. *J. New Music Res.* **47**(2), 114–128 (2018)
17. Della Ventura, M.: The influence of the rhythm with the pitch on melodic segmentation. In: Abraham, A., Jiang, X.H., Snášel, V., Pan, J.-S. (eds.) *Intelligent Data Analysis and Applications*. AISC, vol. 370, pp. 191–201. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-21206-7\\_17](https://doi.org/10.1007/978-3-319-21206-7_17)
18. de la Motte, D.: *Manuale di armonia*, Bärenreiter (1976)
19. Schoenberg, A.: *Theory and Harmony*. Univ of California Pr; Reprint edition (1992)

20. Moles, A.: *Teorie de l'information et Perception esthetique*, Paris, Flammarion Editeur (1958)
21. Louboutin, C.: Multi-scale and multi-dimensional modelling of music structure using polytopicgraphs. *Sound [cs.SD]*. Université Rennes 1 (2019)
22. Della Ventura, M.: Similarity measures for music information retrieval. In: Le, N.-T., Van Do, T., Nguyen, N.T., Thi, H.A.L. (eds.) *ICCSAMA 2017. AISC*, vol. 629, pp. 165–174. Springer, Cham (2018). [https://doi.org/10.1007/978-3-319-61911-8\\_15](https://doi.org/10.1007/978-3-319-61911-8_15)
23. Della Ventura, M.: Speech assessment based on entropy and similarity measures. In: Le Thi, H.A., Le, H.M., Pham Dinh, T., Nguyen, N.T. (eds.) *ICCSAMA 2019. AISC*, vol. 1121, pp. 218–227. Springer, Cham (2020). [https://doi.org/10.1007/978-3-030-38364-0\\_20](https://doi.org/10.1007/978-3-030-38364-0_20)
24. <http://lilypond.org/>
25. Madsen, S., Widmer, G.: Separating voices in MIDI. In: *ISMIR*, Canada (2006)
26. Della Ventura, M.: Using mathematical tools to reduce the combinatorial explosion during the automatic segmentation of the symbolic musical text. In: Nguyen, T.B., Do, T.V., Le Thi, H.A., Nguyen, N.T. (eds.) *Advanced Computational Methods for Knowledge Engineering. AISC*, vol. 453, pp. 281–293. Springer, Cham (2016). [https://doi.org/10.1007/978-3-319-38884-7\\_20](https://doi.org/10.1007/978-3-319-38884-7_20)