Deciphering the Complexities in the Perception, Performance and Comparison of Musical Phrases via Statistical Shape Analysis of the Fractal Nature of Music

Makarand V. Ratnaparkhi
Wright State University
Dayton, Ohio, USA
makarand.ratnaparkhi@wright.edu

1. INTRODICTION

The perception and performance of musical phrases is known to be a complex neurological activity. As a result, while trying to learn music there is always a *frustrating* communication gap between the music student and his/her music teacher. The root causes of such gap is the lack of well-defined quantitative methodology that identifies , almost instantly, the distance between the ideal expected performance and the actual note-by-note musical performance by a music-student ,and the factors that give rise to such distance. In this presentation, an effort is made through the interplay between three subject areas, namely; musicology, fractals (in particular fractal dimensions), and statistical thinking in view of developing a methodology, which with the aid of today‟s computational power ,should be helpful in reducing the above mentioned communication gap and making musical life easier for students and their teachers. Here, the target population is the collection of music-students who are learning to play violin. In particular, our sample consisted of students learning Indian Classical Music. However, the methodology of this paper can be extended to any type of music and such can be considered as of global nature.

The organization of the presentation is as follows:

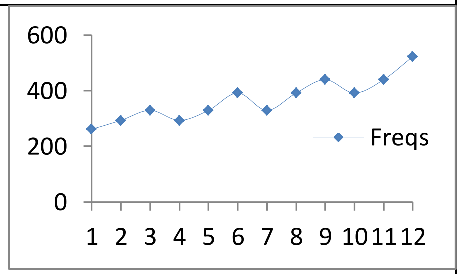
In section 2, first, using an example of short musical phrase, we introduce the concept of the shape of a music composition. Then, we describe the representation of this phrase as a sequential multivariate times series and consider the related note-by-note fractal dimensions (FDs) as quantitative statistical descriptors of the qualitative content of the performance. The configuration of the performance, as a geometric object, is now introduced using the FDs and other variables from musicology such as the duration and the tempo associated with each note in this phrase. In section 3, we introduce the concept of the shape of the performance and the Procrustes analysis technique for finding the distances between the performances and then comparing the shapes of the performances of the main phrase by different subjects (students).Further, the shapes of the performances of the subsets of this main phrase and related distances are discussed, in view of deciphering the complexities in the perception and performance –as indicated in the title of the paper. The description of the data collected for this study and related statistical shape analyses are provided in section 4. The discussion of the results, in view of the main objectives of this study such as the clustering of the performances and identification of the level of perception of the phrases is included in view of demonstrating the usefulness of the methodology introduced in sections 2 and 3.

2. THE FRACTALS AND OTHER DISCRIPTORS OF MUSICAL PERFORMANCES

The music composition is basically a collection of musical notes. We will consider the following simple composition with 12 notes divided into 4 triplets for discussing the objectives of this paper. Note that some of the notes are repeated but with respect to the *aesthetics* of music these have respective effectiveness.

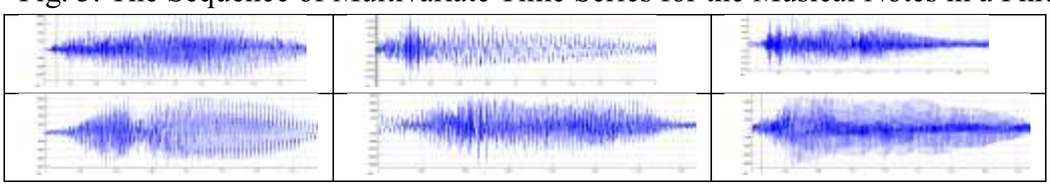
Fig.1. Music Composition (Phrase)
Western **:(C,D,E,|D,E,G,|E,G,A,|G,A,C⁺)**
Indian:(सा रे ग| रे ग प| ग प ध| प ध सां)
(The repeated note is shown in italics)

Fig,2. The Shape of the Music Composition of Fig.1



The above phrase was played on violin. Since, it is known that the violin playing needs special skills such as fingering, bowing, and the enhanced perception for the effective performance the students from the violin class were selected for the data collection. Now, when the above phrase is played on violin and recorded, we get an audio file for 12 notes. This audio file when converted into the digitized data file it consists of a sequential multivariate (12 dimensional) time series. The example of such time series is shown below as Fig.3 (for brevity only six time series for the first six notes from first two triplets played by subject #6 are shown in this figure.)

Fig. 3. The Sequence of Multivariate Time Series for the Musical Notes in a Phrase

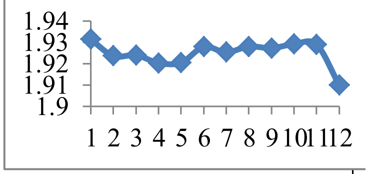


The fractal nature of music in well accepted in musicology. Therefore, for each of the above12 time series (i.e. 12 notes) we calculated the fractal dimension (FD). After studying different methods that are available for calculating the FD of a time series the box method was used for this presentation. We note that ,we considered other methods such as the fast Fourier transforms and the wavelet decomposition were for the above time series data. However, the related analysis becomes more involved for the performances having longer duration. Therefore, the single number, FD in our case, for representing the time series was thought as more suitable. Further, the performance of the FDs as a descriptive measure of qualitative performance of each note was found to be associated with the actual audio-based evaluation of each note. The Table 1 and Fig.4 below show the profile of the FDs of 12 notes played by one of the subjects in this study.

Table 1. The FDs of Twelve Time series (12 Notes) Representing the Performance

Note #	1	2	3	4	5	6
FD	1.93131	1.92355	1.92393	1.9203	1.9204	1.92779
Note #	7	8	9	10	11	12
FD	1.92529	1.92795	1.92726	1.92906	1.92894	1.91004

Fig.4 Sample Plot of FDs for 12 Notes



The comparison of the charts such as those shown in Fig.4 (discussed in the next two sections) for different students provides valuable information about the performance and perception of these students. We note that for even though the comparative studies of the FDs between and within subject are useful we cannot ignore the usefulness of traditional variables such as the duration, tempo associated with each note for assessing the effectiveness of the performance. Therefore, we calculated the duration (in seconds) and the tempo (beats/minute) for each of the 12 notes in the above composition that are played by each student (subject) in this group. These data ,for two subjects, are shown below in Table 2 as a 12x3 matrix for each subject, where the rows correspond to 12 notes the column variables are (FD, Duration, and Tempo). Thus, the matrix configuration provides the profile of the performance by a subject. We note that we have included these variables for demonstrating the application of the methodology that is discussed in next two sections. Further, it is possible to add more variables (both of quantitative or qualitative type) in this configuration as may seem necessary.

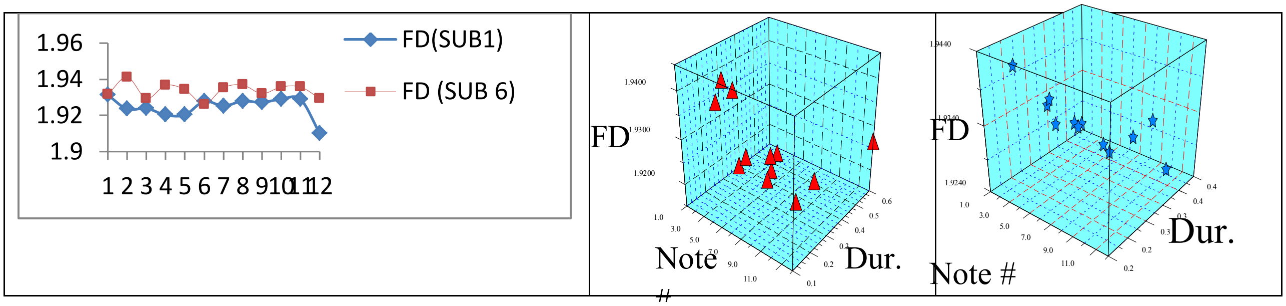
Table 2. Matrix Representation of the Performance Configurations for Subjects #1 and #6.

SUB #1: Matrix of (FD, Duration, Tempo)				SUB #6 Matrix of (FD, Duration, Tempo)			
Note#	FD	Duration	Tempo	Note#	FD	Duration	Tempo
1	1.93131	0.28	214	1	1.93166	0.300	200
2	1.92355	0.27	222	2	1.94123	0.200	300
3	1.92393	0.30	200	3	1.92944	0.320	187
4	1.92030	0.23	261	4	1.93669	0.232	259
5	1.92040	0.24	250	5	1.93426	0.220	273
6	1.92779	0.33	182	6	1.92617	0.335	179
7	1.92529	0.25	240	7	1.93529	0.230	261
8	1.92795	0.25	240	8	1.93685	0.194	310
9	1.92726	0.30	200	9	1.93209	0.365	164
10	1.92906	0.00	273	10	1.93580	0.196	306
11	1.92894	0.28	214	11	1.93575	0.239	251
12	1.91004	0.60	100	12	1.92941	0.302	198

The Shape of the Performance and Related Statistical Analysis

The concept of "shape" of the performance, for brevity, is introduced below (visually) in Fig.5 using the 2D and 3D plots of FDs and other variables from the above Table 2. Then, the formal definition of shape is stated for the statistical shape analysis of performances.

Fig.5. The 2D and 3D Plots of the Performances for Subjects #1 and #6 (Data of Table 2)



Similarly 3D Plots for (Note #,Duration, FD) cab be constructed for 3D shape of performances. The above plots represent the geometric shape (profile) of the performances and also provide a useful tool for note by note comparison of the performances. The formal definition of shape is as follows:

Shape of an object :*The term shape of an object refers to the geometrical information that remains when we filter out translation, rotation and scale information*.

We note that the performance of each subject, as shown in the above Table 2, is a matrix configuration that is a representation of the shape of the performance as shown through the above Fig.5. Thus, the statistical shape analysis of the 15 performances in our data set is equivalent to the shape analysis of the corresponding 15 data matrices(configurations).

Remark 1. The Fig.5 (2D and 3D plots) is a visual description of the performance of 12 notes using the FDs. These plots show that the magnitude of the FDs is not necessarily the same for a note when it gets repeated. For example, the FDs of the note "E" are not the same in any given performance of the above phrase. It changes from one triplet to other. The changes in the magnitudes of FDs are associated with the number of explicit factors such as the fingering and bowing skills that are required for playing and the implicit factors such as the "stimulus " and "perception" that make the playing of each note as " an event in a real time".

Remark 2. The Fig.5 being a geometric figure has a shape of its own and further this shape is associated with the performance.

Remark 3. The 2D and 3D plots for subjects 1 and 6 constructed visually reveal the differences in the performances by these two subjects. Thus, we conclude that the shape of the performance represented by the matrix configuration of Table 2 is meaningful for the practical purposes.

Procrustes Method for Size and Shape Analysis

The Procrustes method is for matching two or more configurations. In our study the configurations are the shapes of the performances of the musical notes that may not have the same origin and same scale.

The matching of two configurations is based on linear regression modeling of one configuration in terms of the other using the similarity transformation. The matching is expressed in terms of the distance between the shapes of the objects that are standardized with respect to the size variable. The formal discussion of Procrustes procedure can be found in Dryden and Mardia (1998). Using this procedure we obtain the distance matrix for the differences in shapes of the objects and use this matrix for further statistical analysis.

4. DISTANCE MATRIX FOR DIFFERENCES IN PERFORMANCES BY 15 SUBJECTS

Here, due the restriction on the length of the paper, it is not possible to reproduce the Table 2 for all the 15 subjects. Also, the distance matrices mentioned in the above section for the complete data are of the order 15x15, and hence cannot be reproduced in this paper. Therefore, for illustrating the usefulness of the methodology we have considered the data for four subjects, and the results obtained using the Proctustes method and related statistical analysis are shown in what follows. However, the extension of the methodology and interpretations can be applied to the shape matrices of any dimension.

Procrustes distance matrix for the shapes of four out of fifteen subjects in this study.

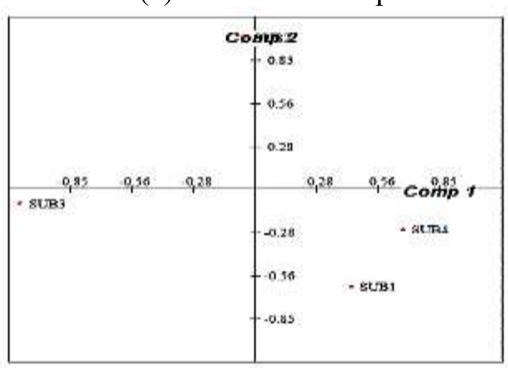
The subjects 1,2,3 and 6 were selected for this limited study. Further, in view of the significance of the performance of the repeated notes in violin playing the distance matrices for all 12 notes and, separately for only the notes “E” (repeated 3 times) were calculated and the corresponding principal componente analysis was performed. The results are shown below in Table 3.

Table 3. Procrustes Distances Using 12 Notes, Principal Components

(a) Distance Matrix(Based on 12 Notes Played by Each Subject)

SUB #	1	2	3	4
1	0.00			
2	6.43	0.00		
3	6.57	6.49	0.00	
4	4.83	6.01	7.02	0.00

(b) Principal Component Analysis Plot of Matrix in (a) with Two Components



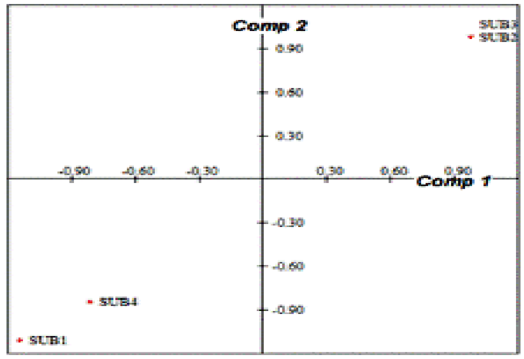
Conclusion: The above table shows that based on the performance of the complete phrase of 12 notes the distance between the performances by subjects 1 and 4 is the minimum and hence their performances have more similarity as compared to that with other players. The distance between the 3rd and 4th players is the maximum. The same conclusion is arrived at by the corresponding principal component analysis. However, the smaller distance between the subjects for the overall performance is not indicative of the perception of the notes that get played in a phrase.Therefore, as an example. the performance of repeated note „E" alone was analyzed for the same subject #s 1,2,3,4. The Procrustes distance and related principal component analysis is shown in Table 4.

Table 4. Procrustes Distances For Repetitions of Note „E" and Principal Component Analysis

4(a) The Procrustes Distances for Repeat Note „E"

	SUB1	SUB2	SUB3	SUB4
SUB1	0.00			
SUB2	0.97	0.00		
SUB3	0.97	0.00	0.00	
SUB4	0.02	0.70	0.70	0.00

4(b) **Principal Components for Table 4(a)**



Conclusion: The playing or the performance of note „E" in all three triplets is very close for the pairs of subjects (2,3) and (1,4). One of the interpretation of this observation is there is closeness between the subjects 2 and 3 with respect to the perception and performance of note „E" with similar conclusion for subjects 1 and 4.

The above conclusion raises an interesting aspect related to the perception and performance of the notes in a given composition and in particular the performance of the repeated notes.

Does such closeness in the performance and (or) perception is due the correlation between the FDs (or the other variables) that are representative of such performances?

To investigate this issue we calculated the correlations of FDs corresponding to different notes. In particular, the pairwise correlations between FDs corresponding to the three repeats of note „E" (occurring in triplet #s 1, 2, and 3), denoted by E_1,E_2 and E_3 respectively, were calculated and the test for zero correlation was carried for each of the three calculated correlations. For all three cases we conclusion was that the correlations between FDs are zero. The result apparently seems surprising, but, given the possibility that the neurological activity that governs the sequential perception and performance of music may have a sort of instantaneous "lack of memory property" or some other more complex unknown property the above observation may not be that surprising to the researchers working the field of neurology. However, from the statistical thinking perspective there seems to be a need to study the sequence of the FDs corresponding a musical performance within the framework of the theory of stochastic processes for exploring the dependence among the sequence of FDs that characterize the note by note performance and the underlying phenomenon of perception.

References

Dryden,I.L. and Mardia,K.L. (1998) *Statistical Shape Analysis*. John Wiley and Sons.