Community detection in visibility networks: an approach to categorize percussive influence on audio musical signals

Dirceu de Freitas Piedade Melo, Inacio de Sousa Fadigas and Hernane Borges de Barros Pereira

Abstract The feature extraction is a very important step in the music audio classification. This task has been performed by renowned descriptors using, in most cases, the time-frequency approach. In this article we propose a descriptor that performs the feature extraction in a set of music audio files labeled in symphonic and percussive music, using parameters calculated within the Euclidean domain. First we calculate the variance fluctuation series of music signal, after we map this series into visibility graphs [13]. At the end each audio track will correspond to a network, where the links are defined by the visibility of variance fluctuations of their respective audio signal. Then, we measure the strength of the partitions of each network in clusters, using calculation of modularity. The results of computation of this parameter in sixty networks showed that percussive and symphonic music can be distinguished and hierarchized on a growing rang, following a direct correlation with modularity.

1 Introduction

Due to the need to develop computational resources for the organization of large digital music libraries, the importance of automatic music classification systems has grown considerably in recent times [19]. Many classification platforms have been proposed [6, 8, 10, 20], and despite efforts to find a new path [9, 12], most feature extraction tools use knowledge of the audio signal processing field [2, 7, 23, 26]. Among the most commonly descriptors used in feature extraction are MFCC - Mel

Dirceu de Freitas Piedade Melo (e-mail: dirceumelo@ymail.com)✉
Department of Mathematics (DEMAT), Nucleus of Studies of Mathematics, Statistics and Education (NEMEE), Federal Institute of Education Science and Technology of Bahia (IFBA), Brazil

Inácio de Souza Fadigas (e-mail: isfadigas@gmail.com)
State University of Feira de Santana (UEFS), Bahia, Brazil

Hernane Borges de Barros Pereira (e-mail: hbbpereira@gmail.com)
State University of Bahia (UNEB), Computational Modeling Program, SENAI CIMATEC, Bahia, Brazil

H. Cherifi et al. (eds.), *Complex Networks & Their Applications V*,
Studies in Computational Intelligence 693,
DOI 10.1007/978-3-319-50901-3_26

Frequency Cepstral Coefficients, Spectral Rollof, Spectral Flux, Zero Crossing Rate, Low-Energy Feature. These algorithms lead their mathematical operations in time-frequency domain in order to extract of the musical signal, three basic characteristics: tone texture (timbre), rhythmic content (time, rhythm, pulse), and tonal content (pitch) [2]. Hoping to cooperate for the growth of new ways to perform feature extraction in musical audio signals, we propose in this article a way to describe musical dynamics [1] in audio tracks following a different direction. To make possible this idea we first captured the loudness of the audio signal from the calculation of the average intensity of their fluctuations in fixed-size windows [12], creating a series of variance fluctuations of the original signal. After this, we mapped this series into a graph, using the geometrical visibility mapping proposed by [13]. In this mapping, if two points of the series see each other in the Cartesian plane, an edge is created in the Euclidean plane. Thus the higher the visibility of a point in the series, the more edges it will have in the graph. At the end of the mapping, the graph inherited in its structure the visibility of all local peaks with their respective neighborhood [15]. Consequently, variance fluctuation series with few local peaks, but very visible, will generate graphs with few hubs, but with a high degree of connections. On the other hand, series with many local peaks with poor visibility will generate graphs with many vertices with lower level of connections. The analysis of modularity will identify if the network structure was created from the series with greater or smaller local visibility. The experiments suggest that the visibility graph generated from the variance fluctuations of audio signals that have a strong influence of percussion activity - like Samba or disco music- have a higher trend to create modules than audio signals whose orchestration has little or no influence of percussion instruments and more dynamics nuances, like a string quartet.

2 Related Works

Researchers at the computer music area have used the structural feature of complex networks to solve various problems related to music information retrieval, such as: musical taste in internet communities [4], algorithmic composition [25], collaborative networks between composers [21], music genre classification [5]. In [25] authors build a network based on pattern analysis of Bach, Chopin and Mozart compositions, linking the duration of two notes in MIDI (Musical Instrument Digital Interface) format which co-occur in a melodic phrase, using universal properties found in these networks to propose rules for algorithmic composition. To analyze the musical tastes of users from their playlists, [4] uses the basic features of networks where the nodes are the song titles, and the edges occur between two song titles, if this title appear in more than one playlist.[5] deals with music genre classification using rhythms extracted from MIDI database, transforming it into complex networks. In [5] each rhythmic cell is a node, while the sequences of notes define the links between nodes,

[1] The varying levels of volume of sound in different parts of a musical performance.https://en.oxforddictionaries.com/definition/dynamics.

according to a Markov model. [11] combines audio analysis and network structures to identify communities of artists on myspace website, establishing links between two artists who have similar tags on social networks, and audio-based similarity using Mel Frequency Cepstral Coefficients, and entropy. [21] studies the topology and evolution of networks of western classical music composers, building links between two composers who co-occur in the same compact disc, linking information about author, period and style extracted from audio file meta-data. A characteristic that can be noticed in most scientific papers that use the mapping of complex networks to understand the music audio phenomena is the absence of structures formed by links where the nodes are non-symbolic elements. With the exception of [11], which use audio data in the network vertex in the first of two phases of the mapping, we have not found in the survey of related work, another study whose network is formed by the relationship between audio signal points. Considering the survey by [22], that shows various approaches for music content analysis, we also note the lack of methodologies that use complex network parameters to perform feature extraction in audio signals.

Visibility graphs have been created bridges between time series analysis and complex networks analysis, opening possibilities on time series field by using a set of new tools. One of this bridges has been used to study long-term correlations, fractal properties, and self-similarity structures [14, 18] and have found applications in temporal observations like Nasdaq and $S\&P500$ daily stock indices [24] and traffic of information packets series [1]. These studies show that the visibility graphs has the ability to capture local trends of time series and measure them through the network analysis. Motivated by these studies, we chose the same type of mapping seeking to identify how much the persistence of an audio signal time series is associated with the dynamics changes influenced by percussive activity of its musical content. This article will show that modularity is able to capture the reflections of the self-similarity and patterns of persistence of loudness embedded in the network, but will not establish a direct relationship with power laws or the Hurst exponent calculations, as in [14].

3 Materials and Methods

In this section we first present the database, after we show the methodological approach to conduct the study of the visibility of an audio signal by using the modularity of complex networks. We take a set of sixty audio samples with 30 seconds long. Each song is represented by a time series W(i). In this series we calculate the subset of variance fluctuations V(j). For each V(j) point is evaluated the "visibility" in relation to their successors and predecessors, according to the slope comparisons [13]. At the end of the process the subset V(j) becomes the graph G(V(n), V(m)), from which is estimated the modularity and the amount of communities.

3.1 Database

The audio files used In this article are divided into two groups: Symphonic Music and Percussive Music. In Symphonic Music were selected thirty compositions for string quartet or large orchestra. The compositions are divided among Bach concertos, Mozart symphonies and string quartets by Debussy and Ravel. To represent the Percussive Music, we chose 30 tracks equally divided into: samba, mangue beat and disco music. The Samba tracks are songs composed for the celebration of the Rio de Janeiro carnival from 2005 to 2014. In Mangue Beat there is an influence of electronic pop-rock music, mixed with a traditional afro-brazilian rhythm called Maracatú. The ten tracks of disco music gives a good overview of the musical scene of the 80s. The symphonic and disco music are chosen from GTZAN [2] database, and samba e mangue beat are from the author's particular collection. The Percussive tracks are labeled as Percussive 1 ... Percussive 30, where disco music occupies the ten first places, samba takes up the next ten, and Mangue Beat the past ten. The Symphonic networks are labeled as Symphonic 1 ... Symphonic 30, where the eight first are Bach concertos, the next sixteen networks are Mozart Symphonies, and the last six are string quartets composed by Debussy, Dutileux and Ravel.

3.2 Calculating The Variance Fluctuation Series

In this section we first calculate the variance fluctuations of a musical signal with the same methodology used by [12, 16], where the authors consider that the loudness can be represented by average intensity of the sound over intervals of 0.01 s. Consider audio music signal represented by the $W(i)$ series, with $i = 1 \cdots N$. The total number of points N is a function $N = SR.t$, where the sampling rate is $SR = 11khz$ and the time is $t = 30$ seconds. The set $W(i) = W(1), \cdots, W(N)$, with $N = 330,000$ is segmented into m-non overlapping boxes $\lambda = 110$. For each box $j = 1 \cdots m$ is calculated by the standard deviation. In j^{th} box we have:

$$V(j) = \sqrt{\frac{\sum_{(j-1)\cdots\lambda+1}^{j\lambda} (W(i) - \overline{W}(j))^2}{\lambda - 1}}, \tag{1}$$

Where the average is given by:

$$\overline{W}_j = \frac{\sum_{(j-1)\cdots\lambda+1}^{j\lambda} (W(i))}{\lambda} \tag{2}$$

This creates the variance fluctuation subseries $V(j) = V1, V2, \cdots, Vm$, with 3000 samples.

[2] Gtzan Genre Collection is a database widely used in musical information retrieval research. It was proposed by 8 and is available at `http://marsyasweb.appspot.com/download/data_sets`

3.3 Transforming Variance Fluctuations in Graphs

Each variance fluctuation point $V(j)$, with $j = 1 \cdots 3000$, is considered to be a vertex of the network. To apply the visibility criterion in the series, we will consider each point of $V(j)$ as an ordered pair (x_j, V_j), where xj is the point position in the series. Two vertex (x_a, V_a) and (x_b, V_b) are connected if there is a point (x_c, V_c) between the m such that:

$$\frac{V_b - V_c}{x_b - x_c} > \frac{V_b - V_a}{x_b - x_a} \tag{3}$$

Equation 3 proposed for [13], provides a comparison between the α_{bc} slope (left side of equation) and α_{ba} slope (right side of equation). Whenever $\alpha_{bc} > \alpha_{ba}$ there is visibility between Va and Vb, and their corresponding nodes are connected in the graph. Otherwise, they do not constitute an edge in the graph. After the equation 3 is applied to all points of the series, following the order $j = 1...300$, we have the visibility of each point of a subset $V(j)$ mapped in a graph $(V(m), V(n)$. This means that, from this stage, each song is represented by a visibility graph.

3.4 Modularity

After mapping the variance fluctuations into visibility graph, the modularity is calculated using the Lovain Method [3], based on GEPHI [3] framework for comminity detection. This algorithm brings a fast unfolding approach for the fundamental modularity defined for [17], whose equation is

$$Q = \frac{1}{2m} \sum_{(i,j)} \left(A_{ij} - \frac{k_i - k_j}{2m} \right) \delta(c_i, c_j) \tag{4}$$

Were i and j are nodes of the network; A_{ij} represents the number of edges between i and j; k_i and k_j are the sum of the the edges attached to i and j; m is the sum of all edges in the graph; c_i and c_j are the communities of the nodes; and (c_i, c_j) is a Kronecker delta function 0 for $c_i = c_j$ and 1 for $c_i \neq c_j$; where c_i and c_j are the communities of the nodes.

To maximize the modularity efficiently, Louvain method proposes a method which uses two stages in iterative repetitions: (1) each node is attributed to their own community. So the change of modularity is calculated for each node i, removing this node from its own community C and moving it to the community of each neighbor i. This value can be easily calculated by:

$$\Delta Q = \left[\frac{\sum_{in} + k_{i,in}}{2m} - \left(\frac{\sum_{tot} + k_i}{2m} \right)^2 \right] - \left[\frac{\sum_{in}}{2m} - \left(\frac{\sum_{tot}}{2m} \right)^2 - \left(\frac{k_i}{2m} \right)^2 \right] \tag{5}$$

[3] GEPHI is a free and open-source software that performs visualization and operation of all types of graphs and networks. Available in `https://gephi.org/`.

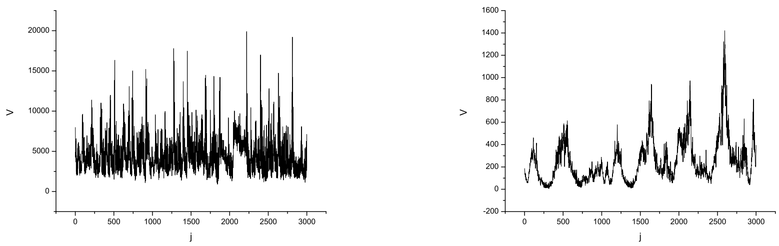
Where $\sum_{in}$ is the sum of the links inside C; $\sum_{tot}$ is the sum of the links incident to nodes in C; $k_{i,in}$ is the sum of the links incident to node i; m is the sum of the links from i to nodes in C and m is the sum of the weights of all the links in the network. In the second stage the nodes belonging to the same community are united, and then it constructed a new network where the nodes are small communities. These steps are repeated until the maximum modularity is achieved and a community hierarchy is produced.

Since the calculation of modularity depends on a random argument, the algorithm each time will return different results. With the tested networks there was very little variation in these results, therefore we considered to all networks a randomly selected result.

4 Experimental results

4.1 Variance fluctuations

Sixty variance fluctuation series were calculated by reducing the original signal, approximately 330,000 points for 3000 points. Figure 1 illustrates the variance fluctuation series of two audio samples. The first represents the Percussive group and the second the Symphonic group. In Figure 1(a) we have a numerical series generated from a song with a strong beat of drums, used in traditional Brazilian rhythm "maracatu", and Figure 1(b) the portion of a Mozart symphony performed by string section of an orchestra, without percussion instruments.



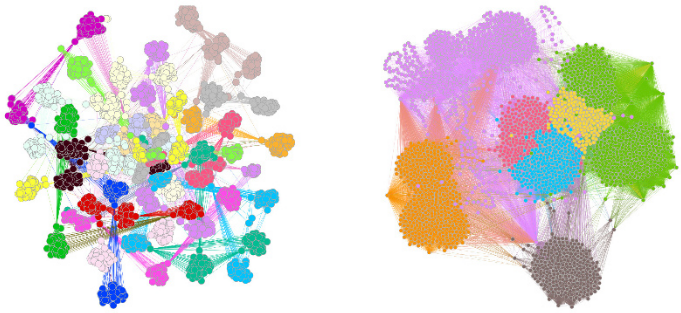
(a) (b)

Fig. 1: Variance Fluctuations series of audio signals corresponding to the files: (a) Etnia by Chico Science & Zoombie Nation; (b) Mozart -Symphony 39 in E flat Major, K 543. Source: Author

We can notice by visual inspection that the first series is denser than the second, with less space between "peaks" and "valleys". We can, even without numerical proof, intuit that these different geometric configurations are associated with the peculiar rhythmic activities to their audio signals. The following results present quantitative basis for characterizing these differences.

4.2 Visibility networks generated from variance fluctuation

We mapped Sixty networks, each with 3000 nodes. The networks are grouped into two types: Symphonic Networks and Percussive Networks. Figure 2 shows two networks, representing respectively the Percussive and Symphonic Networks. The first (Figure 2a) is a mapping of a audio from the 1980s - So Many Men, so Little Time - played by the Canadian singer Miquel Brown. The second (Figure 2b) is a network generated by the mapping of the audio Animé Et Três Décidé - String quartet composed by Claude Debussi, performed by Julliard String Quartet. In these two representations, the modularity classes appear in different colors, indicating the communities formed by each network. Sections 4.3 and 4.4 will present overall results that will give subsidy to infer about trends presented by each group, based on the magnitude of the difference between the amounts of communities formed by the two types of networks.



(a) Percussive Network (Q=0.91 e Nc=24)

(b) Symphonic Network (Q=0.66 e Nc=7)

Fig. 2: Visibility Networks of the variance fluctuations of two audio signals. The colors represent the modularity classes of each network. Source: Author.

The average number of edges of 30 Symphonic and Percussive Networks are, respectively, 60254 ± 10925 and 23827 ± 2899. The results show a significant difference between the mean values of edges generated between the two types of networks. Taking into account that the number of nodes in visibility graphs depends on the visibility of their points in the series. We can infer that, in mapping a set, the greater the number of nodes generated, the higher the visibility of the series. This indicates that, on average, the series that generated the Symphonic networks have greater visibility than the generating series of Percussive Networks.

4.3 Modularity

The results of modularity (Q) of the 60 visibility networks are shown in Figure 3. The networks of each group are indicated with the numbers 1 to 30. We note that all the Q values for Percussive Network ($< Q >= 0.81 \pm 0.08$) are higher than the values calculated for Sympohonic Network ($< Q >= 0.54 \pm 0.13$). The extreme values of modularity are 0.91 for the visibility network of the song Get Up played by the british african-pop band Osibisa, and 0.14 for the network of the Symphony 39 in E flat Major - k 543 composed by Mozart. The Symphonic Networks showed a set of less compact modularity values in the average, with a 12% deviation against the 8% of the Percussive Networks, even so, the average of the two groups showed significant differences with a confidence of 95%, according to the Bonholm test. At this point we can infer, based on the arguments presented in section 4.2, and also on the Q values calculated, that there is an inverse relationship between visibility and modularity.

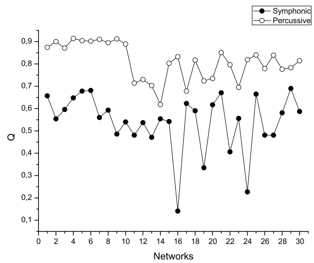


Fig. 3: Modularity of 30 Symphonic (black dots) and Percussive (white dots) Visibility Networks. Source: author.

4.4 Number of communities

Each Q value calculated in section 4.3 is associated with a number of communities (Nc) of the network. Figure 4 shows the Nc values calculated for each Q (Figure 3). Globally the amount of network communities follow the same feature found in the calculation of modularity: exists a very clear distinction between the two classes, where the Percussive Networks outweigh the Symphonic networks for most Nc values. The average values obtained were $< Nc >= 16.5 \pm 4.4$ for Percussive, and $< Nc >= 8.8 \pm 2.2$ for Symphonic networks. Looking locally we can see that in addition to the distinction into two groups, Nc values of Percussive Network can serve as a parameter for stratification within the group, in order, for example, the great distance of the first nine networks Nc values (white dots) to the rest.

4.5 Influence of the randomness in the results

In this section we present the results of a study made about the influence of the randomness factor in the calculation of modularity. As discussed in section 3.4, the calculation of modularity is made based on the comparison of information given for

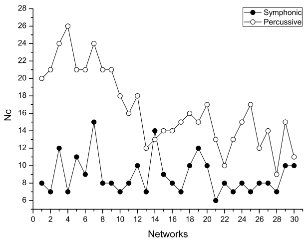


Fig. 4: Number of communities (Nc) of 30 Symphonic (black dots) and Percussive (white dots) of Visibility Networks of audio signal variance fluctuations. Source: author.

edges that exist on the network and edges made randomly. Each time the algorithm is applied, we obtain a value for the modularity and the number of communities. Table 2 shows ten takes from the calculation of modularity and the number of communities for one Percussive Network. In this table we can see that in some cases, the algorithm estimates the same modularity for different Nc values (Takes 2 and 3), and the same number of communities for different Q values (Takes 5 and 9). This shows that, due to the random factor, there is no modularity value associated with a unique modularity class arrangement. In Table 1 we have $< Q >= 0.8420 \pm 0.0022$ and $< Nc >= 15.90 \pm 0.99$ for ten takes. Increasing the number of repetitions to 80 takes we have $< Q >= 0.8422 \pm 0.0024$ and $< Nc >= 16.15 \pm 0.80$. Comparing the results obtained for the two tests, it is clear that the means and variances of Q and Nc do not change significantly with increasing the number of takes, and that for this network there is a great chance that if we choose one of ten or eighty attempts, we find a value of Q and Nc very close to the same average value.

Now we will show the results that investigate the overall impact of the random factor in the calculation of the modularity. We calculate ten repetitions of the Nc of twenty networks (Figure 5). n the x-axes, the networks S1 to S10 (white boxes) are Symphonic Networks, and networks P1 to P10 (dashed boxes) are Percussive Networks. We can see that the overall behavior does not change with the recalculation for each network.

4.6 About sample rate changes

In order to investigate the impact of sample rate changes in results of network parameters we calculate average degree ($< k >$), density (Δ), modularity (Q), number of communities (Nc), diameter (D), average path length (L), clustering coefficient (C), and time processing (TP), of the visibility networks Percussive 11 (samba) and Symphonic 1 (oboé concert), using three sampling rates (SR): 11025 Hz, 22050 Hz and 44100 Hz. We use the framework Gephi 0.9.0 to calculate all parameters.

The results in Table 2 show that SR changes do not bring significant differences in the final statistics, neither alter the trends found in the comparative study between the two musical groups. The computational processing time recorded for this experiment

Table 1: Calculation of the modularity of the network "A walk in the free world" written by Chico Science and the Zombie Nation with ten repetitions. (Source: author).

Take	Modularity (Q)	Communities (Nc)
1	0.844	16
2	0.844	17
3	0.844	16
4	0.839	17
5	0.839	15
6	0.843	16
7	0.842	17
8	0.845	14
9	0.843	15
10	0.844	16

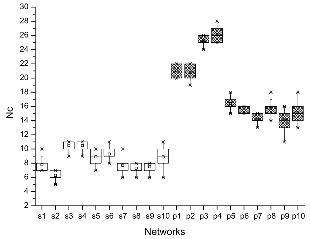


Fig. 5: Number of communities (Nc) of ten Symphonic Visibility Networks (white boxes) and ten Percussive Visibility Networks (dashed boxes), with ten calculations each. (Source: author).

had the decisive influence of the diameter and the average path length calculation. To process only these two parameters, the Gephi spent more than 90% of the total time. If the calculation of these parameters for many networks is required, the rate of 22 and 44kHz are not recommended. To calculate only Q and Nc the Gephi took around 1 sec for each SR.

4.7 Looking closely at some Percussive and Symphonic Networks

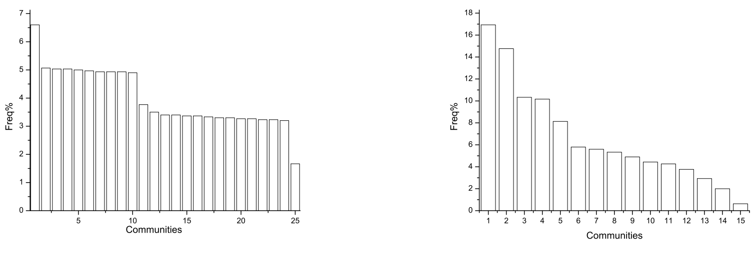
Observing the Figure 3 we can see that some points stood out from the rest of the group because they have reached discrepant or extreme values. Below we will discuss the possible causes of this behavior, putting together musical and statistics similarities.

- *Networks P1 to P10* - They achieved greater magnitude and shorter variance in modularity ($< Q >= 0.897 \pm 0.015$) compared to P11-P30 ($< Q >= 0.767 \pm$

Table 2: Network parameters of networks Percussive 11 (P11) and Symphonic 1 (S1) for 3 different sample rates used in his respective audio samples before network mapping.

Network	SR (hz)	V	E	$<k>$	Δ	Q	Nc	D	L	C	TP (sec)
P11	11025	3000	18866	12.58	0.004	0.857	17	7	3.85	0.837	23
	22050	6000	41500	13.85	0.002	0.875	19	10	3.745	0.845	160
	44100	12000	83179	13.87	0.001	0.918	27	9	4.296	0.849	734
S1	11025	3000	65115	44.74	0.015	0.657	8	4	2	0.860	90
	22050	6000	152262	50.754	0.008	0.714	10	4	1.998	0.878	61
	44100	12000	299272	49.895	0.004	0.682	8	5	2	0.902	3180

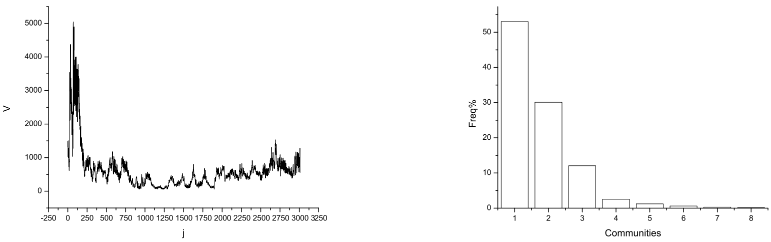
0.064)) and Symphonic networks (Section 4.3). Looking at the distributions of vertices per community, of all networks, we observed higher homogeneity in P1-P10 distributions. This contributed to these networks have obtained greater modularity than the others. Fig 6 shows the distribution of vertices per community of P6, representing P1-P10 networks, and P27, representing the others percussive networks. Comparing the two distributions we can notice greater homogeneity in P6, which reached modularity 0.902, while P27, with less homogeneity got Q=0.839. Musically the P1-P10 networks represent songs of the eighties, which is characterized by danceable groove on every song, dominated by the constant pulse of bass and drums without much dynamics variation. We can speculate that this ”musical homogeneity” may have strongly influenced the statistical uniqueness that made these networks stood out from all others.



(a) (b)

Fig. 6: Distribution of vertices per community of the networks:(a) Percussion 4 - Disco Music, and (b) Percussion 27 - Mangue Beat. Source: Author.

- *Networks S16, S19, S24* - These networks draw attention by having modularity with very low values (0.141, 0.227 and 0.355). Musically, the audio excerpts associated with these networks also have a common feature. In all of them there is a sudden change of dynamics, strongly influenced by the presence or absence of timpani [4]. It created a particular topology in the variance fluctuations of these audio signals, with great ”valleys” followed by high ”peaks”, favoring visibility graphs with big hubs, and cluster distributions with very low amount of nodes in some communities. In consequence, they achieved lower modularity values than the others symphonic networks. Figure 7 (a) shows the variance fluctuations of the audio track Symphonic 16. We can note in Figure 7 (b) that five communities have less than 5% of vertices, while only one community have about 50% of them. The modularity maximization algorithm was not able to merge these small communities into larger communities. This prevented the Q value to stay a bit higher. Anyway the lower Q values found in these three networks, helped to distinguish the particular audio musical behavior that these networks are topologically representing.



(a) (b)

Fig. 7: (a) Variance fluctuations of the track S16 - Symphony 39 in E flat Major, K 543, Mozart. During the first two seconds (j = 200 to 3000), the whole orchestra, including timpani, play a part in fortissimo, and thereafter comes off the timpani, and remain the strings and woods gently touching; (b) Relative frequency of vertices by community of the S16 visibility network. Source: Author.

5 Conclusion and future work

In this article we mapped variance fluctuations of sixty musical audio files into visibility graphs, and through the modularity and the number of communities of each network, we measured the level of dynamics changes influenced by percussive activity of each audio content. We concluded that modularity and number of communities of complex networks has produced useful information for categorization

[4] A set of two or three large drums (called kettledrums) that are played by one performer in an orchestra `http://www.merriam-webster.com/dictionary/timpani`.

into two groups, where audio samples with musical affinities were gathered within the same group according to its high or low percussive activity. Although in this study we have explored the feature extraction with only two categories, the algorithm showed potential for categorizing by more than two labels. Other investigations are in progress in which some network features are performing an audio music hierarchy according to the taxonomy of some musical genres, with a large number of files. To better understand the level of contribution that this algorithm can give to the music information retrieval field, we will conduct an experiment comparing the parameters extracted from the variance visibility networks with rhythm-based tools most used in the literature. Another important issue which is worth be discussed in future work is the evaluation of Pajek adjustment indices (Cramer's V, Rajski and Adjusted Rand Index) in front of the parameters adopted by Gephi, and its influence on the extraction of features proposed by the visibility descriptor of variance fluctuations.

References

[1] Andjelković, M., Gupte, N., Tadić, B.: Hidden geometry of traffic jamming. Physical Review E **91**(5), 052,817 (2015)

[2] Bergstra, J., Casagrande, N., Eck, D.: Two algorithms for timbre and rhythm-based multiresolution audio classification. In: Proceedings of ISMIR (2005)

[3] Blondel, V.D., Guillaume, J.L., Lambiotte, R., Lefebvre, E.: Fast unfolding of communities in large networks. Journal of statistical mechanics: theory and experiment **2008**(10), P10,008 (2008)

[4] Buldú, J.M., Cano, P., Koppenberger, M., Almendral, J.A., Boccaletti, S.: The complex network of musical tastes. New Journal of Physics **9**(6), 172 (2007)

[5] Correa, D.C., Saito, J.H., da F Costa, L.: Musical genres: beating to the rhythms of different drums. New Journal of Physics **12**(5), 053,030 (2010)

[6] Costa, Y.M., Oliveira, L., Koerich, A.L., Gouyon, F., Martins, J.: Music genre classification using lbp textural features. Signal Processing **92**(11), 2723–2737 (2012)

[7] Eronen, A.: Signal processing methods for audio classification and music content analysis. Tampereen teknillinen yliopisto. Julkaisu-Tampere University of Technology. Publication; 817 (2009)

[8] Ezzaidi, H., Rouat, J.: Automatic musical genre classification using divergence and average information measures. World Academy of Science, Engineering and Technology **15** (2006)

[9] Goulart, A.J.H.: Classificação automática de gênero musical baseada em entropia e fractais. Ph.D. thesis, Universidade de São Paulo

[10] Guaus, E., et al.: Audio content processing for automatic music genre classification: descriptors, databases, and classifiers (2009)

[11] Jacobson, K., Sandler, M.B., Fields, B.: Using audio analysis and network structure to identify communities in on-line social networks of artists. In: ISMIR, pp. 269–274 (2008)

[12] Jennings, H.D., Ivanov, P.C., Martins, A.d.M., da Silva, P., Viswanathan, G.: Variance fluctuations in nonstationary time series: a comparative study of music genres. Physica A: Statistical Mechanics and its Applications **336**(3), 585–594 (2004)

[13] Lacasa, L., Luque, B., Ballesteros, F., Luque, J., Nuno, J.C.: From time series to complex networks: The visibility graph. Proceedings of the National Academy of Sciences **105**(13), 4972–4975 (2008)

[14] Lacasa, L., Luque, B., Luque, J., Nuno, J.C.: The visibility graph: A new method for estimating the hurst exponent of fractional brownian motion. EPL (Europhysics Letters) **86**(3), 30,001 (2009)
[15] Lacasa, L., Toral, R.: Description of stochastic and chaotic series using visibility graphs. Physical Review E **82**(3), 036,120 (2010)
[16] Melo, D.F.P.: Análise de flutuações de variância em sinais de áudio agrupados por gênero musical. Proceeding Series of the Brazilian Society of Computational and Applied Mathematics **1**(1) (2013)
[17] Newman, M.E.: Analysis of weighted networks. Physical review E **70**(5), 056,131 (2004)
[18] Nunez, A., Lacasa, L., Valero, E., Gómez, J.P., Luque, B.: Detecting series periodicity with horizontal visibility graphs. International Journal of Bifurcation and Chaos **22**(07), 1250,160 (2012)
[19] Pampalk, E., Rauber, A., Merkl, D.: Content-based organization and visualization of music archives. In: Proceedings of the tenth ACM international conference on Multimedia, pp. 570–579. ACM (2002)
[20] Panagakis, Y., Kotropoulos, C., Arce, G.R.: Music genre classification via sparse representations of auditory temporal modulations. In: Signal Processing Conference, 2009 17th European, pp. 1–5. IEEE (2009)
[21] Park, D., Bae, A., Schich, M., Park, J.: Topology and evolution of the network of western classical music composers. EPJ Data Science **4**(1), 1 (2015)
[22] Schedl, M., Gómez, E., Urbano, J., et al.: Music information retrieval: Recent developments and applications. Now Publ. (2014)
[23] Silla Jr, C.N., Kaestner, C.A., Koerich, A.L.: Automatic genre classification of latin music using ensemble of classifiers. In: Proc. of the 33rd Intigrated Software and Hardware Seminar, pp. 47–53 (2006)
[24] Stephen, M., Gu, C., Yang, H.: Visibility graph based time series analysis. PloS one **10**(11), e0143,015 (2015)
[25] Tse, C., Liu, X., Small, M.: Analyzing and composing music with complex networks: finding structures in bach, chopin and mozart (2008)
[26] Tzanetakis, G., Cook, P.: Musical genre classification of audio signals. IEEE Transactions on speech and audio processing **10**(5), 293–302 (2002)