Acoustic Analysis of Musical Timbre of Wooden Aerophones

Yubiry GONZALEZ
Center of Mathematics, Computer Science and Cognition (CMCC) Federal University of ABC, Av. Dos Estados, 5001, Bangú Santo André, 09210-580, SP, BRAZIL, Yubiry.gonzalez.17@gmail.com

Ronaldo C. PRATI
Center of Mathematics, Computer Science and Cognition (CMCC) Federal University of ABC, Av. Dos Estados, 5001, Bangú Santo André, 09210-580, SP, BRAZIL, ronaldo.prati@ufabc.edu.br

Abstract: - The characterization of the musical timbre, which allows the quantitative evaluation of audios, is still an open-ended research topic. This paper evaluates a set of dimensionless descriptors for studying musical timbre in monophonic recordings of woodwind instruments from the TinySOl audio library, considering the region of frequencies common to all instruments in their three dynamic levels (pianissimo, mezzo-forte, and fortissimo). These descriptors are calculated using the spectra obtained from the Fast Fourier Transform (FFT) using the Python programming language. From the analysis of the distribution of the coefficients, it was possible to verify that the Affinity coefficient (A) allows discrimination in all octaves of musical sounds. The analysis of the data through the empirical distribution of the coefficients shows that the timbral variations due to the dynamics are reflected through the coefficients Sharpness (S) and Mean Affinity (MA). The coefficients are examined using the Principal Component Analysis (PCA), and it was observed that the variability in the distribution of the first principal component is mainly due to the Sharpness (S) and Mean Contrast (MC) coefficients (~55%), and in ~43% by the Affinity coefficient. Similarly, the variation in the second principal component is due to 62 % of the MC coefficient and 49% due to MA. It is concluded that the proposed descriptors are sufficient to differentiate the aerophones studied by octaves and musical dynamics.

Keywords: - Musical timbre, Wood Aerophones, Woodwind instruments, PCA, TinySol

1. INTRODUCTION

Timbre is a general property that allows distinguishing different sounds with the same duration, pitch, and intensity [1]. Timbre is associated with the source of the sound. In musical instruments, it is common to make an analogy with "color" of the sound. Its quantification as a measurable magnitude is an open-ended topic of research, even in the case of monophonic sounds [2-3]. The analysis of musical timbre can be approached from two complementary perspectives: firstly, the one that is related to the psychophysical perception of sound by the listener who discriminates and identifies the source of the sound; second, focused on the acoustics that is related to the description, composition, and distribution of harmonics and partial frequencies (overtones) that accompany a given sound.

This work adopts the second approach, following the original idea of the physicist Georg Simon Ohm (1789-1854) according to which "the difference in timbre of the different sounds comes only from the presence of harmonics and their relative intensity" [4]. The advent of digital technology for the recording and reproduction of musical sounds shows that the collection of frequencies and amplitudes of the spectral decomposition, by means of the Fourier Transform (FFT) contains sufficient timbral elements that allow the univocal reproduction of the generated sounds for musical instruments.

The psychoacoustic aspects of timbre are important for the complete perception of musical sound. However, for the analysis of monophonic sounds (the focus of this paper), the basic characteristics of timbre variations due to changes in the octaves of the musical scale, the type of musical instrument, and intensity variations must also be considered. All these variations are manifested in the FFTs of the digitized sound, independently of the environmental and stimulus-response variations perceived by the listener. For a detailed approach that considers magnitudes and psychoacoustic techniques applied to musical timbre, see Caetano et al. and references there in [3].

An important research question is how to quantify or express musical timbre variations resulting from changes in the octaves of the musical scale, the type of musical instrument, and the intensity level from the

Fourier spectrum? What elements of the timbral variations can be extracted from an analysis based on the FFT of the audio records? To address these questions, the musical timbre will be analyzed from a set of woodwind aerophones audio records, evaluating the change in the timbral coefficients proposed by Gonzalez and Prati [5]. If these sets of coefficients characterize the timbre properties of the FFTs of the musical sounds, then the following question is valid: How the changes in each particular coefficient for each specific timbral variation are distributed?

Aerophones, according to the Sachs-Hornbostel organological classification [6], are a family of musical instruments characterized by the use of air as a vibrating material. Aerophones are made up of tubes or ducts inside which a longitudinal wave of compressions and rarefactions of the air contained in it propagates, forming standing waves that generate the sound. These standing waves have zones of zero vibration and maximum pressure (nodes) and zones of maximum vibration and zero pressure (tops or bellies). In the gas column, sinusoidal pressure variations are out of phase by 90 degrees with respect to particle displacement variations. The lowest note of these instruments is achieved by covering all its holes (by means of keys or fingering) so that the column of air inside it has a maximum length. The column is shortened by uncovering the holes in succession starting at the open end.

Metal aerophones (more precisely brass: trumpets, tubas and trombones, among others) are classified according to the extraction forms of the vibrating material (of the air in the aerophones), where the air is excited by the vibration of the performer's lip, and in wooden aerophones (whose name alludes to their ancient construction) by the vibration of a reed (double or simple) or against an edge (bevel mouthpiece) [6].

This study is limited to Western orchestral music, therefore the sample of wooden aerophones includes the most common aerophones in symphonic orchestration: Bassoon (Bn), Oboe (Ob), Transverse Flute (Fl), and Clarinet (ClBb). The selection of the family of wooden aerophones is justified as they have the greatest timbre diversity and more melodic possibilities working as a solo instrument within a symphony orchestra, but also presenting well-differentiated timbre characteristics when they work as a set of instruments [7]. On the other hand, the set of monophonic sounds in Western orchestral music is a finite number of 96 different tones, which constitute the tempered musical scale, and allows, in principle, to delimit the problem of the acoustic characterization of their timbre. Then, the acoustic analysis is limited to the common tessitura of the mentioned aerophones: B3 to D#5 of the tempered musical scale, that is, between 246.9 Hz and 622.3 Hz.

From the perspective of acoustics, the initial and boundary conditions determine the vibration of a sound body (musical instrument) when producing a sound. This is, in principle, made up of a finite set of frequencies, where the most important one is the fundamental and the others can be concordant (harmonics) with it, or they are discordant (overtones or partials).

Harmonics are modes of vibration characterized by tones that have a natural scale relationship with the fundamental frequency or pitch. On the other hand, overtones do not have this characteristic natural scaling. The number of harmonics and overtones, their relative intensities, and the distribution of sound energy among them determines the specificities of timbre that characterize the sound body and are specific to each musical instrument [8,9]. In acoustics, woodwind instruments are considered resonant tubes where the gas, and not the cylinder that delimits it, is the vibrating body. They all have an open end where reflection occurs, generating the standing wave.

The mouthpiece drive, except in the clarinet, is open-end, so a node for the fundamental sound is formed in the middle of the two ends. Furthermore, the frequency of the nth harmonic follows the relationship $f_n = nf_0 = \frac{nC_s}{2L}$, where C_s is the speed of sound in the air column and L is the length of the tube. In tubes with a closed end there will be a node at the closed end (embouchure) and an antinode at the free end, so the fundamental frequency will have a node twice as long as the open case, and consequently, only even harmonics are present.

The Clarinet behaves as if it were a closed resonant tube: a node is formed near the embouchure (reed) and an antinode is formed a little further from the output end, completely suppressing the second harmonic and attenuating the fourth, although having a single reed embouchure and the semi-conical construction of the clarinet tube (with conical parts) causes some even (6th and 8th) harmonics to be registered [10]. The shape of the tube establishes limitations to the propagation of the plane wave within it, the most used geometries being the straight cylinder (vertical flutes and Clarinet) and conical cylinders such as in the Oboe, Bassoon, Transverse Flute, and Saxophone. The embouchures also differ in the wooden aerophones: double-reed (Bassoon and Oboe), simple (Saxophone and Clarinet), and direct (bevel-shaped like the Transverse Flute).

Therefore, the timbral characteristics of the aerophones will be delimited or distinguished by the particular characteristics of the embouchure, the shape, and the length of the resonant tube. It is expected that the timbre variety of the aerophones is linked to the propagation of the longitudinal wave in the gas column, subject to the specificities of the geometry and the form of vibration generated in the embouchure.

It is intended to study the common characteristics, as an organophonic group, of woodwind musical instruments and their particularities that distinguish the individual timbre of each one, from the perspective of physical acoustics.

For this, the spectra obtained from the audio records will be analyzed in their common tessitura, in monophonic sounds of the tempered musical scale, comparing the Fourier Transforms (FFT) through previously selected timbre quantifiers or acoustic descriptors (section 2). The results of the comparison of the characteristic spectral signatures of the aerophones, varying the dynamics (*pianissimo pp*, *mezzo-forte mf*, and *fortissimo ff*) are shown in section 3. In section 4 the empirical distribution of the coefficients is estimated, in addition to a PCA analysis to assess the contribution of the coefficients in the differentiation of woodwind instruments by octaves and dynamics. In the last section, the conclusions and expectations for the general understanding of the musical timbre are presented.

2. METHODOLOGY

2.1. Specifics of audio recordings

In this study, a subset of the open-source TinySOL [11] sound library is used. This audio library contains recordings of individual sounds in the WAV audio format with minimal loss, sampled at 44.1 kHz on a single channel (mono) at 16-bit depth. TinySOL's audio recordings consist only of sounds played in the so-called "ordinary" style and in the absence of mute. This library has been used in projects related to timbre perception, computer-assisted composition, and intelligent systems for musical orchestration [12-15]. In general, it is a library that can be used as a data set to train and/or evaluate musical information retrieval (MIR) systems, for tasks such as instrument recognition, and estimation of fundamental frequencies, among others [11].

2.2 Techniques for obtaining FFT and PCA

In each of the audios from the TinySOL library, the FFT was calculated and not STFT because this work is only analyzing monophonic audio recordings without considering the temporal variation of the sound. Then, the characteristic frequency spectra were obtained for each sound and musical instrument considered in this study. For this, the SciPy Python library module [16-17] was used, and then the "Scipy.signal.find_peaks" function was used to calculate the local maximums with the parameters height = 0.01 (1% of the maximum height) and distance = 50 (peaks should be at least 50 points apart from each other). This function takes a 1-D array and finds all local maxima by a simple comparison of neighboring values. From this function, we were able to know the amplitude of each overtone.

From this calculation, the tables of the maximum frequencies, expressed in Hertz (Hz), with relative normalized amplitudes with respect to the maximum amplitude value recorded in each FFT spectrum were constructed.

The calculation of the coefficients, for each instrument, was carried out with the frequency spectrum of the common region of the four studied aerophones and from these data, a table (Pandas dataframe) was built on which the empirical distribution of the coefficients was calculated using Principal Component Analysis (PCA). PCA is a common exploratory data analysis and data reduction technique that transforms data from higher dimensions to lower dimensions while preserving as much information as possible. The original data is projected into the Principal Components (PC), promoting a change of basis of the space. It does so by creating new uncorrelated variables that successively maximize variance.

The PCs are found solving an eigenvalue/eigenvector problem [18]. For this analysis, we used the scikit-learn library [16], a free statistical modeling and data analysis library for Python.

2.3 Timbral coefficients

The FFT spectrum of an audio record is essentially a discrete collection of N frequencies (f_i) and N amplitudes (a_i). On the other hand, there are only 96 possible monophonic musical sounds in the tempered scale of Western orchestral music, spread over eight different octaves. Each of these 96 possible musical sounds has a single fundamental pitch or frequency (f_0).

Figure 1 illustrates the Fourier spectra of the selected aerophones for the musical sound A of the fourth octave (A4, f_0= 440 Hz).

It can be observed that: the centroid or average value of the frequencies (black arrow on the horizontal axis) is far from the fundamental frequency in each case; the fundamental frequency does not always have the maximum amplitude (Oboe), the number N partial frequencies are different in each instrument and the f_i frequencies are not always integer multiples of f_0 (as in the third maximum of the bassoon), the maximums are always decreasing in some cases (Clarinet) or increasing at the beginning (Oboe), and the average values of the amplitudes and partial frequencies are different for each musical instrument.

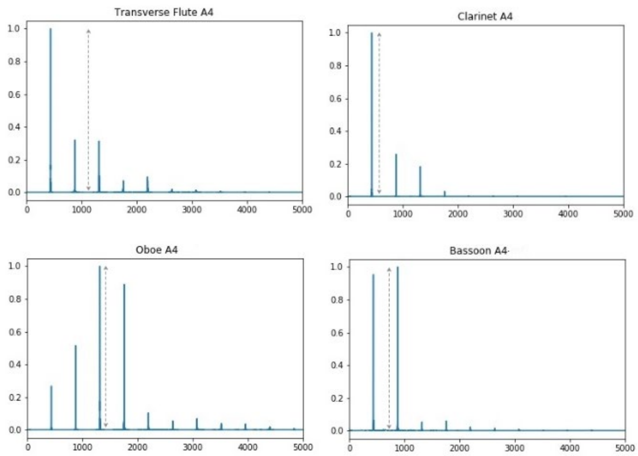

Figure 1. Fourier spectra of sound A4 for the studied aerophones. The dotted line represents the position of the centroid.

Thus, each monophonic sound of a considered musical instrument is characterized by a single FFT that identifies it. Then, the timbre characteristics associated with that particular audio must be contained in the distribution of frequencies and amplitudes of the FFTs, that is, it must be somehow inscribed in the corresponding Fourier spectrum.

Comparing the timbral differences of audio recordings is equivalent to comparing the spectra of FFTs, and this is not a trivial task. This paper proposes a way to quantify the differences in the FFTs through dimensionless coefficients that allow assessing the following three aspects: the fundamental frequency, the shape of the partial frequency distribution (harmonic or not), and the statistics of the distribution.

Table 1 shows the set of six coefficients, proposed in [5], that describe and discriminate the timbral similarities and differences that result from considering these three aspects for the two variables contained in the FFT: frequencies and amplitudes.

Table 1. Timbral Coefficients associated with the FFT of monophonic musical sounds [5].

Coefficient	**Operational definition**	**Description**
Affinity (A)	$A \equiv \frac{1}{f_0}\frac{\sum_{i=1}^{N} a_i f_i}{\sum_{i=1}^{N} a_i}$	Relative measurement of the centroid with respect to the fundamental frequency
Sharpness (S)	$S \equiv \frac{a_0}{\sum_{i=1}^{N} a_i}$	Relative measure of the amplitude of the fundamental frequency. Note that S does not refer to the Zwicker-Sharpness used in psycho-acoustics.
Harmoni-city (H)	$H \equiv \sum_{j=1}^{N}\left(\frac{f_j}{f_0} - \left[\frac{f_j}{f_0}\right]\right)$	Average value of the harmony of the partial frequencies
Monotony (M)	$M \equiv \frac{f_0}{N}\sum_{j=1}^{N}\left(\frac{a_{j+1}-a_j}{f_{j+1}-f_j}\right)$	Deviation from regularity in the distribution of amplitudes with respect to frequencies
Mean affinity (MA)	$MA \equiv \frac{\sum_{i=1}^{N} \lvert f_i - \bar{f} \rvert}{N f_0}$	Mean deviation of the partial frequencies from the average frequency
Mean Contrast (MC)	$MC \equiv \frac{1}{N}\sum_{j=1}^{N} \lvert a_0 - a_j \rvert$	Mean deviation of the partial amplitudes from the amplitude of the fundamental frequency

Note: The square brackets in Harmonicity (H) were defined as the integer division (integer part).

3. RESULTS: ANALYSIS OF THE FFT

The timbre is related to similarities in the sound of several tones of the same instrument. Furthermore, the timbre is related to significant variations between different musical instruments given the same tone and intensity; variations due to the change of octaves for the same instrument; and similar dynamics when it is executed in different modes (*pianissimo*, *mezzoforte*, and *fortissimo*) for the same tone and particular aerophone.

Next, the FFTs of the aerophones of the considered samples are analyzed for each of these aspects, using the timbral coefficients of the FFT of the monophonic audio recordings.

3.1 Tone and Timbre

Given a set of sounds, the timbre of a given musical instrument, a particular aerophone, for example, shall present common characteristics. In this sense, the timbre is transverse to the frequency or tone.

The Affinity (A) timbre coefficient reflects this behavior as shown in Figure 2, where the tonal variation (by frequency) is compared with the aerophones in the common tessitura.

Thus, two aerophones, the Oboe, and the Bassoon are very similar in terms of their construction as conical resonant tubes with open ends and double reed embouchure, and present very different values of affinity among each other, throughout the common tessitura.

The Affinity (A) also decreases in the Flute as the sound becomes higher, while the clarinet increases the value of A in the middle of the scale (note A4 and higher), and its affinity is greater than in the flute for the high sounds and similar to those of this one in the serious sounds.

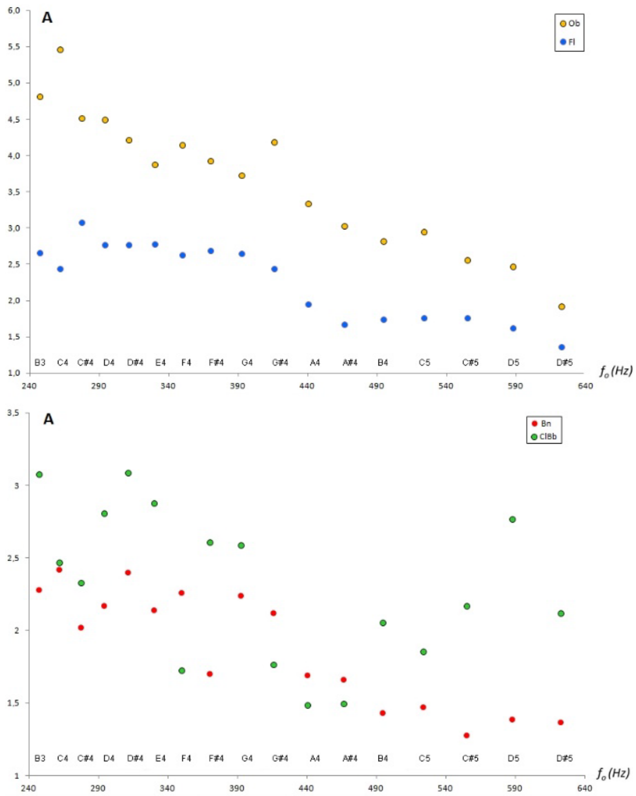

Figure 2. Aerophone Affinity comparison

The timbre coefficient of Sharpness (S) as were defined in Table 1, it`s shown in Figure 3, allows us to distinguish other qualitative timbral aspects of the aerophones. Notice that Sharpness (S) is not the psychoacoustics Zwicker sharpness.

For a given set of sounds (a common region to the considered instruments), the Oboe presents less resolution of the fundamental frequency with respect to the set of overtones of the FFT (lower S value) in relation to the other aerophones.

The Clarinet has higher S values in almost all common tessitura. In the sounds of the Flute and the Oboe, the fundamental frequency becomes more noticeable by having a greater S, proportionally to the sharpness of the tone of the audio record.

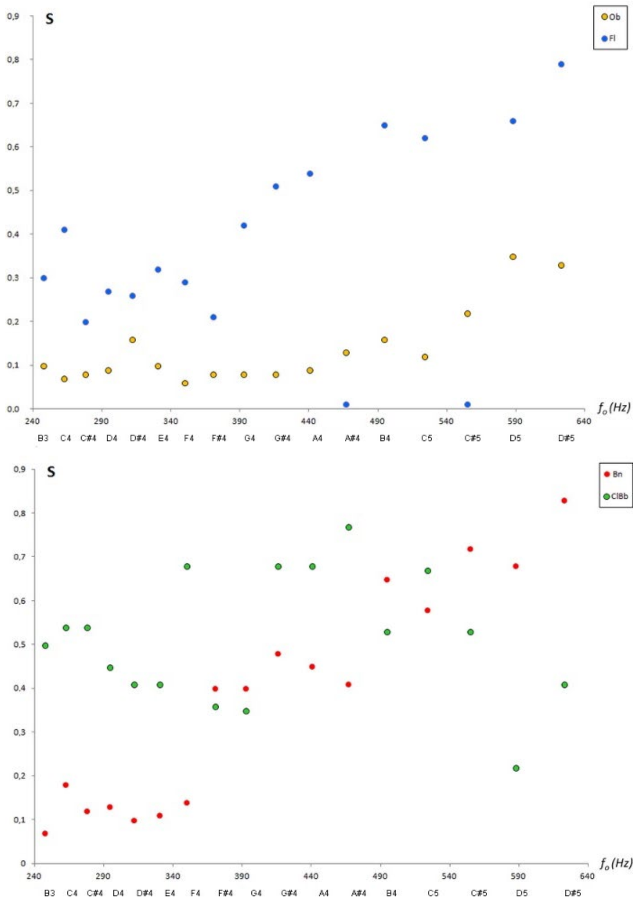

Figure 3. Aerophone Sharpness comparison. Note that S does not refer to the psychoacoustics Zwicker sharpness.

3.2 Spectral signatures

The identification of a sound source in general, and musical instruments in particular, can be done through the analysis of its Fourier spectrum. The set of secondary frequencies that are present in a monophonic sound of a musical instrument objectively characterizes it.

These secondary frequencies (harmonics and overtones) do not always have a multiplicity relationship with the pitch or fundamental frequency, as only the harmonics correspond to integer multiples of the fundamental frequency (f_0).

The set of harmonics depends, among others, on the geometry and boundary conditions of the sound tube of the aerophone [8]. Table 2 compares the harmonics present in the studied aerophones for the common tessitura and dynamic mezzo forte (mf).

Table 2. Set of harmonics depending on the Aerophone and musical sound

Sound	Fl	Ob	ClBb	Bn
B3	1 - 8, 10	1 - 12	1, 3-8, 13	1 – 5
C4	1 – 9	1 - 14, 17	1 - 8*	1 - 8*
C#4	1 -9	1 - 11, 16	1 - 7	1 – 5
D4	1 – 8	1 - 8,1013	1 - 7	1 – 4
D#4	1 - 8	1 -7, 9-14	1 - 7,10	1 – 5
E4	1 - 9*	1 - 13	1 - 9*	1 – 5
F4	1 - 8	1 - 10, 13	1 -5	1 – 6
F#4	1 - 9	1 -12	1, 3 - 7	1 – 4
G4	1 - 8	1 -12	1 - 7	1 - 5, 8, 11
G#4	1 - 8	1 -12	1 - 7	1 - 5, 8, 11
A4	1 - 7	1 - 10	1 - 4*	1 - 4*
A#4	1 - 5	1 - 10	1 - 6	1 – 5
B4	1 - 7	1 - 8	1 - 6*	1 - 6*
C5	1- 6*	1 - 8	1 - 6*	1 – 7
C#5	1 -10	1 - 7	1 - 6	1, 2
D5	1 - 4	1 - 7*	1 - 7*	1 – 3
D#5	1 -3, 5	1 - 5*	1 - 6	1 - 5*

(*) Denotes degeneration: two different instruments with the same set of harmonics.

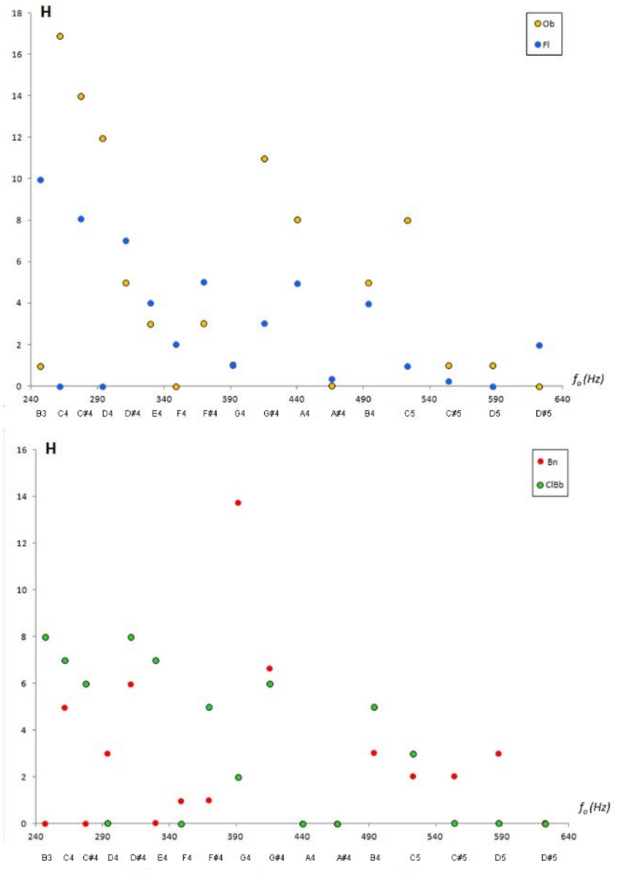

Figure 4. Aerophone Harmonicity comparison

Note that only harmonic frequencies whose amplitude (a_j) is greater than 1% of the amplitude of the fundamental frequency (a_0) were collected, since the sound intensities are proportional to the square of the amplitudes and, consequently, the secondary frequencies would have intensities of the order of 10^{-4} with respect to the fundamental frequency, contributing very little to the formed sound

In general, the presented set of harmonics allows us to distinguish or identify the aerophone in a given monophonic sound. However, there are cases where the sets of harmonics are equal, which has been called "degeneracy".

However, the timbral coefficient H (Figure 4), allows us to assess how harmonic the set of secondary frequencies is or the relative closeness of a certain secondary frequency, compared to its value of integer multiplicity of the fundamental frequency. In addition to the Harmonicity H, the Affinity A of an aerophone (Figure 2) also allows the aerophone to be distinguished, in particular for cases of degeneration. Given a monophonic audio, the aerophone that generated it can be uniquely discriminated by obtaining the set of harmonics present in the FFT and the evaluation of the timbral coefficients.

3.3 Octaves

If the tessitura of each aerophone is considered, it can be observed (Figure 5) that the

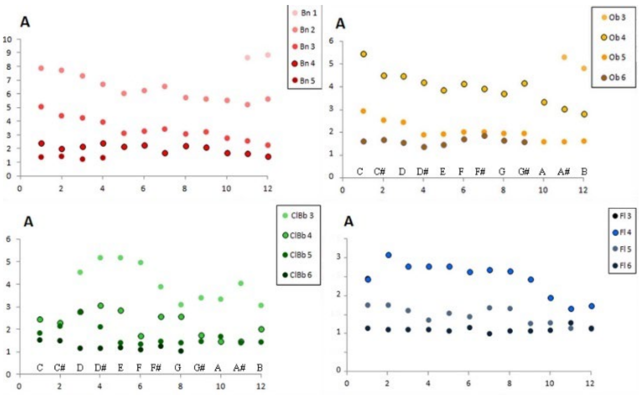

Figure 5. Variation of Affinity in terms of octaves.

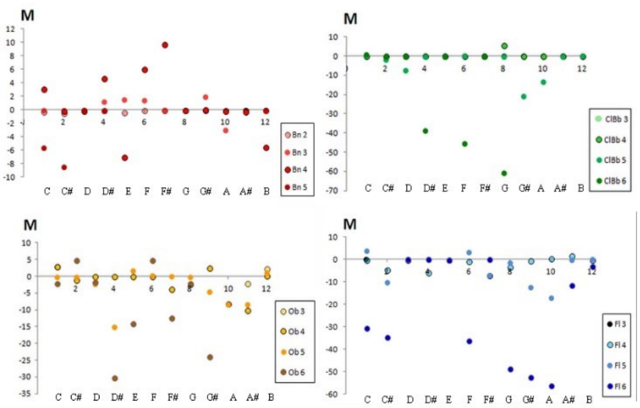

Figure 6. Monotony of the aerophones as a function of the octaves

Affinity (A) decreases monotonically with the increase of the octave for all aerophones, that is, it decreases as the octave is higher. Without loss of generality, for the analysis of the octaves, only the mezzo forte dynamics have been considered, since the graphs corresponding to the pianissimo and fortissimo dynamics behave analogously to Figures 5 and 6.

3.4 Dynamics

The execution modes of monophonic sounds with respect to the average intensity of relative loudness (dynamics) are called mezzo forte (*mf*), pianissimo (*pp*), and fortissimo (*ff*) refer that, in the case of aerophones, the musical interpreter decreases or increases, respectively, the driving pressure of the air flow in the sound tube.

However, in general, depending on the instrument and the performer, such action is accompanied by slight variations in the absolute pitch and, consequently, it is not a mere increase in the level in decibels or physical power (in watts), but results from variations relative to the fundamental frequency of the corresponding tone.

For a particular monophonic register of an aerophone, the variations of the dynamics (*pp*, *mf*, and *ff*) cause a general decrease in the presence of secondary frequencies in the corresponding FFT, suppressing many of them and the harmonics in the pp mode. Similarly, the ff mode is accompanied by an increase in the number of audible secondary frequencies (harmonic or not). Figure 7 shows that the values of Sharpness (S) decrease in the *ff* mode in all aerophones and tessituras, with respect to the recordings in *mf* mode, and these with respect to *pp*.

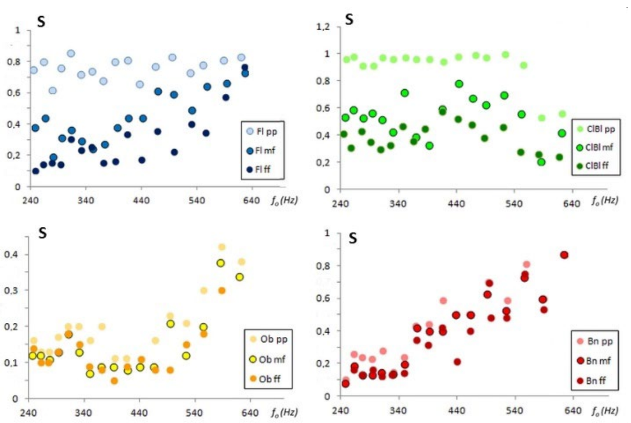

Figure 7. Sharpness (S) as a function of dynamics for the aerophones studied. Note that S is defined in table 1 and is different from the psychoacoustics Zwicker sharpness.

Note the similarity in the variations of (S) for Oboe and Bassoon in accordance with their similar acoustic properties (open conical tube aerophones with double-reed embouchure). Variations in dynamics modify not only the amplitude of the pitch (a_0) but also the amplitude ratio between the fundamental frequency and its partials.

Figure 8 shows how the Mean Affinity varies, which decreases in the dynamics *ff*, with respect to the values of *mf* and increases in the dynamics *ff* with respect to the value *mf* in all the aerophones. Likewise, there is a decrease in the value of MA as the sound becomes less serious, with inverse dependence on the fundamental frequency, due to the very definition of the MA coefficient.

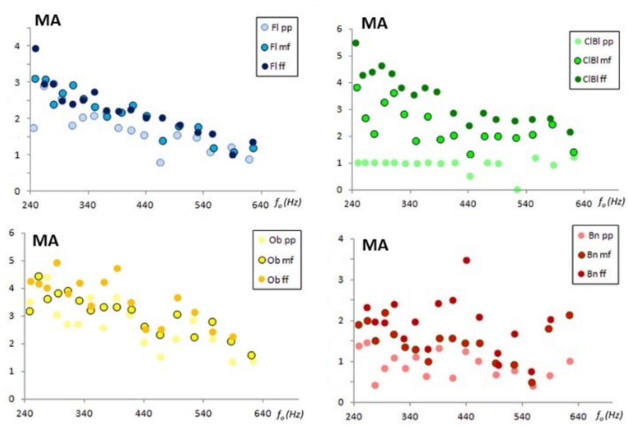

Figure 8. Mean Affinity MA according to the dynamics of the aerophones, common tessitura: B3-D#5.

It should be noted that some deviations observed in particular sounds (A4 fortissimo in the Bassoon or A#4 pianissimo in the Flute, etc.) are due to playing techniques, which go beyond the mere increase in driving force or air pressure in the resonant tube of the aerophone.

4. DISCUSSION.

The empirical distribution of the descriptors can be estimated for the data set of the timbral coefficients of the studied aerophones, considering their complete tessituras.

The results are shown in Figure 9. We observe that in effect the Affinity coefficient (A) allows discrimination in each case the octaves of the musical sounds.

Note that the increase in the octave is proportional to the increase in variance in affinity across all aerophones.

The analysis of the data shows that the timbral variations due to the dynamics are reflected through the coefficients S and MA (Figure 10), in accordance with Figures 7 and 8. The separation of the data is more conspicuous between dynamics *fortissimo* and *pianissimo*.

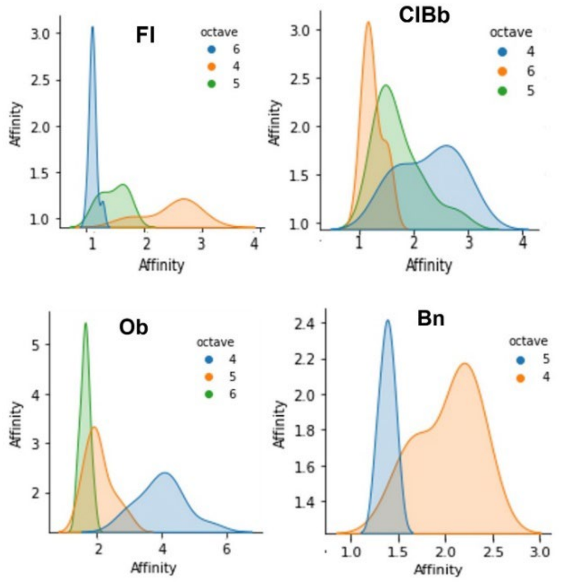

Figure 9. Distribution density analysis considering the Affinity by Octaves, for the complete tessitura of the selected aerophones.

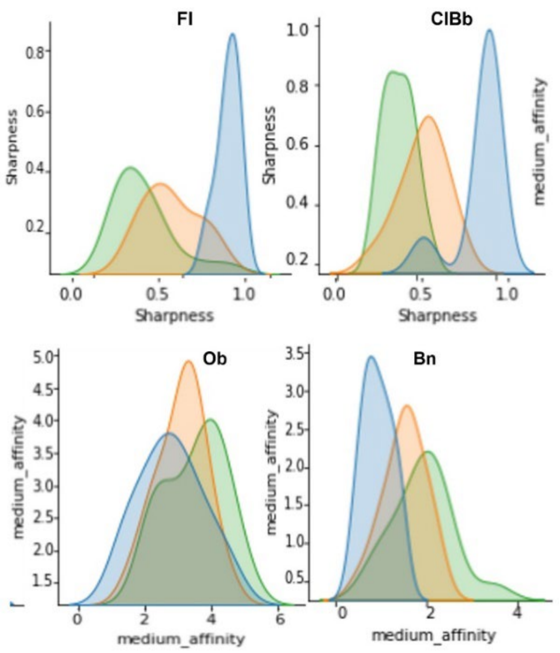

Figure 10. Analysis of variance of the coefficients S (Left. Flute and Clarinet) and MA (Right. Oboe and Bassoon) by dynamics in the common tessitura.

Table 3. Correlation between timbre coefficients and reduced PCA variables

	X	Y
A	0.426	-0.318
S	-0.545	-0.178
H	0.348	-0.488
M	0.128	0.038
MA	0.287	0.487
MC	-0.550	-0.624

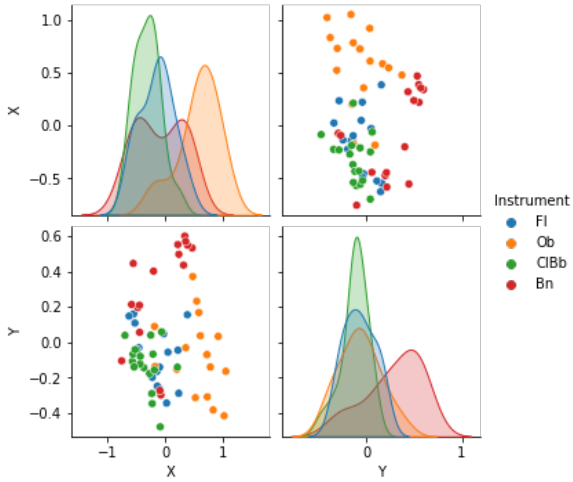

Figure 11. Principal Component Analysis (PCA) in the common tessitura region.

Taking into account the common dynamics of tessitura and mezzo forte, the PCA shows that the timbral differences between the aerophones can be reduced to a two-dimensional space of components X and Y (Figure 11), whose dependence on the proposed timbral coefficients is shown in Table 3.

Note that M contributes little to the distribution of the data and that S and MC contribute similarly to the variable X, as expected since they are variables associated with amplitudes (Table 1), while Affinity (A) and Mean Affinity (MA), both associated with frequencies (Table 1), are proportionally linked in both operational variables X and Y of the Principal Components Analysis (PCA). On the other hand, the PCA shows that 62% of the data is contained in the operational variable X and therefore the variability in the distribution of X is mainly due to the coefficients S and MC (~55%), and in ~43% by the coefficient A. Similarly, the variation in Y is due in 62% to the coefficient MC and in 49% due to MA, in agreement with the results of figures 8 and 10.

5. CONCLUSIONS

It is concluded that the musical timbre in monophonic audio records can be characterized using the FFT through six dimensionless coefficients that we have proposed here. These timbre coefficients incorporate the information about the distribution of harmonics and overtones and their amplitudes in relation to the fundamental frequency. The values of these coefficients allow devising a timbre space to describe the variations due to the type of instrument, the musical note, the octaves, and the dynamics in the selected wooden aerophones (Oboe, Clarinet, Transverse Flute, and Bassoon).

In the common tessitura of the aerophone sample (B3-D#5) the timbre can be distinguished by means of the relative measure of the centroid with respect to the fundamental frequency (Affinity A) varies for each one, being greater in the Oboe and less in the bassoon, despite their similar acoustic properties (aerophones with open conical tubes with a double reed mouthpiece). Also, in the common tessitura, the aerophones are distinguished timbrally through the relative measurement of the amplitude of the fundamental frequency with respect to the set of amplitudes of the partial frequencies: major in the Clarinet and minor in the Oboe. Remember that, although we use the same name, in this paper Sharpness S does not correspond to the Zwicker psychoacoustics definition of sharpness.

The differences in the composition of the harmonics present in each aerophone and the average value in the harmony of the partial frequencies (coefficient H) allow each aerophone to be distinguished timbrally in the common tessitura, even when the set of harmonics is the same for two instruments in a particular sound.

For each instrument, in tessitura, the timbral variations due to the octave change are shown through the coefficient of affinity (A) and monotony (M) in all the aerophones. Affinity decreases monotonically with the increasing octave, and Monotony increases in absolute value as the octave increases.

Timbral variations due to dynamics are reflected through the Relative measurement of the amplitude of the fundamental frequency (S) and the average deviation of the partial amplitudes with respect to the amplitude of the fundamental frequency (MA). The values of S and MA decrease in the *ff* mode in all aerophones and tessituras, with respect to the *mf* dynamics, and those with respect to *pp*.

In the common tessitura, the timbral variations of the set of aerophones can be reduced to two dimensions: X and Y, through PCA, preserving 80 percent of the data variability from the original space. It was observed that the variability in the distribution of X is mainly due to the S and MC coefficients (~55%), and ~43% by the A coefficient. Similarly, the variation in Y is due to 62% to the MC coefficient and 49% due to MA. It is concluded that the proposed timbre descriptors are effective to differentiate the studied aerophones by octaves and differences in musical dynamics.

ACKNOWLEDGMENTS

"This research was supported by Supported by Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES)".

REFERENCES

[1] AMERICAN NATIONAL STANDARDS INSTITUTE, *Psychoacoustic terminology S3:20*. New York, NY: American National Standards Institute, 1973

[2] Antoine A., *An Investigation into the Use of Artificial Intelligence Techniques for the Analysis and Control of Instrumental Timbre and Timbral Combinations*. PhD Thesis University of Plymouth. 2018, http://hdl.handle.net/10026.1/1195.

[3] Caetano M., Saitis C., Siedenburg K., Audio content descriptors of timbre. In Timbre: *Acoustics, Perception, and Cognition*. Springer, 2019.

[4] Pastor-Garcia V., *Estudio y análisis sobre la acústica y organología del Clarinete y su optimización.* PhD Thesis. Universidad Politécnica de Valencia, España. 2005. https://riunet.upv.es/handle/10251/59071.

[5] Gonzalez Y., Prati, R. C., Acoustic Descriptors for Characterization of Musical Timbre Using the Fast Fourier Transform. *Electronics*, Vol.11, No. 9, 2022, pp. 1405

[6] Pérez de Arce J., Gili F., Clasificación Sachs-Hornbostel de instrumentos musicales: una revisión y aplicación desde la perspectiva americana. *Revista Musical Chilena*, Vol. LXVII, No. 219, 2013, pp. 42-80.

[7] Adler S., Hesterman, P., *The study of orchestration* (Vol. 2). New York, NY: WW Norton, 1989.

[8] Rossing T. D., *The Science of Sound* 2nd Ed, Addison-Wesley, 1990.

[9] Sethares W. A., *Tuning, timbre, spectrum, scale*. Springer Science Business Media, 2005.

[10] Muñoz A., La Acústica del Clarinete. *Innovación y experiencias educativas* No. 23, 2009, pp. 1-8. https://archivos.csif.es/archivos/andalucia/ensenanza/revistas/csicsif/revista/pdf/Numero_23/ANGEL_MUNOZ_MUNOZ01.pdf

[11] Cella C. E. et al OrchideaSOL: a dataset of extended instrumental techniques for computer-aided orchestration. *arXiv preprint* arXiv:2007.00763, 2020.

[12] Lostanlen V. et al, Time–frequency scattering accurately models auditory similarities between instrumental playing techniques. EURASIP *Journal on Audio, Speech, and Music Processing*, Vol 2021, No.1, 2021, pp 1-21.

[13] Soden C., *Orchestrational combinations and transformations in operatic and symphonic music.* PhD Thesis. McGill University, 2021. https://escholarship.mcgill.ca/concern/theses/b5644x33p

[14] Cella C.E., Dzwonczyk L, Saldarriaga-Fuertes A., Liu H., Crayencour H. C., A Study on Neural Models for Target-Based Computer-Assisted Musical Orchestration. In *2020 Joint Conference on AI Music Creativity (CSMC+ MuMe)* Stockholm, Sweden, 2020.

[15]. Sridhar S., Lostanlen V. Helicality: An Isomap-based Measure of Octave Equivalence in Audio Data. *arXiv preprint arXiv:2010.00673*. 2020.

[16]. Pedregosa F., Varoquaux G., Gramfort A., et al., Scikit-learn: Machine Learning in Python, *Journal of Machine Learning Research* vol 12, 2011, pp 2825-2830

[17] Virtanen P., Gommers R., Oliphant T.E. et al., SciPy 1.0: fundamental algorithms for scientific computing in Python. *Nature methods*, Vol. 17, No. 3, 2020, pp. 261-272

[18]. Jolliffe I. T., Cadima J., Principal component analysis: a review and recent developments. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, Vol. 374, No. 2065, 2016, pp.20150202.