POLITECNICO DI MILANO

Scuola di Ingegneria dell'Informazione

Master of Science in
Computer Engineering

Feature-based Analysis of the Violin Tone Quality

Candidate

Luca Trapasso
Student Id. number 740700

Thesis Supervisor

Prof. Augusto Sarti

Assistant Supervisor

Dr. Massimiliano Zanoni

Academic Year 2012/2013

POLITECNICO DI MILANO

Scuola di Ingegneria dell'Informazione

Laurea Magistrale in
Ingegneria Informatica

Analisi della Qualità del Violino basata sull'Estrazione di Descrittori Audio

Candidato

Luca Trapasso
Matricola 740700

Relatore

Prof. Augusto Sarti

Correlatore

Dr. Massimiliano Zanoni

Anno Accademico 2012/2013

Feature-based Analysis of the Violin Tone Quality
Master thesis. Politecnico di Milano

This thesis has been typeset by LaTeX and the smcthesis class.

Author's email: trapasso.lu@gmail.com

Dedicated to
my Family and Valentina

Sommario

La classificazione della qualità timbrica di un strumento è una delle sfide di ricerca più ardue e perseguite da anni. Quest'ambito è reso ancora più interessante se ha come protagonista uno strumento affascinante e antico come il violino.

Esso è stato reso famoso dai maestri liutai italiani e la leggenda vuole che il celebre suono prodotto dai loro violini sia determinato anche dalla ricetta delle loro vernici. Tuttora, l'impatto della vernice sulle qualità timbriche e acustiche del violino rimane quasi sconosciuto.

Anche le corde danno un grande contributo alla caratterizzazione del suono del violino. Per questo motivo, capire come queste modificano le qualità timbriche può aiutare i musicisti nella ricerca di un particolare suono.

In questo lavoro quindi, viene esaminata la risposta del violino al variare di alcuni parametri costruttivi. Quindi estraiamo quante più informazioni possibili dal suono prodotto dal violino, correlandole con due parametri costruttivi. Il primo riguarda il processo di verniciatura ed il secondo è rappresentato dalla tensione delle corde.

I precedenti studi effettuati sull'effetto della verniciatura hanno fornito solo alcune considerazioni.

In questo lavoro si propone una soluzione innovativa. Analizziamo le variazioni timbriche introdotte prima dal turapori e dopo dalla vernice, su un violino appositamente costruito. Le analisi vengono eseguite usando delle risposte in frequenza integrate con dei descrittori audio. Allo stesso modo, viene analizzato l'impatto che diversi set di corde hanno sulla qualità timbrica.

Parte dei risultati è stata confrontata con quelli trovati in letteratura a scopo di validazione. In aggiunta, vengono individuati diversi effetti sulle qualità timbriche e acustiche del violino introdotti dal turapori e dalla vernice. Tali effetti possono essere tenuti in considerazione dai liutai durante il lavoro di costruzione del violino.

Infine, è stata determinata una correlazione tra gli effetti timbrici introdotti dalle corde ed il loro valore di tensione. Questi risultati possono essere utilizzati dai musicisti per agevolarli nella scelta delle corde adatte alle loro esigenze.

Abstract

The classification of the tonal quality of an instrument is one of the most difficult and pursued research challenges that lasts for years.This topic becomes even more interesting if it has as protagonist a so fascinating and so old instrument as the violin.

It became famous thanks to the great Italian masters. The legend says that the celebrated sound produced by their violins is also determined by the recipe of their varnishes. Nowadays, the impact of the varnish on both the timbre quality and the acoustics of the violin remains almost unknown.

Strings also provide a great contribute to the characterization of the violin tonal quality. For this reason, understanding how the strings affect the tonal qualities can help the musicians to research a characteristic sound.

To this purposes, in this work we inspect how the violin responds to the changes performed on several manufacturing materials. In order to do this, we extract as much information as possible from the sound produced by the violin, correlating them with two manufacturing parameters. The first concerns the varnishing process and the second is represented by the physical properties of the strings and in particular by the tension.

The previous studies on the effects of varnishing have provided just a few considerations.

In this paper we propose an innovative solution. We analyze the tonal changes introduced firstly by the ground coat and secondly by the varnishing on a brand new violin. The analysis are performed using frequency responses integrated with audio cues. In the same way, the impact on the sound of different string sets is analyzed.

Part of the results are compared with those found in the literature for validation purposes. In addition, we identify several effects on the quality timbre and the acoustics of the violin introduced by the ground coat and the varnish. These effects can be taken into account by violin makers during the making process of the violin.

Finally, a correlation is determined between the timbral effects introduced by the strings and their tension value. These results can be used by musicians to facilitate the selection of suitable strings to customize their sound.

Acknowledgments

Vorrei ringraziare il Prof. Sarti per avermi dato la possibilità di perseguire questo lavoro di tesi e il Dr. Massimiliano Zanoni per avermi costantemente seguito e supportato durante tutto il mio percorso. Ringrazio per il loro indispensabile contributo e la loro disponibilità Marco Malagodi, Tommaso Rovetta e Claudia Invernizzi del Laboratorio Arvedi del Museo del Violino di Cremona. Un grazie ai liutai Elena Bardella e Sebastiano Ferrari per la loro preziosa collaborazione e senza i quali non si sarebbe potuto realizzare questo lavoro di ricerca. Ringrazio la Fondazione Stradivari del Museo del Violino per aver ospitato e appoggiato questo lavoro. Grazie a Bruno Di Giorgi e a tutti i dottori e dottorandi dell'ISPG che si sono resi disponibili nei momenti di bisogno.

Ringrazio tutti gli amici che mi sono stati vicini, quelli conosciuti a Como e Milano (in particolare Stefano Bider) e quelli di Catanzaro, in particolare Danilo per il suo supporto morale fuori e dentro Skype (...). Un doveroso ringraziamento a Max per tutte le volte che è stato ad ascoltarmi e sopportarmi nei momenti difficili ed allo staff degli Jam Session Music Studios di Milano.

Un ringraziamento particolare alla mia famiglia, Maria, Salvatore e Giovanna, che anche se lontani mi hanno sempre fatto sentire la loro vicinanza e mi hanno dato il loro conforto.

Infine, un grazie speciale a chi da anni mi sostiene e mi affianca in tutte le mie esperienze di vita, a chi ormai fa parte di me e quindi anche del frutto di questo lavoro. Perchè senza di te non ce l'avrei fatta, grazie Valentina.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

Often celebrated as the instrument sounding closest to the human voice, the violin has a special place in musical history. It is the chief of the bowed string instrument family.

The first ancestors of the violin dating back to the Middle Ages, indeed, its name comes from the Medieval Latin word vitula, meaning stringed instrument. The violin, as it is known nowadays, was built in the early 16th century and in this climate the viola and the cello also emerged. Almost everything that is explained about the violin and its structures can be referred to the whole family of stringed instruments. This family has been developed in order to satisfy new ideas of sounds that emerged at that time in Italy.

It was with the Cremonese makers that the violin and its family reached its zenith and, although technical innovations have been applied through time, the basilar design is still in use today. Cremona was the home of the most famous of all violin makers: the families Amati and Guarneri, Antonio Stradivari, the families Ruggeri and Bergonzi. For more than 150 years, violins made by Stradivari and Guarneri have been the most desired concert instruments.

Nowadays, Cremona hosts the Museo del Violino managed by the Fondazione Museo del Violino Antonio Stradivari. The museum is flanked by a research centre which is led by Politecnico di Milano and Università di Pavia. The topics covered by Università di Pavia mainly concern the analysis of materials involved in the violin making process, whereas the Politecnico di Milano aims to inspect the acoustics and timbre properties of musical instruments. The research center started a new project that has the goal of investigate the correlation between the materials characteristics and the acoustics and timbre properties of the instruments. This work of thesis is part of this research project.

All the information gathered from this project will be stored and linked in a Database created from the scratch for this purpose. The design and implementation of the DB was part of this work of thesis and will be presented in Appendix B.

From an acoustic point of view, the charm of the violin comes from its manufacturing complexity. Indeed, over 70 different pieces of wood are put together to compose the modern violin. This makes the research studies very difficult, but also highly interesting. We can identify two main components that are at the basis of the violin acoustics: the body and the strings, respectively identifiable as the resonator and the exciter [5]. These two elements produce sound through the interaction between the bow and the strings. The strings vibrate at a certain frequency (pitch), but can't produce any easily audible sound itself. Indeed, the vibrations produced are transmitted through the bridge to the body. The bridge acts like a filter and is determinant for the resulting sound [6] and the body amplifies the vibrations in a way according to the Helmholtz resonator principle.

How the body reacts to pitched stimuli, is highly related to the top and back plates manufacturing, which are modeled from the violin maker according to several guidelines. The most important of these is to apply a vibration to the plates, before they are glued together. After that, it is possible to study the resulting modes of the vibration in order to deal with the thickness and shape of the wood.

Another peculiar characteristic of the violin are the f-holes situated on the top-plate. The area of these holes affects the Helmholtz resonance frequency [7] of the body.

Last but not the least, there are the strings that are responsible for the frequency generated by the bowing process. The strings have fixed length that can be changed only through the finger pressure onto the neck. This causes a pitch variation being the latter dependent on the length of the string. Therefore, the only physical parameter that we can change on the strings is the tension.

The acoustics effect behind the parts mentioned so far are well-known in the literature. On the other hand, there's another important element that needs to be better studied: the varnish. The varnish is shrouded in mystery, making even more interesting the studies on it. This is due to the myth according to which the secret of the glorious sound produced by the master violins, such as Stradivari ones, is related to the recipe of their varnishes.

Several studies on the violin have been done, in order to inspected the role of its main components in influencing the final sound (body, soundpost, bridge, etc.) [8] [7] [9]. The goal of these studies was to describe how these components affect the timbre of the violin. Many discoveries have been made about the relationship between manufacturing process and the sound quality, but there is still much to be done. For example, only a small number of these studies were conducted on the impact that the varnish has on the sound quality [10] [11]. In this regard, a common question that the community of enthusiasts and researchers of the violin instruments arises is: how a very thin layer of material, such as the varnish, affects the timbre of this instrument?

The answer to the question will be inspected in the first part of this work. For this purpose we mainly use audio analysis, which in turn are supported by analysis on the materials. All the analysis are performed on a violin specifically built for the purpose. The making process is identified and analyzed in three main steps. The first is represented by the violin "in the white", i.e. with all the parts glued together and ready for the varnishing. The second is considered after the ground coat application and the third after the complete drying of the last layer of varnish applied. Both audio and materials analysis are performed for each of the three steps and compared in chapter 5. The final goal is to outline what is the impact of each making step on the final sound characteristics.

From the audio point of view, the frequency response is surely the first tool that can be used. Most of the past works have been done analyzing the frequency response curves computed on violin recordings [6] [4], in order to understand how a violin changes its sound properties depending on the construction parameters.

We integrate the frequency responses with another interesting and more recent analysis tool: Audio Features Extraction. The features analysis is the basement for what concerns the Musical Information Retrieval field [12] [13] [14]. The information acquirable from audio features or audio descriptors, is more sophisticated and is closer to the human perceptive aspects. Due to the reduced data obtained by the features analysis, is possible to investigate in an easier way a larger number of aspects related to the timbre, rather than we can do with the frequency response.

In order to improve and make more relevant our results, the analysis on the material is performed. Two microscopy techniques are used. The first exploits the Polarized Light Microscopy principle [15] in order to inspect the morphological aspect of the materials. The second one uses the Scanning Electron Microscopy and Energy Dispersive X-ray principles [16], in order to get also elemental information.

The second part of this work is related to the impact that different types of string sets have on the timbre. The strings are one of the few violin components that can be adjusted once the making process of the violin is terminated. We conduct this study taking into account a particular physical property of the strings introduced above, the tension. The tension is a relevant property of the strings because it is the variable used by manufacturers and musicians in order to distinguish different kinds of sets present on the market. Finally, the study of the tension can be very interesting since, no other works are present in the literature on this topic.

The audio data were acquired engaging a recording session for each of the three steps and during which a professional musician played the violin in a semi-anechoic room. The analysis of the materials are performed sampling two wooden boards on which the violin maker performed the same varnishing process reserved for the violin.

For the second part of work, we use a brand new viola. We mount four different sets of strings and then we follow the same procedure for the audio data acquisition as for the first part of the work.

The rest of the thesis is organized as follows: In Chapter 2 we will treat the theoretical basis in order to understand the main principle behind the sound production of bowed string instruments, including the relationship between constructive components and the timbre of the sound (with related state-of-the-art) and showing the manufacturing process. In Chapter 3 we will describe the tools employed for this work, both for the audio and for the materials analysis. In Chapter 4 we will discuss the methodology we adopted in order to acquire both audio and materials data and the correlation analysis that we use to find a relationship between the former and the string tensions. In Chapter 5 we will show the final experimental results, comparing them with that found in literature, in order to validate our conclusions. Finally, in chapter 6 we will draw some conclusions about our methodology, discussing some applications and ideas for possible future works.

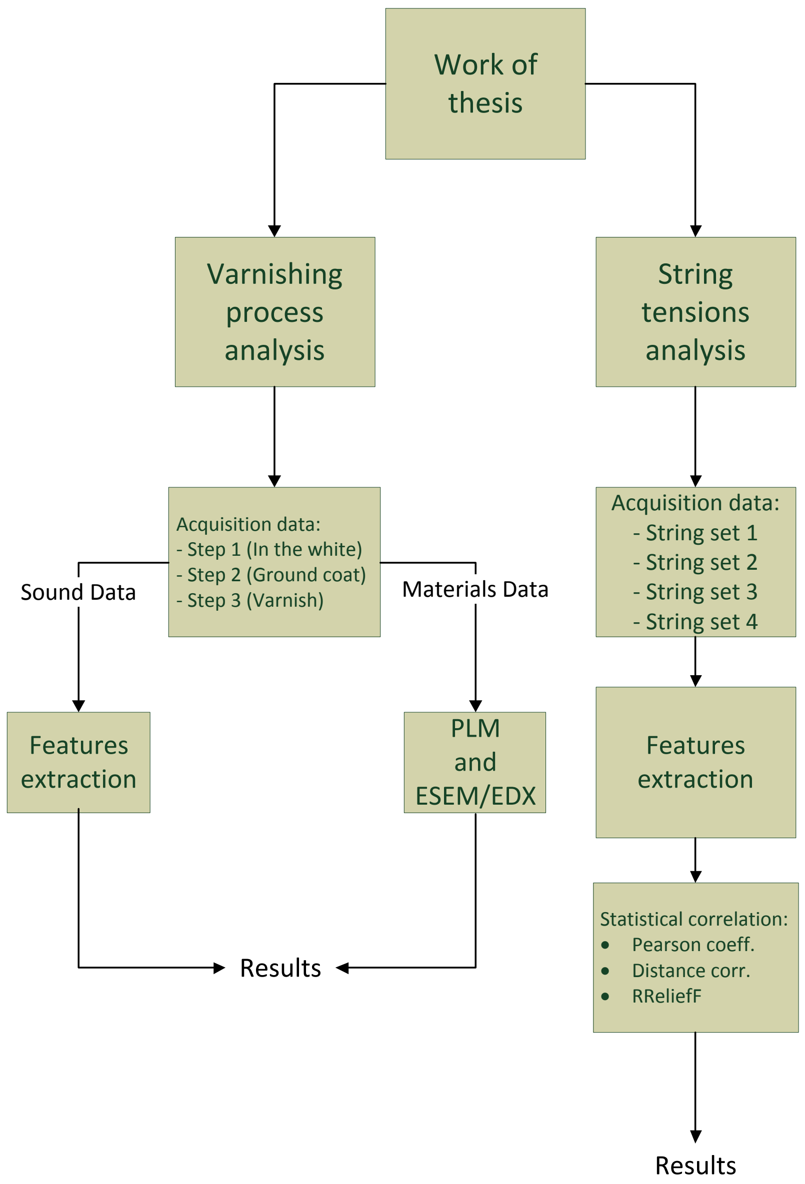


Figure 1.1. Block diagram representing the overall process adopted in this thesis.

Chapter 2

Background

The ability to evaluate objectively the sound quality of a musical instrument is a hard issue. This is particularly true for the violin, which complexity makes it a fascinating bowed instrument. Several studies have already been done on the violin parts, like on: strings, bridge, soundpost, bass bar and body.

Due to the large number of acoustic variables that characterize the violin instrument family, for the sake of simplicity, in this chapter we will concentrate on two main parts, i.e. the strings (the exciter) and the body (the resonator). The acoustic principles behind these two components was explained by Helmholtz with his theories about *string motion* and the *resonance* [5]. The working process between these two components starts from the interaction between the bow and the strings. This interaction produces an oscillation (Fig. 2.1) along the strings and these oscillations are not sinusoidal, but similar to triangular waves. This means that each oscillating string contains several harmonics (ideally infinites), each one with amplitude determined by the Fourier series. The oscillations are then filtered by the soundboard according to its acoustic response, obtaining the resulting final timbre that will be subsequently affected by the acoustics of the surrounding environment. The acoustic response of the soundboard is determined by its resonance modes (see sec. 2.2.2), that describe how the soundboard moves under vibration stimuli. In particular, a resonance mode tend to amplify the harmonic frequencies nearest to its characteristic frequency and, on the other hand, to damp the frequencies which are out of resonance. Therefore, the harmonic components of the original waves are modified in a different and complex way, in general depending on the fundamental frequency excited, i.e. the played note.

The first aspect we want to analyze in this work is how the body behavior changes according to some manufacturing steps. We identified the three main steps of the varnishing process: the violin "in the white" (after all wooden parts are glued together), the application of the ground coat 2.3.2 and the application of the final varnish 2.3.3. For each step we analyze the acoustic behavior, by retrieving a set of timbral cues.

The second aspect that we analyze is the correlation between strings and the sound quality. This is done starting from Mersenne's formula 2.8 and reformulating it in function of the tension. This is important because the tension is used by musicians to distinguish different sets of strings on the market. Indeed, once having fixed the length of the string, the tension is the only physical parameter that can be changed.

The rest of this chapter is organized as follows: in section 2.1 we treat some physical acoustic theory in order to understand the main principle behind the bowed string instruments sound production; in 2.2 we introduce better the violin starting from its history and giving an overview on its main acoustic characteristics; afterwards, in section 2.3, we illustrate the making process of the violin, by taking a look to the materials and techniques used by the violin makers; finally, in 2.4 we review the state of the art in terms of violin sound quality related to its constructive relations.

2.1 Fundamentals of Acoustics

In this section we discuss some basic acoustic principles related to the violin sound production. In particular, we introduce the interaction between the strings and body of the instrument, modeled through the Helmholtz motion and resonance laws, which are the basic principles behind the violin sound production. In the conclusion we take a closer look to the strings acoustic principles, exploiting the fundamental wave relationship by linking it to the string tension parameter.

2.1.1 Helmholtz motion

In order to understand how the string is affected by the bow gestures and how the timbre and dynamics can be manipulated by these, it is of vital importance to have a clear concept of how the string moves under the bow.

The kinematics of the bowed string was first observed and described by Helmholtz in the second half of the 19th century [5]. Since then, the understanding of this phenomenon has gradually increased, and today the motion of the bowed string is the only example of vibration excited by friction which can claim to be well understood [17].

When a bowed string oscillates in steady state, the string takes the shape of two (nearly) straight lines connected by a kink or *Helmholtz corner* that travels along a parabolic path, see figure 2.1. The interaction between the bow and string switches between two conditions: sticking and slipping. During the slipping phase, which is when the kink is traveling between the bow and the bridge, the string moves in a direction opposite to that of the bow. The velocity is then $(\beta - 1)/\beta$ times higher than the (steady) bow velocity supplied by the player, where β denotes the ratio between the distance of the bow from the bridge, and the total string length. During the sticking phase, the string is stuck to the bow hair and consequently takes the speed of the bow.

Real strings have stiffness, and as result, the Helmholtz corner is not perfectly sharp. This leads to at least two phenomena of interest to string players. One is a slight flattening of pitch as the bow force increases. The other is a type of noise called *jitter* that is caused by random variations in the period of the string. The amount of jitter is roughly proportional to the amount of corner rounding [18].

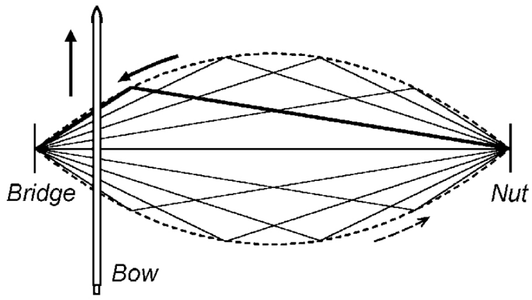


Figure 2.1. The string in Helmholtz motion. Every time the corner of the string passes under the bow on its way to the bridge, the string slips on the bow hair. When passing the bow again on its way to the nut, the string is captured wherefore it stays stuck for the rest of the period.

2.1.2 Helmholtz Resonance

The Helmholtz resonance is the phenomenon of *air resonance* in a cavity, such as when one blows across the top of an empty bottle. The name comes from a device created in the 1850s by Hermann von Helmholtz, the *Helmholtz resonator*, which the author used to identify the various frequencies or musical pitches present in music and other complex sounds [5]. An Helmholtz resonator or Helmholtz oscillator is a container of gas (usually air) with an open hole (or neck or port). A volume of air in and near the open hole vibrates because of the "springiness" of the air inside.

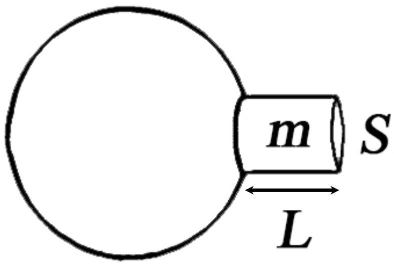


Figure 2.2. Helmholtz resonator with the neck of length L

In the Helmholtz resonator, shown in figure 2.2, the mass of air in the neck and the spring constant of the confined air are given by the expressions

$$m = \rho S L \tag{2.1}$$

and

$$K = \rho \frac{S^2 c^2}{V} \tag{2.2}$$

where S is the area of the neck hole, L is the length of the neck, ρ is the air density, V is the air volume and c is the speed of sound.

We know that a sound wave is created as a result of a vibrating object. Nearly all objects, when hit or struck or plucked or strummed or somehow disturbed, will vibrate. This vibration is performed at a particular frequency or a set of frequencies, known as the natural frequency of the object. The vibration in the case of an Helmholtz resonator is due to the "springiness" of air: when it is compressed, its pressure increases and it tends to expand back to its original volume.

In general, the frequency has a relationship with the speed of the sound and its wavelength described from the following formula:

$$f = \frac{c}{\lambda} \tag{2.3}$$

where c is the speed of sound and λ is the wavelength.

Starting from the eq. 2.3, is possible to obtain the natural frequency of vibration for an Helmholtz resonator [5], that is given by:

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{K}{m}} = \frac{c}{2\pi}\sqrt{\frac{S}{VL}} \tag{2.4}$$

where K is the spring constant, m is the mass of air in the resonator, c is the speed of sound, S is the hole area, V is the volume and L is the length of the neck. Changing these parameters the resonant frequency of the violin changes. In the violin case the mass of air is clearly inside the body, and the area S corresponds to the overall f-holes area. These components are responsible of two well-known resonances of the violin: the *air resonance* (around 300 Hz) and the *wood resonance* (around 500 Hz).

2.1.3 Wave relationship and the Mersenne's frequency formula

A single frequency wave has the form of a sine wave. A snapshot of the wave in space at an instant of time (see figure 2.3) can be used to show the relationship of the wave properties frequency, wavelength and propagation velocity.

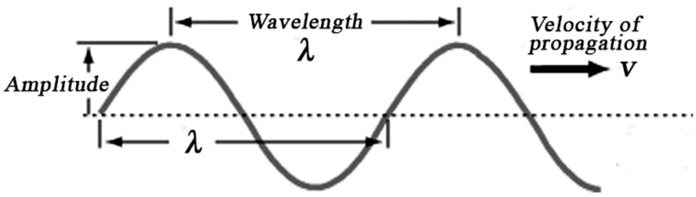


Figure 2.3. Snapshot of a traveling wave.

The motion relationship for a space traversed s at velocity v in time t, is the key to the basic wave relationship and is defined as:

$$s = vt \tag{2.5}$$

At this point we can consider the wavelength λ as the space traversed and using $f = 1/T_w$, where T_w is the time period of the wave, we obtain the standard wave relationship:

$$v = f\lambda \tag{2.6}$$

This is a general wave relationship which applies to sound and light waves, other electromagnetic waves, and waves in mechanical media.

The fundamental vibrational mode of a stretched string is such that the wavelength is twice the length of the string. Applying this concept to the 2.6 and considering the sound speed we obtain the expression for the fundamental frequency:

$$f_0 = \frac{c}{2L} \tag{2.7}$$

where c is the sound speed and L the length.

In [19], the author investigates the proprieties of strings when plucked or bowed. The study outlines that the correlation between the tune frequency f_0, the length L, the weight p ($p = m/L$) for unitary length and the tension T is given by

$$f_0 = \frac{1}{2L}\sqrt{\frac{T}{p}} \tag{2.8}$$

The eq. 2.8 explain the tuning process on the string instruments, showing the dependence of the frequency by the tension. In fact, in most string instruments all the parameter are fixed, except the tension that can be changed, e.g. by rotating the tuning pegs in the case of a violin. In section 5.3 we will investigate how the tension variation can affect the timbre of the violin.

2.2 The violin instrument

In this section we present an overview about the violin history and its main acoustic properties.

2.2.1 Brief history

Often celebrated as the instrument sounding closest to the human voice, the violin has special place in musical history. It is the chief of the bowed string instrument family. The first ancestors of the violin dating back to the Middle Ages, indeed, its name comes from the Medieval Latin word vitula, meaning stringed instrument. The violin, as it is known nowadays, was built in the early 16th century. In this climate the viola and the cello also emerged and almost everything that is explained about the violin and its structures can be referred to the whole family of stringed instruments. This group of instruments has been developed in order to satisfy new ideas of sounds that emerged at those times in Italy. It was with the Cremonese makers working in this environment that the violin and its family reached its zenith, and although technical innovations have been applied through time, the ground plan and its basic form are still used today. Cremona was home to the most famous of all violin makers: The families Amati and Guarneri, Antonio Stradivari, the families Ruggeri and Bergonzi. For more than 150 years, violins made by Stradivari and Guarneri have been the most desired concert instruments.

2.2.2 Acoustic overview of the Violin

From an acoustic point of view, the charm of the violin comes from its manufacturing complexity, indeed, over 70 different pieces of wood are put together to form the modern violin. This makes the research studies more difficult, but at the same time, very interesting. The main parts that compose the violin are showed in figure 2.4. Below we will explain what we think are the most influential parts of the acoustics of the violin, according to our research in the literature.

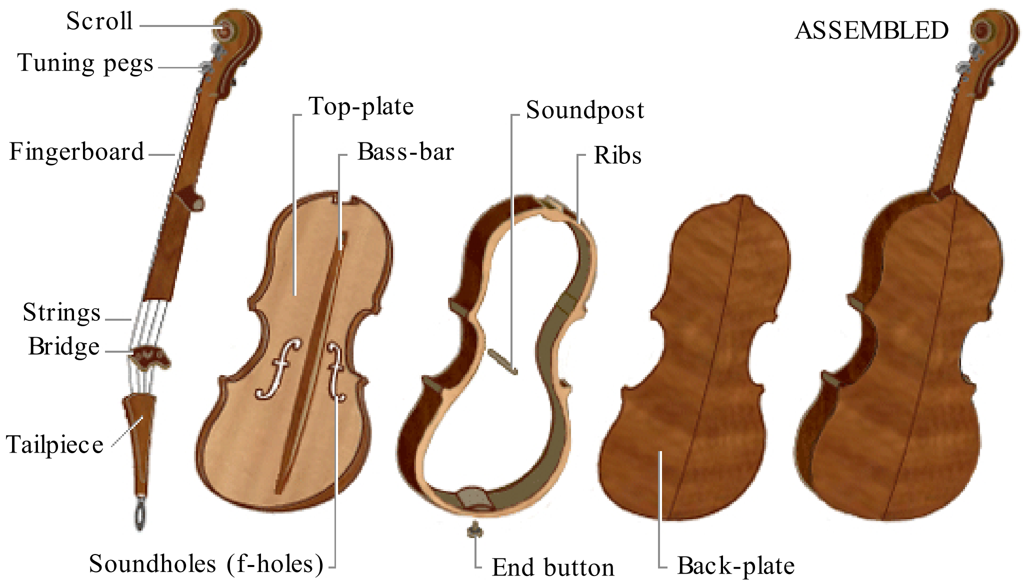


Figure 2.4. The main parts of the violin.

A standard violin is tuned in sequence to E(`659.3Hz`), A(`440Hz`), D(`293.7Hz`) and G(`196Hz`) and the strings are thicker as you go down from the E to G strings. The portion free to vibrate is the same for all four strings and, as anticipated in section 2.1.3, given the standard tuning, the only related parameter that can be modified is the tension. The violin produces sound by bowing one or more of the four strings stretched along its length, with the strings terminated at one end by the supporting bridge and at the other by the end-nut or the player's finger, used to shorten the vibrating length and hence the pitch of the bowed note.

The vibrating string (a linear dipole source) radiates a negligible amount of sound because its diameter is very much smaller than the acoustic wavelengths involved. To produce sound, energy has to be transferred to the radiating body of the instrument via the supporting bridge. Therefore, the bridge is never a perfect node of string vibration, so that the harmonicity of the string modes is perturbed. The bridge itself is very effective at transmitting power to the body at frequencies from about one to four kHz, which is where the human ear is most sensitive. This is one of the reasons for the bright timbre of the violin. Furthermore, the bridge is the component that characterize what is called *bridge hill*, a prominence in the 2000-3000 Hz range. More precisely, this hill arises from a combination of resonance of the bridge and response of the violin body at the bridge foot positions [9][6].

At low frequencies, below around 400–500 Hz, there are no strongly radiating structural resonances. In order to boost the sound of the lowest notes, which extend a whole octave lower (down to 190 Hz), the f-holes perform the function of the opening needed for the Helmholtz air resonance at around 270–280 Hz (see section 2.1.2). The size of the violin becomes comparable to the wavelength at about one kHz so that the breathing action is important for most of the fundamentals of notes played, especially below 600 Hz [7]. If the f-holes area is enlarged the resonant frequency goes up and vice versa (see eq. 2.4). Small variations in volume and hole area are not going to have a large effect. The area of an f-hole can be expressed in terms of an equivalent ellipse for the purpose of calculation [20] as used by Helmoltz [5]. In addition, the f-holes are an additional parameter that influence the bridge hill as evaluated in [21].

The bowed string is excited via the visco-elastic frictional force of the rosin coating both the moving bow hairs and stretched strings. The resultant waves excited on the strings are not the simple sine waves generally used to illustrate wave motion on stretched strings, but are similar to triangular waves like explained in section 2.1.1. The timbre of an instrument will therefore be determined by the overlap of the spectrum of the bowed string waveform with the multi-resonant response of the bridge, the body of the instrument and even the acoustic into which the sound is radiated, as illustrated in fig. 2.5.

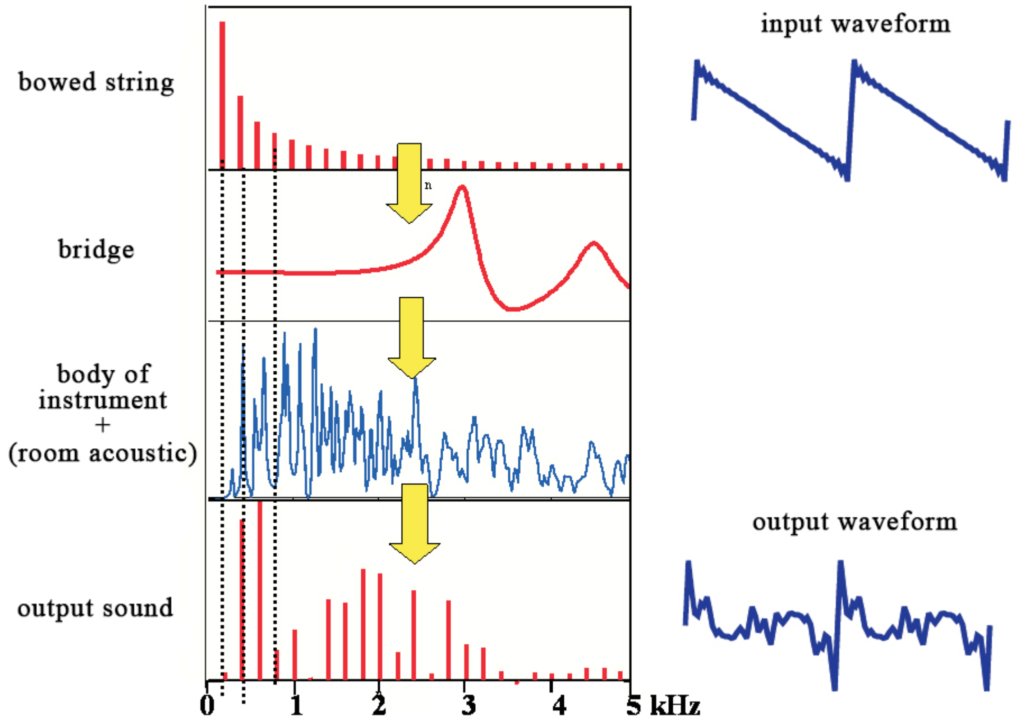


Figure 2.5. A simplified block diagram of the dependence of the sound of the violin on the input bowed waveform and spectrum and their modification by resonances of the bridge and body of the instrument (figure taken from [1]).

The perceived quality of a violin is associated with global features of the acoustic spectrum and not on the response at particular resonant frequencies [22], since the spectrum varies so dramatically from note to note, as it is confirmed from our results (see section 5.2).

The bridge has the main role of transmit the vibration of the strings to the body. The vibrational behavior of the latter, is very important for the sound quality and playability of a string instrument [23].

The top and back plates are made so that they can easily vibrate up and down and each violin is characterized by a number of characteristic modes [24] [25]. The normal modes of vibration or eigenmodes of a violin are determined mainly by the coupled motions of the top plate, back plate, and enclosed air. Further contributions are made by the ribs, neck, fingerboard, strings type, soundpost, etc..

The plates can support longitudinal, flexural and torsional modes of vibration. Of these the flexural and torsional are the only motions that involve acoustically radiating displacements perpendicular to the surface of the plates. The acoustic properties of the plates are determined by their geometric shape, the density and elastic properties of the wood, variations in thickness across the plate and geometric arching. It is the maker's skill in controlling all such parameters that determine the acoustical properties and ultimately the quality of sound of the assembled violin.

Mode shapes appear to differ rather widely in different violins and the most consistent mode is the "air resonance" or "f-hole resonance" anticipated before. Several prominent resonances and antiresonances occur in the 400 to 800 Hz octave [23].

Makers continuously test the elastic properties of the plates thinning them down from the solid. Traditionally, this was done by feel, as the plates were flexed and twisted by the hands, and by listening to the sound when they were tapped or rubbed around the edges by the thumb. More recently, largely inspired by Carleen Hutchins [8], many modern makers use more quantitative, scientific measurements to monitor the modal frequencies and modal line-shapes of the individual plates as they are carved from the solid. From the observed changes in frequencies and nodal line shapes on selective thinning in particular regions, the maker continuously refines the thickness gradations across the plates. The aim is to end up with prominent free-plate modes having a particular set of final frequencies and well-defined nodal line shapes.

Chladni pattern measurements still provide the simplest, most convenient and least expensive way of making such measurements. To obtain Chladni patterns, the plate is lightly supported at the nodal points of the particular mode to be measured and is placed over a loudspeaker cone driven by a sine-wave oscillator [26]. Christmas glitter or some other such light material is sprinkled over the surface. When the frequency of the sound from the loudspeaker strongly excites a resonant mode, the glitter moves to the nodal line positions. Figure 2.6 shows Chladni patterns for a freely supported viola back plate illustrating the three modes that tend to be used in the gradation of thicknesses.

The treble foot of the bridge (the one under the E string) is quite near the soundpost, which is a small post connecting the relatively flexible top plate of the violin to the much stiffer back plate. This post couples the vibrations of the plates [27] making the body stiffened. In [28], the author demonstrated the rise in the main air resonance of about 30 Hz and other studies on this aspect were done in [29]. In addition, the soundpost introduces an asymmetry to the vibrational behavior of

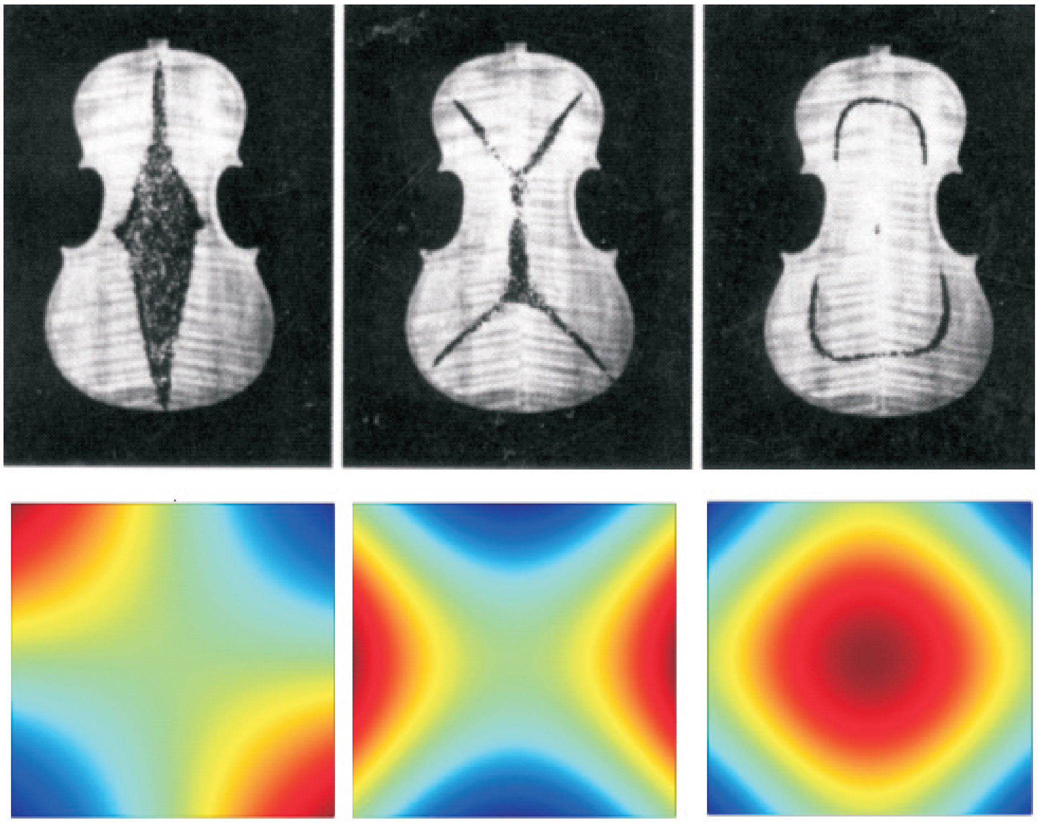


Figure 2.6. An example of Chladni patterns for a viola back plate (figure taken from [1]). At the bottom equal amplitudes of vibration are indicated by different colours (e. g. from red-positive to blue-negative).

the instrument that is essential for it to radiate sound effectively at the lower end of its range. The connection introduced by the soundpost restricts the motion of the treble foot considerably. The bass foot of the bridge is much easier to move up and down. As a result, when a string is driven from side to side by the action of the bow, the bridge tends to pivot about the treble foot. The bass foot moves up and down a little, moving part of the belly with it. The soundpost and its position therefore play a very significant role in determining the intensity and quality of sound produced by an instrument [27], as long recognized by French violin makers, who refer to it as l'âme or *soul* of the instrument. Makers will sometimes move it slightly to change the response of the instrument and small changes can have a noticeable effect.

The asymmetry of the front plate of a violin is further enhanced by the bass bar, which increases its ability to support the downward force of the stretched strings on the bridge without collapsing. It extends beyond the f-holes and thus transmits the motion of the bridge over a large area of the top plate affecting the vibrational modes.

2.3 Violin making process analysis from the maker point of view

There are a lot of elements that affect the sound involved in the violin making process. We will focus only on a part of them: the ground coat and the varnish. In order to do this, we identified three main steps: the violin "in the white" (section 2.3.1), the violin after the application of the ground coat (section 2.3.2) and the violin after the varnishing (section 2.3.3).

Violin making represents more than a craft and involves much more aspects than the mere construction of a violin. Before the violin maker starts to produce an instrument, he must first of all consider the choice of materials and define a strategy of design.

The operation of gluing the front and back plates obtained from a rough piece of wood is very important. In order to shape the outline of the thin strips of maple called the ribs, a shaped container is made, which is later removed from the structure. The front and back plates must be worked fairly thin to be able to resonate easily, but because they must also support a lot of tension, they are shaped into a very exact arching. The edges of the thin plates are then reinforced by the inlaying of the purfling which also gives an adding visual style.

Once the outside arching of the plates has been shaped, the hollowing of the inside can begin. The graduation of the plates to the proper thickness along with the cutting of the sound holes and fitting the bassbar are really the most difficult aspects of violin making. No two pieces of wood have the same properties and must therefore be worked differently every time.

The carving of a scroll is made on the end of the neck and it has a long tradition and probably has symbolic archetypal significance.

After the finished instrument has been coated in a fashionable layer of tinted varnish, it can be adjusted tonally by means of the sound post and other aspects of the final fitting up.

2.3.1 Materials

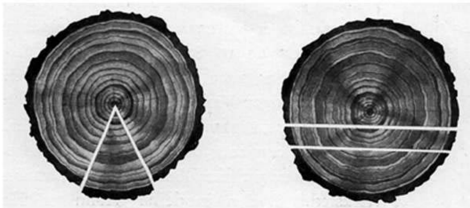
The woods typically employed for the construction of the violin family instruments are maple for the back plate, ribs and neck, and generally spruce for the top plate. The maples are medium sized trees with long-stalked, palmately lobed leaves. The spruce used most often for stringed instrument fronts is the Picea abies or the Picea excelsa, the common Christmas tree, with its spiral branches and narrow needle.

It is sometimes said that wood suitable for violins is that which has grown at high altitudes and has had to suffer harsh conditions such as cold weather and poor soil. While this might seem slightly myth oriented, it is nonetheless true that wood which has grown too quickly in lush environments and rich soil, generally tends to be less resonant and less able to withstand the stresses it is subjected to in the finished state.

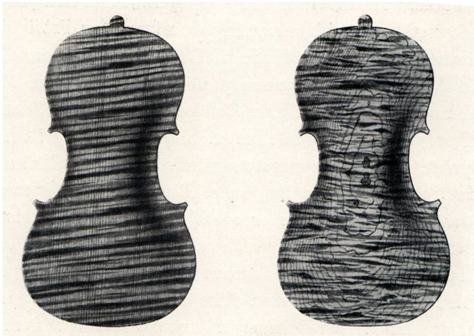
It is also a well-known fact that air dried wood, seasoned for some years, without being kiln-dried is far better choice. This is especially true for musical instruments which are shaped to a thin form and must bear the considerable tensions of the taut strings. Normally 8-10 years are considered necessary to season quality tone wood.

If fresh wood is used it will invariably distort, check and split. The violin maker bites on his wood to try to tell whether it will be strong enough. He lets it fall and listens for its ring. He will try to go by feel.

There are generally two ways to cut pieces of wood from the tree trunk for violin family plates. One is radially from the core of the trunk and is usually referred to as being "on the quarter", and the other is tangentially on the perimeter and usually called "from the slab". In Fig. 2.7 is possible to see an example of the two kinds of cutting and the resulting back plates.



(a) Tree trunks



(b) Violin back plates

Figure 2.7. The two ways to cut a piece of wood. In Fig. 2.7(a) is represented the quarter cut on the left and the slab cut on the right. In Fig. 2.7(b) is possible to view the related aspect results for the back-plate.

2.3.2 Ground coat

Once that the violin making is "in the white" (i.e. all woodcarving and gluing are completed), the next step is the varnishing. The latter consists of several steps related to: the preparation of the wood, the highlighting of the grains and hues and the preparation and application of the colored film and its polishing.

First, the wood must be treated in order to remove resinous substances in excess. In this way the violin maker creates the optimal conditions that allow the wood to absorb and tolerate, in a properly way, the substances that will cover it in a homogeneous and uniform way.

After this initial process, is possible to apply the ground coat that is one of the most important and skilled challenge for a maker. The ground coat stops up the pores of the wood, preventing an excessive absorption of the varnish by the wood, otherwise the latter may assume an unpleasant aesthetic aspect.

In order to obtain an efficient and homogeneous action on the entire surface, the ground coat need to be made up of at least three fundamental elements: a suitable diluent (water, alcohol, etc.) that allow a uniform distribution of the solution; a vehicle substance able to exploit filmogenous properties and an inert and dusty substance that penetrates in the pores filling them (excipient).

The ground coat has a considerable influence on the behavior of the sound by affecting the stiffness of the plate. In section 5.2 we will show this influence according to our results.

2.3.3 Varnish

The varnishing has two basic objectives. One is to protect the wood from the dirt and sweat and the second is to cover the instrument with a colored coating, which has a purely aesthetic function.

The colored varnish that lying on the top of the ground and fades over time, is meant to enhance the appearance of the wood, but can be detrimental to the freedom of vibration of the plates if it is applied too thickly or if it has too hard consistency. The varnish should ideally be thin and light enough not to constrict the instrument.

The myths associated with varnishes never stop to inspire the imagination. The "lost secret" of varnish making has a charming appreciated by many. Indeed, the varnish used on classical Cremonese instruments is one of the great achievements of violin making history. Whether it is possible or not to recreate exactly the varnish of the classical masters is a subject of fervent debate. Is interesting to denote that in the Italian renaissance, the term "secret" didn't have quite the meaning of the word today, the meaning having been more on the lines of "knowledge". It seems that apprentices were on oath, not to divulge the knowledge they had gained in the workshop of their master to other regions, and the oath if broken could have serious consequences.

In section 5.2 we will present our result regarding the influence that the varnish has on the timbre.

2.3.4 Strings

As discussed in the introduction of this chapter, the strings play a fundamental role for the sound production of the violin family instruments.

There exists different types of strings and each of them has its own special characteristics, which highly influence the sound of the violin (see section 5.3). These characteristics make changes in the quality, playability, volume and responsiveness of the instrument. Some instruments respond best to a certain kind of string and less well with other types. Each instrument has its own personal characteristics. A string that works well with one instrument may not produce the best sound with another brand.

The three main type of strings are: Gut strings, Steel core strings and Synthetic core strings. For centuries, all musical strings were made of pure sheep gut. In the 16th century, the lower strings (which were the thickest) were wrapped with silver wire to increase mass. Today, gut strings have a gut core and are not entirely made of gut.

Gut strings are known for having a warm, rich sound with many complex overtones. These considerations have been demonstrated by our results (*Pirastro Eudoxa*®strings in section 5.3). Gut strings tend to take longer to stretch than synthetics, and once stretched they are generally stable, but can react to changing weather conditions and generally require more tuning than synthetic core strings. Different string gauges for gut strings can change the quality and power of tone drastically. Usually a smaller gauge gut string will have less carrying power and be rather bright sounding, whereas a thicker gauge gut string will be more powerful, gritty and with a higher string tension. Musicians playing Baroque or early music often prefer gut strings for the sound.

Soon, steel strings became more popular than gut among non-classical players. Strings made of steel core have a direct, clear sound, and few overtones, although those that are wound can have more interesting overtones. They are much more stable in pitch than gut. They also last longer. They are very bright sounding, and sometimes thin, although again the thinness can be negated by windings. They are also good for smaller, entry-level or beginner instruments. In our experiments on strings (section 5.3), all A (pitched at 440 Hz) strings are made of steel and characterized by a typical bright sound.

Synthetic core strings were invented in the early 1970s when Thomastik-Infeld began producing the revolutionary Dominants, made from nylon perlon. Since then, manufacturers have created many new brands of synthetic strings using other high-tech nylons and composite materials. Synthetic core strings have the warm sound qualities of gut, but are much more stable pitch. These sound quality are confirmed by our experiments (section 5.3) for the sets *Pirastro Obligato*®, *Pirastro Evah Pirazzi*®and *Larsen*®.

Almost all strings are available in different thickness or gauges. The majority of string players use the medium gauges, the same used for our experiments (section 4.2.1). In general, a thicker than normal string will require more tension in order to bring it up to pitch (see eq. 2.8). This increase in tension will produce more volume and sometimes a fuller sound, but with a slower response. A thinner string requires less tension and will give a faster response, but with less volume and a thinner, slightly more focused sound.

In conclusion, synthetic core strings are by far the most popular type of strings because they are more stable than fickle gut strings, but have most of the tonal colors of gut strings. Gut core strings are regarded as having the best tone, but they need to be tuned more often and react to changes in the weather. Steel strings are generally for specialized uses.

2.4 Violin sound quality and its constructive relations: State-of-The-Art

The literature that studies the relations between the construction parameters and the timbre of the violin is very poor.

Several studies have been done to inspect the impact of the varnish on the physical characteristics of the violin. Others studies have investigated the possibility to describe the sound quality of the violin, through the judgment expressed by skilled musicians [30] [31]. As the best of our knowledge, no prior studies was conducted on the impact that the materials involved on the making process have on the sound quality perception of the instrument.

From the physical point of view, in order to investigate relations between constructive elements and the violin sound quality, first approaches have made use of frequency response curves (we will see them in section 3.1). These curves were often measured by several researchers using drive method, i.e. by exciting the violin with vibrational external force or bowing the strings. In the second half of the 19th century, Meinel [2] correlated the response curves to the kind of wood and to the influence of the varnish, without finding particular relationships. In addition the author has established some conclusions on the sound quality perception, indicating several frequency bands of interest.

Acoustical effects of violin varnish was analyzed also by Schelleng [11] in relation to the mass, stiffness, and internal friction that were examined qualitatively and quantitatively. Neither the observations by Meinel nor the measurements done by Schelleng confirm the popular view of violin varnish as a major contributor to acoustical excellence [11]. Small benefits can occur if the wood used is really overly resonant or if damping by varnish increase dramatically above 4000 or 5000 Hz, but of neither of these suppositions have been established.

In [10] the author starts from the Schelleng and Meinel works in order to investigate the properties of the wood before and after varnishing in terms of dynamic modulus and internal friction. Also in this work was not easy to conclude if the varnishing was beneficial acoustically, but an additional final consideration was done. Indeed, it was stated that for a low dynamic coefficient value of the wood, the varnishing can improve the quality of the instrument and the sound power above the 300 Hz, conversely, for a high dynamic coefficient value the result is inverse, proportionally to the thickness of the varnish layer.

From a perceptive point of view, several years later, in [4] the authors measured the frequency responses of 43 violins with highly different sound qualities in order to develop a model that predict the sound quality of a violin, after a learning process. The same instruments were subjectively evaluated by competent experts and classified into five classes of sound quality. The study predicted the violin sound quality conducting three phases. In the learning phase, the data from acoustic measurement and subjective ratings of sound quality are employed for construction of a sound quality predictor. Secondly, in the application phase, data from acoustic measurement of an unknown instrument are used for predicting its sound quality. At the end, in the verification phase, subjective rating of a new instrument is compared with its predicted sound quality.

The bridge-hill [9] [21] was recently suggested as a major physical measure for evaluate the quality of a violin. In this regard, in [3] the author evaluates three violins of professional quality and one soloist's instrument made by Guadagnini in 1778. Four notes were recorded for each violin: the open string G (196 Hz), the fifth note A4# (466 Hz), the 12th note A5# (932 Hz), and the G6 (1568 Hz). In order to get a subjective judgment on the timbre quality, the notes were played back and filtered by an equalizer in octave sub-bands. Each band was isolated and classified by the author according to the magnitude of influence on the overall timbre. The results highlighted that for each note played, the band which has the strongest influence on listening is the one that includes the bridge-hill prominence ($\sim$ 2.5 kHz).

In the last decade of the 19th century and the first of the 20th century, the earliest study exploiting the audio feature extraction and analysis (see Sec. 3.2) has been presented, in order to investigate the sound quality perception. In [32] Łukasik et al., find the dissimilarity factors of the timbre of various master violins, i.e. the features that enable to automatically distinguish between instruments or their groups. The features were searched in the spectral domain using the auditory model of human ear. Two methods for the feature extraction have been applied: based on the harmonic spectral parameters (brightness, relative energy of odd and even harmonics, three tristimulus coefficients [33]) and based on the auditory model of human ear. The originally set of features was very large, so feature reduction stage has to be introduced. Two approaches to the reduction have been applied: one was based on heuristics related to the interpretation of averaged filter responses and the other, automatic, based on Singular Value Decomposition (SVD) of the matrix of features. Machine ability to perceive the differences of violin timbre from isolated tones of instruments has been demonstrated using a multidimensional scaling (MDS) paradigm [34], where the multidimensional feature space is reduced to two dimensions, preserving the original distance relationship between the objects.

In [22] the author extracts several harmonic features (see section 3.2.3) from a collection of 53 violins recordings. A set of linguistic descriptor of the violin timbre are related to these features. The result of the analysis is the allocations of the recorded violins in several semantic categories. The study also shows that the four strings have different values for the same feature. Therefore, each string has to be evaluated singularly.

2.5 Conclusions

In this chapter we outlined the theoretical basis in order to better understand our study.

At first we illustrated the basic acoustic principles behind the violin sound production. We have taken into consideration the two main components that characterize the violin instrument, the exciter and the resonator, i.e. the strings and the body. For the former we carried out the Helmholtz motion equation, starting from the triangular wave generated by the bowing movement on the string. We saw that in this context a characteristic phenomenon comes out, which is called stick-and-slip that causes the jitter noise.

After that, we talked about the Helmholtz resonance illustrating the acoustic principle of an Helmholtz resonator as model for the violin body. In the equation obtained (eq. 2.4) we noticed how the fundamental resonant frequency depends on the body volume and the f-holes area. The former and the latter are related to a prominence slightly below the 300 Hz (*air resonance*) and a prominence below and above 500 Hz (*wood resonance*).

Then, we focused even more on the acoustics of the strings, starting from the wave relationship equation and obtaining the Mersenne's formula that relates the latter to the string tension. We noticed that increasing the mass per unit length of the string, the tension needs to be increased. For this reason, if we have a fixed length for the string, as for the case of a violin, the only way in order to change the pitch is to increase the tension.

Going forward we presented an overview on the components that make the acoustic of the violin. The bowing gesture generates the exciting waves on the string that are transmitted by the bridge to the entire violin body. Therefore, the final sound is the result of a complex chain represented in a simplified way in fig. 2.5. This is the reason why the study of the sound quality of a violin is so difficult.

Since our first aim is to correlates the varnishing process to the timbre of the violin, we gave an overview on the making process and the materials used by the makers. The origin of the wood plays a fundamental role for the characteristic sound. In addition to the provenience of the wood, also the way of cutting is one possible variable. All these parameters contribute to the responsiveness of the wood that in several studies was described in terms of internal friction and dynamic coefficient. As well as the wood, we also shown other possible contributors to the resulting timbre of the violin such the ground coat, the varnish and the strings type.

In the past studies, in order to study the sound properties of the violin, at the beginning the researchers used frequency response curves. More recently, the audio feature extraction was employed to have a more compact and less redundant data.

In the next chapter we are going to illustrate how the audio feature extraction and analysis can provide a measure of the sound properties strictly related to the concepts illustrated in the present chapter.

Chapter 3

Analysis tools

The ability of a human to interpret and describe sounds and music has been subject of studies in many disciplines including signal processing, music information retrieval, acoustics, musicology and psychology. Although an exhaustive knowledge of the perceptual mechanisms is not reached, many studies show that it is highly related to many simple descriptors (audio features) [35][36].

The aim of the audio features is to simplify the analysis that would otherwise be performed on the more large, noisy and redundant methods of representation, like waveform envelope or spectrum. In this way is possible to attribute a basic semantic meaning to the audio acquired data, facilitating the interpretation of the same.

Therefore, this process gives measurable properties of the observed phenomenon, usually containing information relevant for pattern recognition. Normally one feature is not enough to reconstruct the listening experience and is useful to combine several features into feature vectors, describing a multidimensional space. The choice of the latter is strictly related to the case study and the correctness of this selection is still an open issue in audio research field.

In section 3.1 we show how to obtain a spectrum in the digital domain, starting from the basis theory of the Fourier transform; in 3.2 we treat the audio features theory in order to obtain a compact and less redundant informations about the audio data; finally, in section 3.3, we give an overview on the microscope analysis tools used for this work, in order to obtain informations about the morphology of the materials and its elemental composition.

3.1 Fourier transform

The Fourier Transform is a tool that allow to express a generic waveform into an alternate representation, based on sine and cosines. Indeed, through this tool, is possible to demonstrate that any waveform can be represented as the sum of sinusoidal functions. The Fourier transform of a continuous-time signal $s(t)$ may be defined as:

$$S(\omega) = \int_{-\infty}^{+\infty} s(t)e^{-j\omega t}dt, \qquad \omega \in (-\infty, +\infty) \tag{3.1}$$

where ω is the angular frequency (radians per second) and t is the time instant. We can replace the infinite integral in the Eq. 3.1 with a finite sum, obtaining the

Digital Fourier Transform (DFT):

$$S(\omega_k) = \sum_{n=0}^{N-1} s(t_n) e^{-j\omega_k t_n}, \qquad k = 0, 1, 2, ..., N-1 \tag{3.2}$$

where ω_k is the kth frequency sample, N is the number of time samples or frequency samples, $s(t_n)$ is the input signal amplitude at time t_n, t_n is the nth sampling instant (an integer ≥ 0) and $S(\omega_k)$ is the spectrum of s (complex valued) at frequency ω_k.

In the field of digital signal processing, signals and spectra are processed only in sampled form, hence the DFT is considered.
The frequency ω_k is defined as:

$$\omega_k = k\Omega = k\frac{2\pi}{NT} \tag{3.3}$$

where T is the sampling interval. In the signal processing literature, it is common to write the DFT and its inverse in the more pure form below, obtained by setting $T = 1$ in the 3.3:

$$S(k) = \sum_{n=0}^{N-1} s(n) e^{-j2\pi nk/N}, \qquad k = 0, 1, 2, ..., N-1 \tag{3.4}$$

$$s(n) = \frac{1}{N} \sum_{k=0}^{N-1} S(k) e^{j2\pi nk/N}, \qquad n = 0, 1, 2, ..., N-1 \tag{3.5}$$

where $s(n)$ denotes the input signal at time (sample) n, and $S(k)$ denotes the kth spectral sample.

Recall the Euler's Identity:

$$e^{j\theta} = \cos(\theta) + j\sin(\theta) \tag{3.6}$$

Euler's Identity is the key to understanding the meaning of expressions like:

$$x_k(t_n) = e^{j\omega_k t_n} = \cos(\omega_k t_n) + j\sin(\omega_k t_n) \tag{3.7}$$

Such expression defines a sampled complex sinusoid.

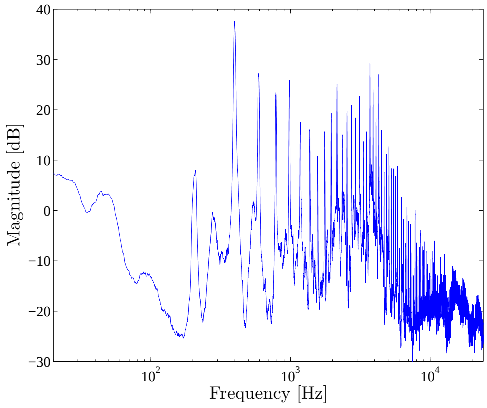
Finally, we need to understand what the summation over n is doing in the definition of the DFT. This can be interpreted as the computation of the inner product of the signals s and x_k defined above, so that we may write the DFT, using inner-product notation as:

$$X(k) = \langle s, x_k \rangle \tag{3.8}$$

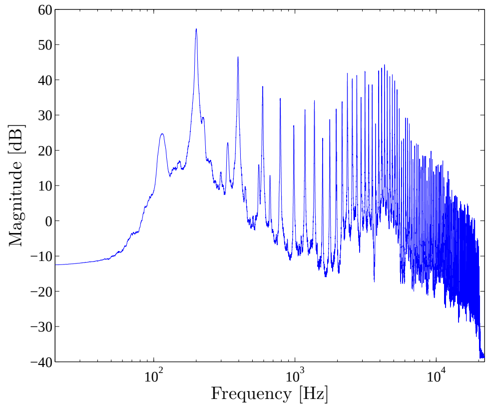
where $x_k(n) = e^{j2\pi nk/N}$ is the sampled complex sinusoid at normalized radian frequency $\omega_k = 2\pi k/N$.

The inner product of s with the kth basis sinusoid x_k is a measure of "how much" of x_k is present in s and at "what phase" (since it is a complex number).

The operation of decomposing a sequence of values into components of different frequencies, is useful in many fields, but the direct implementation of the definition



(a) Viola



(b) Electric guitar

Figure 3.1. STFT related to an open G string (196 Hz) played on a viola Fig. 3.1(a) and on an electric guitar Fig. 3.1(b), computed using the FFT algorithm.

is often too slow to be used in practice. An FFT (Fast Fourier Transform) is a way to compute the same result more quickly: computing the DFT of N points, performing the direct implementation of its definition, takes $\mathsf{O}(\mathsf{N}^2)$ arithmetical operations, while a FFT can compute the same DFT in only $\mathsf{O}(\mathsf{NlogN})$ operations. There exists several algorithm in order to compute the FFT, but the most commonly used is the Cooley-Tukey algorithm [37].

The DFT "architecture" extends to the STFT (Short Time Fourier Transform) that is related to the Fourier transform and is used to determine the sinusoidal frequency and phase content of local sections of a signal as it changes over time. In the continuous-time case, the function to be transformed is multiplied by a window function which is nonzero for only a short period of time. The Fourier transform of the resulting signal is taken as the window is slid along the time axis. Mathematically, this is written as:

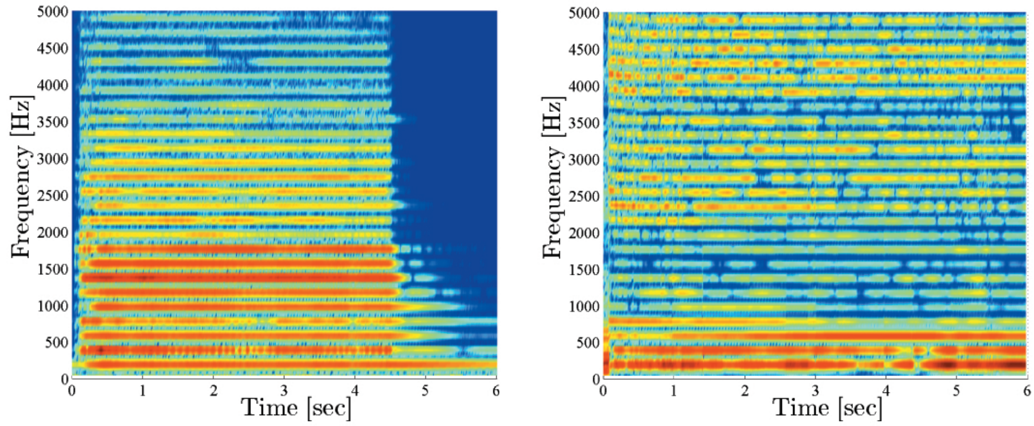
$$S(\tau,\omega) = \int_{-\infty}^{+\infty} s(t)w(t-\tau)e^{-j\omega t}dt \tag{3.9}$$

where $w(t)$ is the window function and $s(t)$ is the signal to be transformed. $S(\tau,\omega)$ is essentially the Fourier Transform of $s(t)w(t-\tau)$, a complex function representing the phase and magnitude of the signal over time and frequency.

In the discrete case, the data to be transformed could be broken up into chunks or frames (which usually overlap each other, to reduce artifacts at the boundary). Each chunk is Fourier transformed, and the complex result is added to a matrix, which records magnitude and phase for each point in time and frequency. This can be expressed as:

$$S(m,\omega) = \sum_{n=-\infty}^{+\infty} s(n)w(n-m)e^{-j\omega n} \tag{3.10}$$

In this case, m is discrete and ω is continuous, but in most typical applications the STFT is performed on a computer using the FFT explained above, so both



(a) Viola **(b)** Electric guitar

Figure 3.2. Spectrogram related to an open G string (196 Hz) played on a viola Fig. 3.2(a) and on an electric guitar Fig. 3.3(b).

variables are discrete and quantized. The magnitude squared of the STFT yields the spectrogram of the function:

$$Spectrogram\{s(t)\}(\tau, \omega) \equiv |S(\tau, \omega)|^2 \tag{3.11}$$

3.2 Low level audio features

Low-level descriptors are the basic components (MPEG-7 foundation layer [35]) in MIR applications and sound and music classification [38] [39] [40] [12] [13] and retrieval [14] [41] [42]. They are also involved in the definition of semantic descriptors [43] [44] [45] [36]. Due to their importance, several automatic extraction tools have been developed in the past few years: MIR toolbox [46], Psysound [47], Marsyas [43], Echonest, MA toolbox [48], Music annotator [49]. For the experiments in this thesis we use the MIR toolbox. Further features, such as Spectral Contrast [50], are also implemented. According to the captured acoustic characteristics low-level features are categorized as followings: Energy features, Temporal features, Spectral features. The list of used LLF that we use in this thesis is reported in Table 3.1.

3.2.1 Energy

The acoustic Energy (or intensity) is related to the amplitude. Amplitude indicates the amount of the force applied over an area and in acoustics it's typically measured in Newtons per square meter (N/m^2). One Newton is the amount of force it takes to accelerate a 1-kilogram object by one meter per second (m/s).

In order to calculate the amplitude of a generic sinusoid over a period of time, a simple average of the instant amplitude values is not sufficient. For example, if we try to calculate the average amplitude of a simple sine wave, it is equal to zero irrespective of its amplitude amount. Indeed, the sine function rises and

Feature set	Features
Energy	RMS
Temporal	Attack Time, Attack Slope, Attack Leap, Zero Crossing Rate
Spectral	Spectral Centroid, Spectral Spreads, Spectral Skewness, Spectral Kurtosis, Spectral Rolloff85, Spectral Entropy, Spectral Flatness, Spectral Roughness, Spectral Irregularity, Spectral Inharmonicity, MFCC, Brightness, Spectral Contrast

Table 3.1. Full list of LLF used in this thesis. LLF are divided into three categories, according to the acoustic cues they capture: Energy, Temporal, Spectral.

falls symmetrically above and below the zero reference. For this reason, a more meaningful method called root-mean-squared or RMS is used. The RMS can be obtained squaring the amplitude values of each point of a waveform and then taking its mathematical average. It can be formalized as:

$$F_{RMS} = \sqrt{\frac{1}{N}\sum_{n=1}^{N} s(n)^2} = \sqrt{\frac{s(1) + s(2) + ... + s(N)}{N}} \tag{3.12}$$

where s(n) is the audio signal at a specific audio sample n and N is the number of samples.

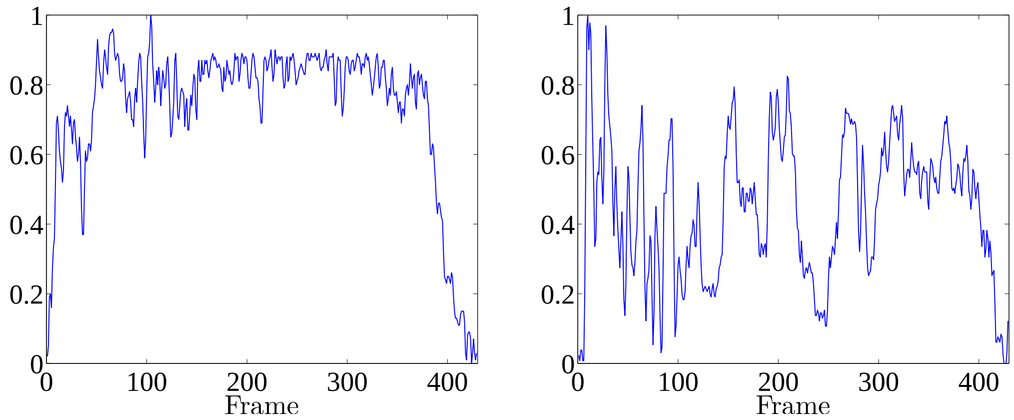
3.2.2 Temporal Features

In order to measure the temporal characteristics of the audio sources, two common features are adopted in this thesis: Zero Crossing Rate (ZCR) [51] and Attack Time [52].

ZCR is defined as the normalized frequency at which the audio signal $s(n)$ crosses the zero axis. It is defined as follows:

$$F_{ZCR} = \frac{1}{2}\left(\sum_{n-1}^{N-1} |sign(s(n)) - sign(s(n-1))|\right)\frac{F_s}{n} \tag{3.13}$$

where N is the number of samples in $s(n)$ and Fs the Sample Rate. ZCR is a measure of the signal noisiness and it has widely been used in different audio applications, e.g. for audio signal classification [38] [39]. As an example in Fig. 3.3 it's possible to notice the difference between the same played note (i.e. same frequency) on a viola and on an electric guitar with an overdrive effect (i.e. the original signal is distorted), in terms of Zero Crossing Rate. In the case of the viola the ZCR value is more constant than that of the electric guitar. This is due to the more harmonicity and less noisy properties of the viola sound than the distorted guitar.



(a) Viola **(b)** Electric guitar

Figure 3.3. Zero Crossing Rate related to an open G string (196 Hz) played on a viola Fig. 3.3(a) and on an electric guitar Fig. 3.3(b).

3.2.3 Spectral Features

Spectral features are highly related to the timbre. The timbre is associated to the tonal qualities that defines a particular sound/source and it can refer to class (e.g. violin or piano) or quality (e.g. bright, rough). Oftentimes the timbre is defined comparatively as an attribute that allows us to differentiate sounds of the same pitch, loudness, duration and spatial location [53].

Is possible to define a timbre space empirically measuring the perceived dissimilarity or similarity between sounds and project this measure to a low-dimensional space where dimensions are assigned to a semantic interpretation (brightness, temporal variation, synchronicity, etc).

The Spectral Features are all computed through spectral analysis techniques on Short Time Fourier Transform (STFT), explained above in the section 3.1, of the audio signal [54] and they represent a set of measures of the shape of the magnitude spectrum. In order to increase the temporal accuracy of the analysis, audio is decomposed into short pieces (frames) of the signal. The technique is called frame decomposition.

As for all distribution functions, the magnitude spectrum can be described by the first four statistical moments (an example is given in the figures 3.4 and 3.5): Spectral Centroid [55], Spectral Spread [56], Spectral Skewness [57], Spectral Kurtosis [57].

In statistics the first moment, called the mean, is the geometric center of the distribution. The Spectral Centroid has the same meaning and is computed on the spectrum distribution. Given a frame decomposition of the audio signal, Spectral Centroid is formalized as:

$$F_{SC} = \frac{\sum_{k=1}^{K} f(k) S_l(k)}{\sum_{k=1}^{K} S_l(k)} \tag{3.14}$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, $f(k)$ is the frequency corresponding to kth bin and K is the total number of frequency bins. Looking at Spectral Centroid, is possible to measure how the magnitude spectrum tends to low or high frequency and it is often associated with the brightness of the sound. In the first quadrants of Figures 3.4 and 3.5, the centroid is compared for an open G string played on a viola and on a electric distorted guitar (both frequencies at 196 Hz). The former is characterized by discontinuities due to the high harmonic content and the values tend to be smaller than the latter. This because the distortion effect applied on the guitar expands the frequencies of the sound upon the maximum frequency audible from a viola, presenting more high frequencies content and more noise.

Spectral Spread is the second moment of the distribution and it is a measure of the standard deviation of the magnitude spectrum with respect to the Spectral Centroid. For a given signal frame it is defined as follows:

$$F_{SS} = \sqrt{\frac{\sum_{k=1}^{K}(f(k) - F_{SC})^2 S_l(k)}{\sum_{k=1}^{K} S_l(k)}} \tag{3.15}$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, $f(k)$ is the frequency corresponding to kth bin, K is the total number of frequency bins and F_{SC} is the Spectral Centroid at the lth frame (see eq.(3.14)). Spectral Spread gives a measure of the compactness of the magnitude spectrum around the Spectral Centroid. A spread distribution of the frequency components generally refers to noisy sounds. For this reason, Spectral Spread provides a measure of the noisiness of a sound source. This can be noticed clearly in the second quadrant of the Figures 3.4 and 3.5, where the viola presents values of spread lower than the electric distorted guitar, highlighting the more noisy content of the latter.

Spectral Skewness is the third moment of the distribution and it gives an estimation on the symmetry of the magnitude spectrum values. A positive value of Spectral Skewness represents an asymmetric concentration of the spectrum energy on higher frequency bins that means a presence of a long tail on lower frequency bins. Vice versa negative Spectral Skewness coefficients represent a distribution with a long tail on the lower frequency bins. The perfect symmetry corresponds to the zero Spectral Skewness value. For a given signal frame it is defined as:

$$F_{SSK} = \frac{\sum_{k=1}^{K}(S_l(k) - F_{SC})^3}{K F_{SS}^3} \tag{3.16}$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, K is the total number of frequency bins, F_{SC} is the Spectral Centroid at the lth frame (see eq.(3.14)) and F_{SS} is the Spectral Spread at the lth frame (see eq.(3.15)). Looking at the third quadrant of the Figures 3.4 and 3.5, is possible to note that the skewness values of the electric guitar are smaller then the viola. This means that the latter has more brilliant sound, having a preponderance of high harmonics above the centroid.

Spectral Kurtosis is the fourth moment of the distribution and describes the size of the tails of the distribution of the Magnitude Spectrum values. Positive Spectral Skewness values mean that the distributions have relatively small tails, distributions with large tail have negative kurtosis, and normal distributions have zero kurtosis. For a given signal frame it is defined as:

$$F_{SK} = \frac{\sum_{k=1}^{K}(S_l(k) - F_{SC})^4}{KF_{SS}^4} - 3 \qquad (3.17)$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, K is the total number of frequency bins, F_{SC} is the Spectral Centroid at the lth frame (see eq.(3.14)) and F_{SS} is the Spectral Spread at the lth frame (see eq.(3.15)). The offset -3 at the end of the formula is a correction to make the kurtosis of the normal distribution equal to zero [46]. In the Figures 3.4 and 3.5, is evident that the values of Spectral Kurtosis of the viola are larger than the guitar, due to the more spiky spectrum of the former. in other word this means that the viola has an high harmonic content, whereas the electric distorted guitar is much more noisy.

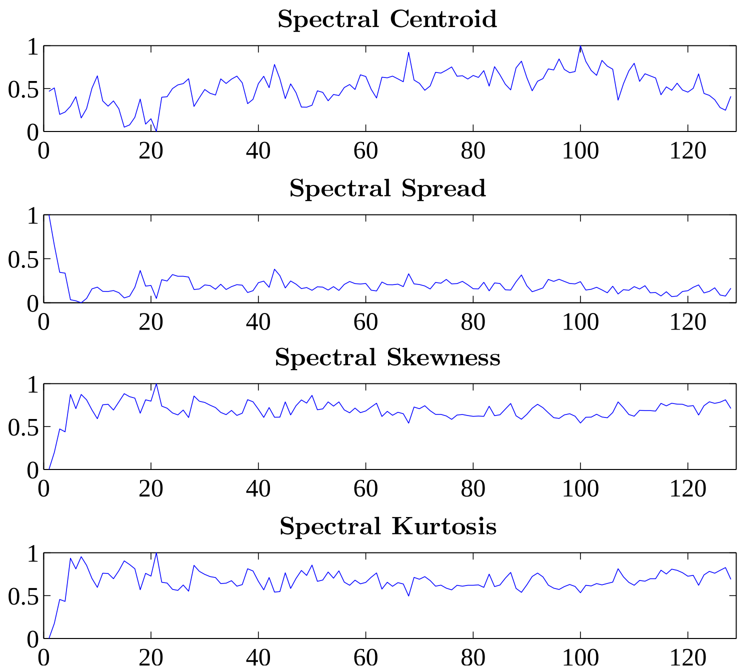


Figure 3.4. Spectral Moments related to an open G string played on a viola.

Quality of sounds often relates to high frequency components of the spectrum. An indicator of the amount of high-frequency bins in the signal is the high-frequency energy. In this thesis we adopted two Spectral features in order to measure the high-frequency energy: Spectral Rolloff [43] and Brigthness [14].

Spectral Rolloff is defined as the lowest frequency at which the accumulated sum of the whole lower frequency power spectrum values reach a certain amount of the total sum of the magnitude spectrum. For a give signal frame it is defined as:

$$F_{SR} = min\{f_{K_{roll}} | \sum_{k=1}^{K_{roll}} (S_l(k)) \geq R \sum_{k=1}^{K} (S_l(k))\} \qquad (3.18)$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, K is the total number of frequency bins, K_{roll} is the frequency bin index corresponding to the estimated rolloff frequency, $f_{K_{roll}}$ is the estimated rolloff

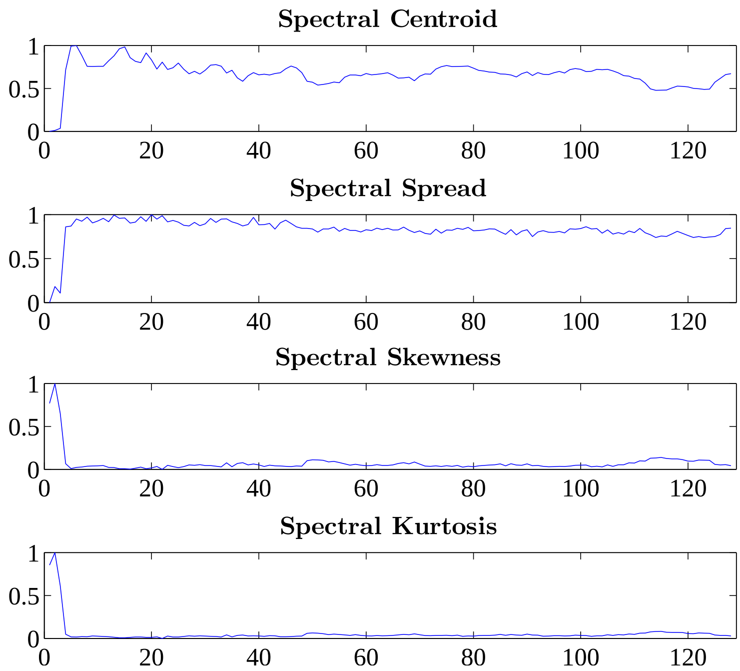


Figure 3.5. Spectral Moments related to an open G string played on a electric guitar.

frequency and R is the frequency ratio. In [43] authors consider R at 85%, whereas in [58] R is fixed at 95%.

The dual of the Spectral Rolloff is the Brightness descriptor that is defined as the measure of amount of energy above a fixed frequency $f_{K_{br}}$. Mathematically is defined as:

$$F_{BR} = \sum_{k=K_{br}}^{K} (S_l(k)) \tag{3.19}$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin, K is the total number of frequency bins, K_{br} is the frequency bin corresponding to the fixed frequency bound.

The magnitude spectrum derives many of the audio features presented in this section, which are the most used in MIR context. However, in some contexts, a linear representation of frequency progression is not sufficient to best characterize sounds [59]. This is especially true in applications where the human perceptual model plays an important role. In those cases a spectrum where the frequency bands are positioned logarithmically can better approximate the response of the human auditory system: Mel-Frequency Cepstrum Coefficients (MFCCs) [35]. MFCCs are based on Mel-Frequency scale [60], which is a scale of pitches judged by listeners to be perceptually equal in distance one from another. MFCCs are obtained as the coefficients of the discrete cosine transform (DCT) applied a reduced Power Spectrum. The reduced Power Spectrum derived as the log-energy of the spectrum is measured within the pass-band of each filter of a mel-filter bank. The cepstral coefficients are finally formalized as:

$$c_i = \sum_{k=1}^{K_c} \left\{ \log(E_k) \cos\left[i \left(k - \frac{1}{2} \right) \frac{\pi}{K_c} \right] \right\} \qquad 1 \leq i \leq N_c \tag{3.20}$$

where c_i is the ith MFCC component, E_k is the spectral energy measured in the critical band of the ith mel filter and N_c is the number of mel filters, K_c is the number of cepstral coefficients c_i extracted from each frame.

In order to summarize useful information in few easier-to-use coefficients, many features are computed by some reduction process of the spectrum components. One of this features is the Audio Spectral Envelope (ASE). The ASE is a log-frequency power spectrum and it is obtained by summing the energy of the original power spectrum within a series of frequency bands. The bands are logarithmically distributed between two frequency edges $loEdge$ and $hiEdge$. The spectral resolution r of the frequency bands within the $[loEdge, hiEdge]$ interval can be chosen from eight possible values, ranging from 1/16 of an octave to 8 octaves. Lower and upper limit of each band are given by:

$$loF_b = loEdge * 2^{(b-1)r} \qquad 1 \leq b \leq B_{in} \tag{3.21}$$

$$hiF_b = loEdge * 2^{br} \qquad 1 \leq b \leq B_{in} \tag{3.22}$$

where b is the index of the band in which the ASE is computed. Finally we can obtain each sub-band value defined as:

$$ASE(b) = \sum_{k=loK_b}^{hiK_b} S(k) \qquad 1 \leq b \leq B_{in} \tag{3.23}$$

The reduction process, however, has the disadvantage to reduce the spectral information. In order to compensate it in the past few years Spectral Contrast coefficients [50] have been used in many MIR applications [59] [61]. Spectral Contrast coefficients attempt to capture the relative distribution of the harmonic and non-harmonic components in the spectrum. The FFT is performed on the audio signal and the result is divided in sub-bands. In order to ensure the steadiness of the feature, the peaks and spectral valleys are considered as the mean value computed on the neighbors of the exact maximum and minimum values. For this reason a factor α is introduced in order to describe the small neighborhood. Once that the FFT is performed, for the kth sub-band we obtain a vector that is $\{x_{k,1}, x_{k,2}, ..., x_{k,N}\}$. The latter is sorted in descending order obtaining a new vector that can be represented as $\{x'_{k,1}, x'_{k,2}, ..., x'_{k,N}\}$, where $x'_{k,1} > x'_{k,2} > ... > x'_{k,N}$. At this point the peaks and spectral valleys can be calculated as follow:

$$Peak_k = \log\{\frac{1}{\alpha N} \sum_{i=1}^{\alpha N} x'_{k,i}\} \tag{3.24}$$

$$Valley_k = \log\{\frac{1}{\alpha N} \sum_{i=1}^{\alpha N} x'_{k,N-i+1}\} \tag{3.25}$$

Finally, the Spectral Contrast can be calculated as their difference:

$$SC_k = Peak_k - Valley_k \tag{3.26}$$

where N is total number in kth sub-band.

In the field of Spectral Features is possible to identify a subset of features strictly related to the harmonic content of the sound. The harmonic content is one of the fundamental proprieties in describing the voice of a bowed instrument. Often harmonic proprieties can strongly characterize the use of the instrument in a given music genre and its use in a solo or in orchestra. To measure the harmonicity of the sound, common features that can be used are: Spectral Entropy, Spectral Flatness, Inharmonicity and Spectral Irregularity. The first two audio features measure the amount of noisiness in the audio signal. Obviously knowing the amount of noise is useful to evaluate the harmonic content, because the more the sound is noises, the less its spectrum is harmonic.

Spectral Entropy is a measure of the flatness of the magnitude spectrum by the application of Shannon's entropy commonly used in information theory context:

$$F_{SE} = -\frac{\sum_{k=1}^{K} S_l(k) \log S_l(k)}{\log K} \tag{3.27}$$

where $Sl(k)$ is the magnitude spectrum at the lth frame and the kth frequency bin and K is the total number of frequency bins. A totally flat magnitude spectrum corresponds to the maximum uncertainty and the entropy is maximal. On the other hand, the configuration with the spectrum presenting only one very sharp peak and a flat and low background corresponds to the case with minimum uncertainty, as the output will be entirely governed by that peak.

Spectral Flatness estimates the similarity between the magnitude spectrum of the signal frame and the flat shape inside a predefined frequency band. As showed in Fig. 3.6, higher values of Spectral Flatness correspond to a noisy sound and vice versa. Mathematically it is defined as the ratio of the geometric mean of the arithmetic mean of the magnitude spectrum:

$$F_{SF} = \frac{\sqrt[K]{\prod_{k=0}^{K-1} S_l(k)}}{\sum_{k=1}^{K} S_l(k)} \tag{3.28}$$

where $Sl(k)$ is the Magnitude Spectrum at the lth frame and the kth frequency bin, K is the total number of frequency bins.

The last three audio features useful to describe the harmonic content are particularly related to the dissonance. When tones are sounded simultaneously, the result may be considered as either pleasant or unpleasant that can be also described in terms of consonant or dissonant. The first is the Spectral Irregularity that expresses the variation of the successive peaks of the spectrum. In the implementation by Jensen [62] it is computed as the square of the difference in amplitude between adjoining partials. Mathematically it is formalized as:

$$F_{IR} = \frac{\sum_{k=1}^{K} (S_l(k) + S_l(k+1))^2}{\sum_{k=1}^{K} S_l(k)^2} \tag{3.29}$$

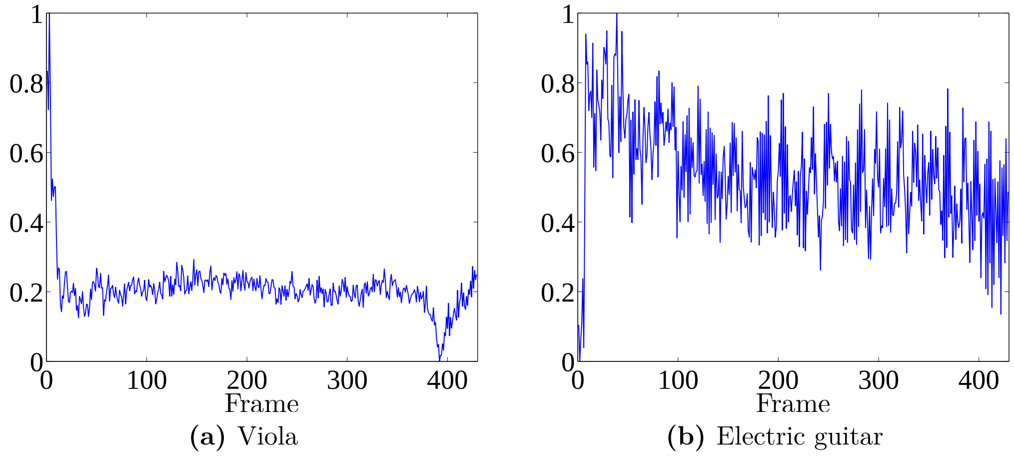


Figure 3.6. Spectrum Flatness related to an open G string (196 Hz) played on a viola Fig. 3.6(a) and on an electric guitar Fig. 3.6(b).

where $Sl(k)$ is the Magnitude Spectrum at the lth frame and the kth frequency bin.

Inharmonicity is a measure of divergence of the partials respect to the purely harmonic signals, which are only made of multiples of the fundamental frequency $f0$. It is formalized as:

$$F_{SI} = \frac{2}{f_0} \frac{\sum_{k=1}^{K} |f_k - kf_0|(S_l(k))^2}{\sum_{k=1}^{K} (S_l(k))^2} \tag{3.30}$$

where $Sl(k)$ is the Magnitude Spectrum at the lth frame and the kth harmonic f_k of the fundamental frequency f_0.

The third feature related to the dissonance is the Spectral Roughness. In [63], the authors have proposed an estimation of the sensory dissonance, or roughness, related to the beating phenomenon whenever pair of sinusoids are closed in frequency. The authors propose as a result an estimation of roughness depending on the frequency ratio of each pair of sinusoids. In this study, the estimation of the total roughness is performed by computing the peaks of the spectrum, and taking the average of all the dissonance between all possible pairs of peaks [64].

3.3 Analysis tools for the materials

In this section we briefly introduce the analysis tools used in this work for the inspection of the materials. The aim is to characterize the penetration of the materials involved both in the finishing making process and varnishing of the violin.

3.3.1 Polarized light microscopy

The polarized light microscopy (PLM) includes different type of techniques and all of these include the illumination of the sample with polarized light. Polarization refers to limitations on the direction of wave oscillation.

First, the light passes the polarizer, which acts like a filter to allow only light oscillating in one orientation to pass [15]. Then, the polarized light passes through an analyzer which contains another filter that can be oriented, in order to capture in different ways how the light is transmitted or reflected by the sample.

The PLM is capable of providing information on absorption color and optical path boundaries between minerals of differing refractive indexes and it can also distinguish between isotropic and anisotropic substances. Furthermore, the polarized light interacts with the sample and so generating contrast with the background. This contrast-enhancing technique exploits the optical properties specific to anisotropy and reveals detailed information concerning the structure and composition of materials. In this work were used both white and UV light.

3.3.2 Scanning Electron Microscopy-Energy Dispersive X-ray spectroscopy (SEM-EDX) (SEM/EDX)

Scanning electron microscopy (SEM) [16] is one of the most used analytical techniques. The electron microscope, compared to the optical one, offers several advantages including high magnification, large depth of focus, great resolution and ease of sample preparation and observation. Electrons generated from an electron gun enter a surface of a sample and generate many low energy secondary electrons. The intensity of these secondary electrons is related to the surface topography of the sample. An image of the sample surface is therefore constructed by measuring secondary electron intensity as a function of the position of the scanning primary electron beam.

In addition to secondary electrons imaging, backscattered electrons imaging and Energy Dispersive X-ray (EDX) Analysis are also useful tools widely used for chemical analysis. The intensity of backscattered electrons generated by electron bombardment can be correlated to the atomic number of the element within the sampling volume. Hence, qualitative elemental information can be revealed. The characteristic X-rays emitted from the sample serve as fingerprints and give elemental information of the samples including semi-quantitative analysis, quantitative analysis, line profiling and spatial distribution of elements. SEM has wide ranges of applications both in industry and research.

3.4 Conclusions

In this chapter we shown the analysis tools used in our study. From the audio point of view, we illustrated the Fourier Transform that is at the basis for our audio analysis. The latter describe the frequency content of a generic sinusoid in the continuous time domain. In the digital domain the time is not continuous, but discrete. For this reason we introduced the digital version of the Fourier Transform that is the DFT. In order to obtain a reasonable computational cost the FFT algorithm is used instead of the direct DFT implementation. Starting from that, we shown how to obtain the STFT in order to get frequencies content of local sections of a signal as it changes over time, by using a window function. At this point we illustrated how is possible to compute an audio spectrum, from which is possible to extract several low level descriptors (or audio features). These features are useful to obtain a more compact and less redundant data about the audio under analysis.

From the materials point of view, we briefly presented two analysis tools. The first is the Polarized light microscopy, that exploits the light properties employing a polarizer filter on a optical microscope, that allow to analyze the morphological aspect of the material samples. The second tools we presented, is the Field Emission Scanning Microscopy and Energy Dispersive X-ray Analysis. By using an electron bombardment onto the sample, through this tools is possible to got a qualitative and semi-quantitative elemental informations about the samples analyzed.

In the next chapter we will see how we implemented these tools for our analysis on the violin.

Chapter 4

Methodology

In the previous chapter we presented the analysis tools used in this work. Their goal is to extract from the violin both audio information as well as information on materials involved in the making process.

The violin used for this study, was obtained thanks to the collaboration of an expert violin maker who realized a violin from the scratch, following the guidelines of a Stradivari 1715. In order to allow needed invasive analysis on the materials, two boards made of the same pieces of wood used for the violin was preserved. One board corresponds to the top-plate and is made of spruce wood and the second one corresponds to the back-plate and is made of maple wood. Both boards have been subject to the same varnishing procedure reserved for the violin. This means that all the layers of ground coat and varnish applied on the violin was also applied on the two boards, with the same technique and the same environmental conditions.

In this chapter we will illustrate how the data were acquired and how the analysis tools were applied on it.

First, we will see the procedure adopted in order to acquire the sound produced by the violin. From an acoustic point of view, in this step, is mandatory to operate into an environment as neutral as possible, in order to reduce the possibility of acquisition error. For this reason the recordings were realized with an high quality recording system and the performances were realized in a semi-anechoic room. The same procedure was adopted also for the second part of this work, related to the string tension.

Finally, we will briefly describe the procedure adopted for the preparation of the material samples and how the analysis tools were applied on them. As for the audio data, also in this case several precautions were adopted in order to reduce possible errors. The data acquisition was performed in a controlled environment room, paying attention against possible external agents action onto the samples.

4.1 Feature-based analysis for violin making steps

The aim of the first part of this work is to analyze the impact that the varnish process has onto the sound quality of a violin. In order to do this, we took advantages from the collaboration with an expert violin maker with which we identified three steps to be analyzed and that are explained in detail in Sec. 2.3. Figure 4.2 shows the violin realized in the three different steps. The first one is the violin "in the white" (i.e. all wood carving and gluing completed), the second is the instrument after the application of the ground coat and the third is characterized by the application of the final varnish. As explained before, all steps were realized in parallel on the violin and on the wooden boards visible in Figure 4.1. For each of the above steps, we have performed both audio and materials data acquisitions on the violin realized by the maker, specifically for this work of thesis.

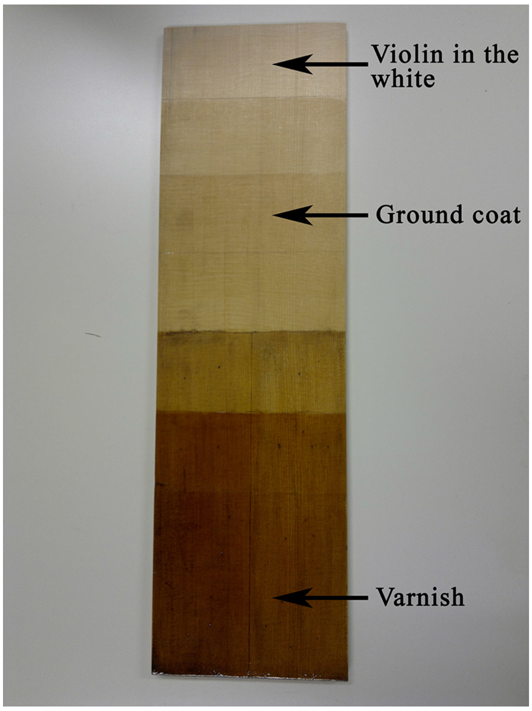


Figure 4.1. The wooden board of spruce corresponding to the top-plate. The arrows indicate the three main steps identified for this work.

4.1.1 Audio data acquisition

The audio data was acquired during a recording session in which the violin was played by a skilled musician. All the recordings were realized into a semi-anechoic room with controlled environment, that is part of the Museo del Violino laboratory, in Cremona.

After a preliminary analysis, we decided to place a couple of omni-directional microphones in correspondence of the f-holes area and in front of the center of the neck (as visible in Fig. 4.3), in order to acquire as better as possible all the sound properties.

The microphone signals were routed into an high quality recording system (see Appendix A), composed by a microphone pre-amplifier and an A/D converter settled to a sample frequency of 48kHz. The *Adobe Audition CS6*® software was used for the acquisition.

For each recording session, the musician was asked to play a sequence of notes (see Table 4.1), selected in order to excite as many frequencies as possible by moving

Figure 4.2. The violin in the three steps of the making process. In the white Fig. 4.2(a), with ground coat in Fig. 4.2(b) and varnished in Fig. 4.2(c).

along the entire neck. In this way we performed a "sampling" of the latter in three equidistant points.

Open string name	1st	5th	10th
G	G3 (196 Hz)	D4 (293.7 Hz)	B4 (493.9 Hz)
D	D4 (293.7 Hz)	A4 (440 Hz)	F5# (740 Hz)
A	A4 (440 Hz)	E5 (659.3 Hz)	C6# (1109 Hz)
E	E5 (659.3 Hz)	B5 (987.8 Hz)	G6# (1661 Hz)

Table 4.1. Full list of single notes played on the violin. For each string the 1st, 5th and 10th was played three times in up and down directions. The notes are represented by the name comprehensive of the octave number combined with the related pitch frequency in parenthesis.

The notes were played, with almost the same intensity, with a slow bowing both in up and down directions and waiting for a complete decay of the sound before starting a new performance. In order to attenuate the impact of possible noises, interfering signals, and performance-dependent features, each performance has been repeated three times. Multiplying the number of the performances by the two bowing directions, we obtained a total number of six different instances for each note listed in Table 4.1. In addition, in order to emphasize as many timbre characteristics as possible, a more expressive performance was done, executing a major scale and two short pieces of song with a duration of more or less one minute.

Between each step, the maker had to remove the strings in order to apply the ground coat and varnish. This action implied the displacement of the soundpost and bridge. In order to keep as constant as possible the settings of the violin, for each recording session a re-setting procedure was adopted by the maker. The soundpost was repositioned with the help of a marker done on the back-plate in the inner

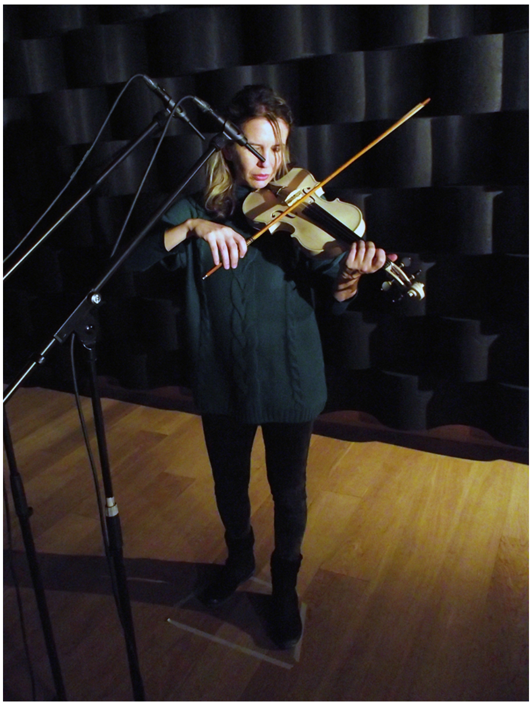


Figure 4.3. The violinist playing during the first recording session. Is possible to note the violin in the white and the location of the two microphones.

part of the body. The same principle was adopted also for the bridge location and inclination.

4.1.2 Audio data analysis and features extraction

Once the data were acquired, the *Matlab*® software was used to perform the audio analysis. We extracted the audio features presented in Section 3.2 making use of the MATLAB *mirtoolbox* [46], following the procedure explained below and illustrated in Figure 4.4. First of all a window-based frame decomposition was done on each audio file using a window length of 2048 samples with hopsize of 50%. Afterwards, the features were extracted for each frame in order to have a more accurate evaluation. This produced a feature vector for each frame of analysis. At this point the average of these set of vectors was computed, obtaining a single vector containing a the feature values for a single performance. In order to have a single meaningful features vector for each note, the average on the six feature vectors corresponding to each performance was computed.

4.2 Feature-based analysis for viola strings identification

In section 2.1.3 we have shown how the tension parameter plays an important role regarding to the sound of the string instruments. More precisely, we shown how the tension is the unique parameter of the string that can be changed given a fixed

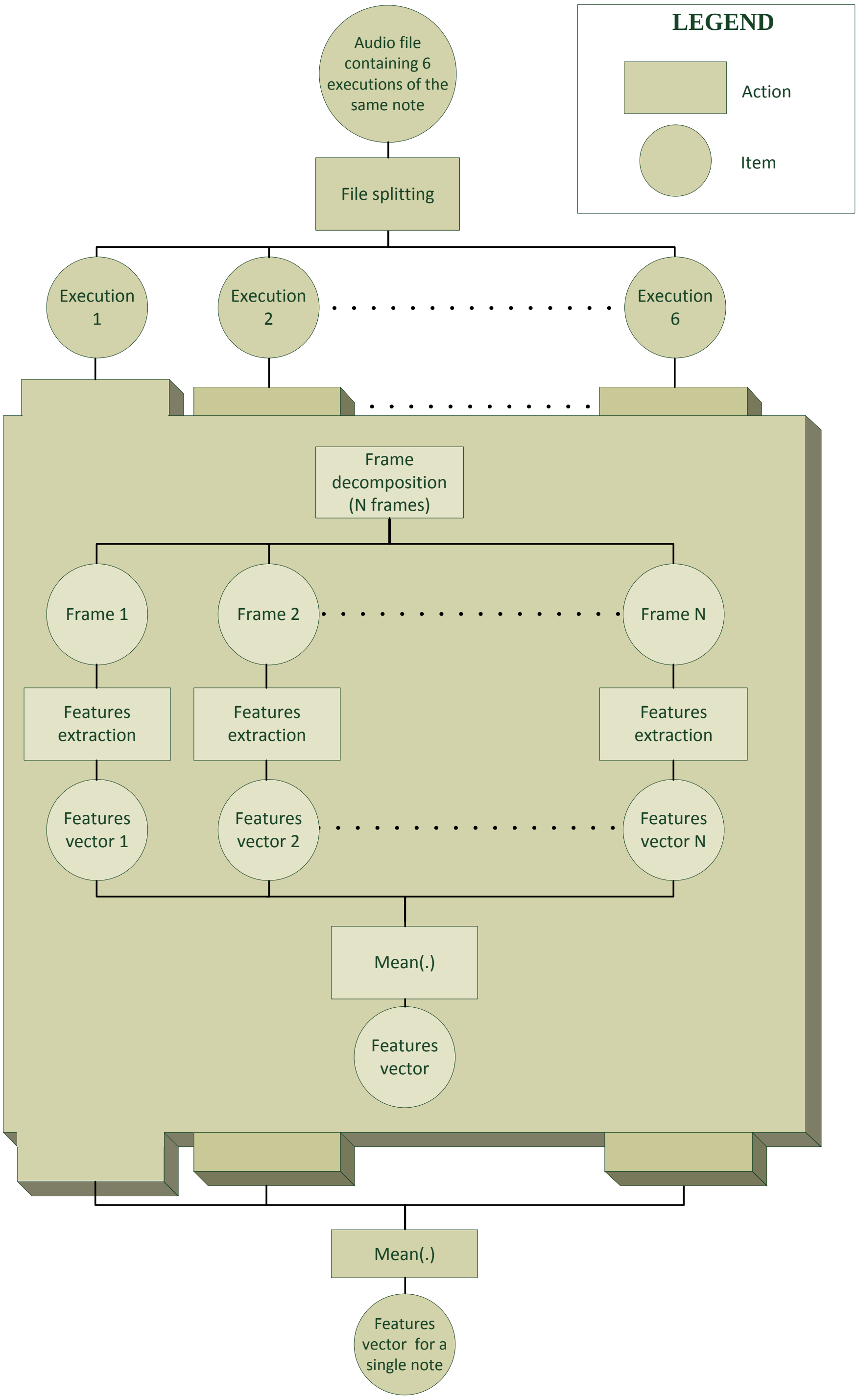


Figure 4.4. Reading the present flow diagram from top to bottom is possible to understand the entire process of features extraction performed for each single note. The initial input is represented by the audio file containing the six executions and the final output is a single vector made up of the features related to the note under analysis.

length of the string self. Furthermore the tension is the principal characteristic of the string used in commerce by musicians and manufacturer companies, in order to distinguish among different set of strings.

4.2.1 Audio data acquisition

In order to assess the relationship existing between the tension of the string and the audio descriptors, we have used a brand new viola on which four different kinds of pre-settled strings have been mounted. More specifically, the string sets are *Pirastro Eudoxa*®, *Pirastro Obligato*®, *Pirastro Evah Pirazzi*®and *Larsen*®. All of them are of medium gauge. For each set and for each string properly tuned, the manufacturers provided us the nominal values of the tension. During the recordings, we also verified that the nominal values of the tension were correct in order to discard possible faulty sets.

Open-string sounds have been recorded with an omnidirectional microphone. After a preliminary analysis, we decided to locate the microphone in proximity of the f-holes. For the reasons explained in Sec. 4.1.1, the performances have been accomplished in a silent semi-anechoic room by a professional musician.

The musician was asked to play each open string at time, with the same intensity, with a slow bowing in up direction and waiting for a complete decay of the sound before starting a new performance. Unlike the recordings done for the violin, this time each performances has been repeated five times.

4.2.2 Audio data analysis and features extraction

As for the analysis performed on the violin, also in this case the *Matlab*® software was used for the audio analysis on the viola recordings. The procedure to extract the features was identical of that described in Figure 4.4. We split the files in order to obtain each single execution for each note. The features was extracted after a preliminary segmentation and than averaged between frames and for each execution. The result is a single features vector for each note or song/scale. The only difference, respect to the procedure described in sec. 4.1.2, was that this time we had five instances of the same note bowed only in up direction instead of six executions played in up and down direction.

4.2.3 Statistical correlation

In this section, the tension of the strings and the audio features are correlated through a set of statistical correlation measures. Since we want to estimate the correlation existing between each acoustic feature and a reference parameter such the tension (t), we resort to using binary correlation indexes. Within this context, the most used indexes are the Pearson coefficient and the Distance Correlation.

Let us denote with f_i the audio feature under study. The sample Pearson correlation index for this feature is P_i and it is computed as

$$P_i = \frac{\mathrm{E}[f_i\ t] - \mathrm{E}[f_i]\mathrm{E}[t]}{\sqrt{\mathrm{E}\{(f_i - \mathrm{E}[f_i])^2\}\mathrm{E}\{(t - \mathrm{E}[t])^2\}}}\ , \tag{4.1}$$

where E[$\cdot$] represents the expectation operator. It is known that

$$-1 \leq P_i \leq 1$$

and more specifically a value close to 1 or -1 denotes a dependency between f_i and t. Since for the experimental data used in our work the expectation of the features under analysis is not available, we approximate the expectation operator with the sample estimate. In particular

$$E[f_i] \approx \frac{1}{R} \sum_{r=1}^{R} f_{i,r} \ ,$$

$$E[f_i \ t] \approx \frac{1}{R} \sum_{r=1}^{R} f_{i,r} \ t \ ,$$

where $f_{i,r}$ is the rth realization of the feature f_i and R is the total number of realizations.

An important drawback of the Pearson correlation index is that it can exhibit positive or negative linear relationship also for the case of nonlinear relationship between the variables under analysis. In order to address this issue, we have also used the Distance Correlation index. The Distance correlation index between f_i and t is defined as

$$D_i = \frac{\mathrm{dCov}_n(f_i \, t)}{\sqrt{\mathrm{dVar}_n(f_i)\mathrm{dVar}_n(t)}} \ , \tag{4.2}$$

where $\mathrm{dCov}_n(\,\cdot\,)$ and $\mathrm{dVar}_n(\,\cdot\,)$ are the sample distance covariance and the sample distance variance, respectively. For two variables x_r and y_r observed for R repetitions, the sample distance covariance is computed as

$$\mathrm{dCov}_n(x, y) = \sqrt{\frac{1}{R^2} \sum_{r_1=1}^{R} \sum_{r_2=1}^{R} \|x_{r_1} - x_{r_2}\| \|y_{r_1} - y_{r_2}\|} \ , \tag{4.3}$$

whereas the distance variance for a variable x_r observed over R repetitions is given by

$$\mathrm{dVar}_n(x) = \mathrm{dCov}_n(x, x) \ . \tag{4.4}$$

Both Pearson and Distance correlation indexes, however, may give inaccurate results when not enough samples are present in the population under study. For this reason, we adopt a third statistical descriptor, the RreliefF algorithm [65], aimed at finding a possible regression between two variables. The RreliefF algorithm uses the regression model to define a ranking and weight factor for each feature. It can therefore be used as a correlation index. Basically, the RreliefF algorithm is based on the probability that the predicted values of two instances are different. More specifically, given a feature f_i, its weight is defined as the probability of predicting different values of the feature, given the nearest instances and two different values of f_i. An exhaustive description of the algorithm can be found in [65].

4.3 Materials analysis

In order to analyze the materials used in the violin making process, a sampling is required. A sampling performed directly on the violin is certainly not possible without compromising the acoustic properties of the violin. For this reason we decided to prepare a couple of wooden board section (one for the top-plate and one for the back-plate) taken from the same piece of wood of the violin, on which the maker performed all the same treatments applied on the violin used in this thesis, i.e. the ground coat application and the varnishing. These treatments have taken place at the same time both for the violin and the wooden boards, in order to keep the environmental conditions equals for the two. At the end of the entire making process of the violin, for each board we have a total number of nine wood samples, three for each step : the violin in white, after the application of the ground coat and after varnishing.

4.3.1 Samples preparation and analysis techniques

The sampling was realized on the two boards for each step. After that, the samples were embedded in a block of acrylic resin which has a "support" function, i.e. allows the subsequent proper sectioning using a diamond blade and polishing of the sample with carbon papers. The resulting polished cross-sections, visible in Figure 4.5, were first observed at the optic microscope exploiting the polarized light presented in Section 3.3.1. For this analysis we used a Leitz Laborlux polarizing microscope (VIS and UV lamps). This observation allowed to inspect the level of the ground coat penetration into the wood.

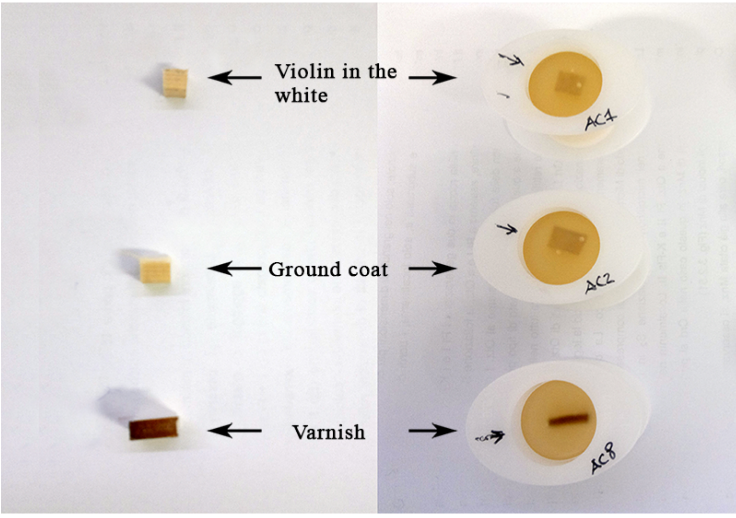


Figure 4.5. On the left the samples taken from one of the two boards. On the right the same samples insrted in the resin ready for sectioning and polishing.

The cross sections previously observed with the polarized light microscopy have been sputtered with an Au coating, using a Cressington 208HR sputter, in order to make the surface conductive. Scanning Electron Microscopy (SEM) images and energy-dispersive X-ray spectra (EDX) were collected by using a Tescan FE-SEM instrument (MIRA XMU series) equipped with EDAX spectrometer. The micro-analyses were performed maintaining the current electron beam at 20 kV, with counts of 100 s per analysis. The semi-quantitative data were obtained by processing the measurements with the EDAX Genesis software.

Chapter 5

Results and evaluations

In this chapter we will show the experimental results based on the methodology presented in Chapter 4. As explained before, the analysis were performed separating three steps that make the varnishing process and which are illustrated in Table 5.1. For each step several performances were executed by the skilled player. These performances are presented in Table 5.2.

Since a large number of audio file was analyzed and a lot of features were extracted, only the most relevant results will be presented. The Temporal Features (see Table 3.1) were discarded due to their strong dependence on the human execution. This problem occurred even if the performances were executed by a skilled player. The Spectral Spread, Spectral Skewness and Spectral Entropy exhibited no changes between the three steps, large enough to be considered significant.

The first part of this chapter is related to the results on the analysis of the varnishing process, presented in section 4.1. In section 5.2.1 we will analyze the varnishing process starting from the energy produced by the violin in different steps; in section 5.2.2 we will use the frequency response tool, comparing our results with the literature; afterwards, in section 5.2.3 we exploit the audio features that reduce the audio spectrum, comparing our results with the literature; in section 5.2.4 the high frequencies will be analyzed through the related audio features; finally, in 5.2.5 we show the results about the dissonance property. An overview on the results is illustrated in Table 5.3.

The second part of this chapter (section 5.3) shows the results on the analysis of the string tension. As we anticipated in 2.4, in this case there are no studies in the literature to compare with our results. We use both the audio feature extraction tool and the frequency responses. At the end of the section we show the results about the correlation analysis.

STEP INDEX	DESCRIPTION
t1	Violin in white (all the wooden parts glued together)
t2	Violin with ground coat (prepared for the varnish)
t3	Violin varnished

Table 5.1. Varnishing steps analyzed.

STRINGS INVOLVED	PLAYED NOTES	LABEL
G	open string	G*
	fith note	GV*
	tenth note	GX*
D	open string	D*
	fith note	DV*
	tenth note	DX
A	open string	A
	fith note	AV
	tenth note	AX
E	open string	E
	fith note	EV
	tenth note	EX
G - D - A - E	major scale	scale**
G - D - A - E	song 1	s1**
G - D - A - E	song 2	s2**

Table 5.2. List of the performances recorded for each step illustrated in Table 5.1. For each performance is indicated which strings are involved and which notes were played. The labels associated to the performances will be used in the rest of this work.

*The note was played six times for each step. Three times bowing in up direction and three times bowing in down direction.
**It was played one time for each step.

5.1 Material results on the varnishing analysis

In this section we present the results obtained from the analysis performed on the materials employed in the varnishing process. The related methodology is explained in section 4.3. In order to allow a destructive sampling, two wooden boards were used: the first is made of spruce and the second is made of maple. These two boards were taken from the same pieces of wood and respectively employed for the construction of the top-plate and the back-plate of the violin body.

The varnish used in this work is alcohol-based and its recipe is illustrated in Table 5.4.

The shellac is an animal resin and it is the main element of the varnish. For its nature, the shellac is a natural polymer and it is chemically similar to synthetic polymers. Therefore, it can be considered a natural form of plastic. The dry flakes of shellac are dissolved in alcohol to apply it on the violin wood. In order to enhance its elastic properties, elemi and the oil of spike are added. In particular the last creates a thin plastic film. Since the shellac has a high reflective coefficient, just for aesthetic purposes, the bonzoe is used to opacify the violin. Propoli and other coloring substances can be arbitrarily added to get a characteristic color.

For the reasons just presented, the varnish behaves as a plastic material with elastic properties and this explains several results of this work.

The results that come out from the materials analysis are almost identical for both the spruce and the maple boards. In order to confirm this equivalence, for the sake of example we show the Figure 5.1 related to the EDX analysis (see section

	Materials tool	Audio Tool	Step	Effect	Prior study
ENERGY	PLM, SEM	RMS, ASE	Ground Coat	Energy reduction for small quantity	
MEDIUM FREQUENCIES		ASE	Ground Coat	Formant A (700-1200 Hz) damped	Bazant [4]
		ASE	Varnish	Formant A (700-1200 Hz) enhanced	Bazant [4]
		ASE	Varnish	Gap G2 (1200-1800 Hz) damped	Bazant [4]
BRIDGE HILL (MED. FREQ.)		Freq. Spectrum	Ground Coat	Additional peaks in 1.2 - 2.5 kHz	Jansson [3]
		Freq. Spectrum	Varnish	Harmonics in 1500-4000 Hz shifted forward	Jansson [3]
		ASE	Ground Coat	Formant E2 (1800-3000 Hz) damped	Bazant [4]
		ASE	Varnish	Formant E2 (1800-3000 Hz) slightly enhanced	Bazant [4]
		Specral Contrast	Ground Coat	Additional peaks in 1.6 - 3.2 kHz	
LOW FREQUENCIES		Freq. Spectrum	Varnish	Low frequency shifted forward	Meinel [2]
		Freq. Spectrum	Ground Coat	Low frequency damped	
		Freq. Spectrum	Ground Coat	1st Air resonance (280 Hz) shifted forward	Bazant [4]
		Freq. Spectrum	Ground Coat	1st Valley above 300Hz damped of 3dB	Bazant [4]
HIGH FREQUENCIES		Freq. Spectrum	Varnish	High frequency damped	Meinel [2]
	SEM	Freq. Spectrum	Ground Coat	Additional peaks in 5 - 8 kHz	
		ASE	Varnish	G4 (> 4000 Hz) damped	Bazant [4]
	SEM	Centroid, Rolloff, Brightness	Ground Coat	High frequencies enhanced (calcium effect)	
DISSONANCES		Roughness	In the white	Presence of many large lobes. High Roughness.	
		Inharmonicity	Varnish	Harmonicity enhanced	

Table 5.3. The table shows the list of the results on the varnishing process analysis.

INGREDIENT	MAIN CHARACTERISTCS	QUANTITY
Shellac	Main component of the varnish. Natural polymer (plastic).	100 g
Benzoe	Opacifies the varnish.	10 g
Elemi	Elasticises the varnish.	8 g
Propoli	Coloring substance.	10 cc
Oil of spike	Helps the application of the varnish. Remains soft over the years.	5cc
Alcohol	Solvent.	50cc

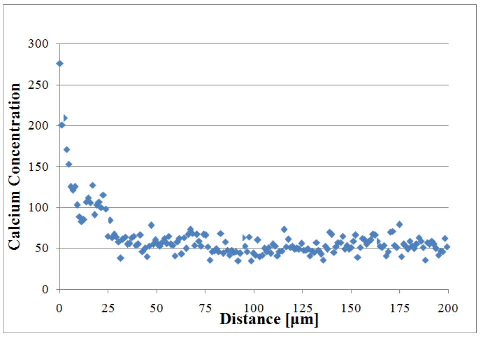
Table 5.4. The main characteristics of the ingredients included in the varnish recipe used in this work.

3.3.2). The Figure illustrates the amount of the calcium concentration recognized after the ground coat application. The trend of the concentration is almost the same for the two cases. Indeed, the differences are enough small to be neglected. Therefore, only the spruce board will be taken into account in this chapter.

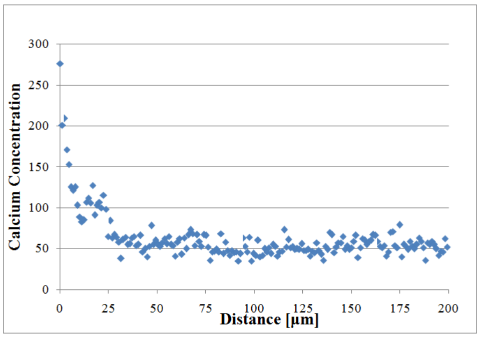
The Polarized Light Microscopy (PLM) analysis on the materials, presented in 3.3.1, allows to recognize the new layer of varnish lying on the wood after that the varnishing process was completed. The Figure 5.2 shows the spruce sample before and after varnishing with a magnification of 20x.

The layer identified includes both the ground coat and the varnish. In order to quantify the thickness of the ground coat and the varnish layers, the SEM/EDX analysis presented in section 3.3.2 has been performed. The Figure 5.3 has a magnification of 498x and illustrates the new layers after both ground coat and varnish applications, indicating the measures of the two.

The average of the approximated thickness of the ground coat and the varnish

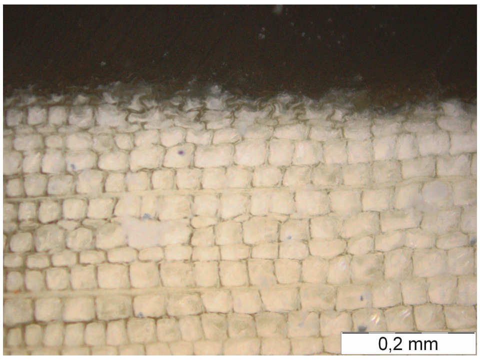


(a) Spruce sample

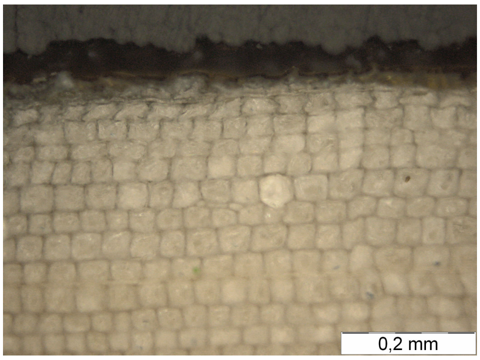


(b) Maple sample

Figure 5.1. The percentages of calcium recognized with the EDX analysis after the ground coat application. The values are analyzed starting from the top and going towards the inner wood. The Fig. 5.1 (a) shows the amount of calcium related to the spruce sample and the Fig. 5.1 (b) shows the amount of calcium related to the maple sample.



(a) Violin in the white



(b) Violin after varnishing

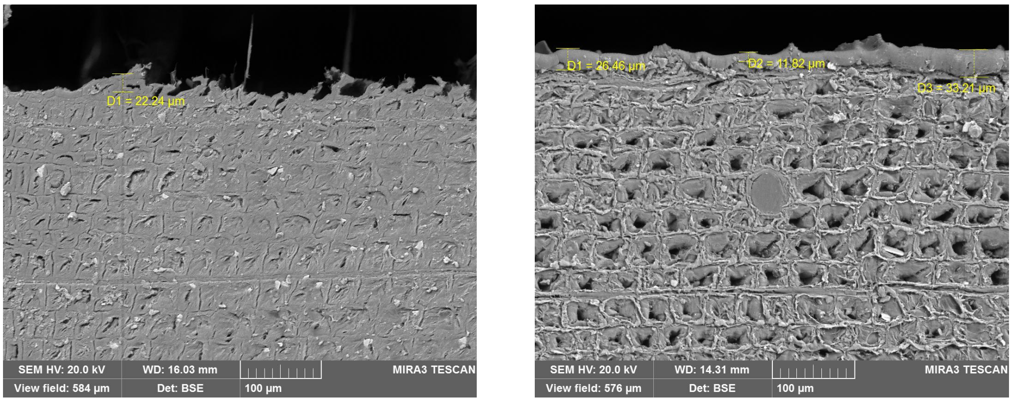
Figure 5.2. The Fig. 5.2 (a) shows the original wood aspect. The Fig. 5.2 (b) shows the sample after the varnishing step.

is respectively of $20\mu m$ and $25\mu m$. We can affirm that the ground coat lying on the wood is a very small quantity. Indeed, it corresponds to the 1% of the plate thickness. The same observation can be done for the varnish.

We explained in section 3.3.2 that the EDX analysis can give some elemental information. The Figure 5.4 shows the EDX Spectrum obtained from the electron microscopy analysis.

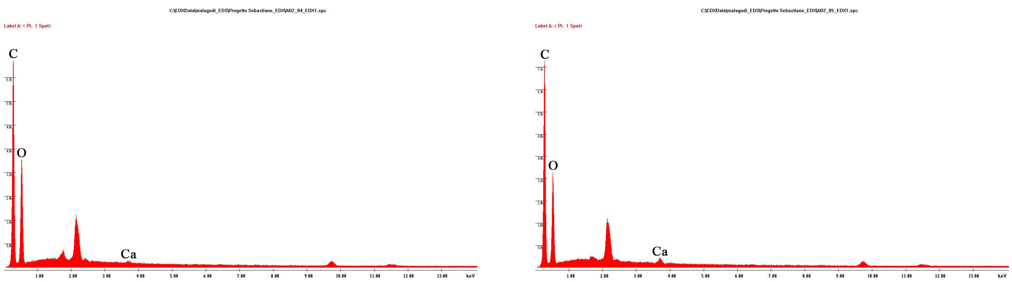
Looking at the spectrum after the ground coat application, can be noticed that there's a larger quantity of calcium (Ca) rather than the violin in the white. This is due to the substance used for the ground coat called casein that contains a large quantity of this element. The calcium has a well-known effect of hardening. The harder the material, the greater are the vibrations at the high frequencies. This explain the high frequencies enhancing performed by the ground coat.

In addition, the calcium creates a layer that after drying exerts a restriction on the underlying material. This confirm what we will see about the reduction of the overall energy in section 5.2.1.



(a) Sample with ground coat **(b)** Sample with varnish

Figure 5.3. The Fig. 5.3 (a) shows the thickness of the ground coat layer. The Fig. 5.3 (b) shows the thickness of the varnish layer.



(a) Violin in the white **(b)** Ground coat

Figure 5.4. In Fig. 5.4(a) the EDX spectrum of the violin in the white. In Fig. 5.4(b) the EDX spectrum of the violin after the ground coat application.

5.2 Audio results on the varnishing process

In this section we discuss the results of the timbral analysis on the varnishing process. In order to immediately inspect the frequency content of our recordings, the first audio analysis tool that we will use is the frequency responses (see section 3.1). The second audio tool we will use is the features extraction, in order to get in a compact form several specific cues on the timbral content (see section 3.2). Whenever is possible, the audio analysis tool will be correlated with the materials analysis.

The setup for the acquisition data and the tools used was described in the Chapter 4.

We explained previously that 53 features were extracted for each performance. The averages of these features were computed across the performances values, in order to get a single vector of 53 features for each step. The Figure 5.5 gives an overview about the audio results, showing the values obtained after this calculation.

On the x-axis are showed the indexes of all the features extracted. The correspondences between the indexes and the features are illustrated in the caption of

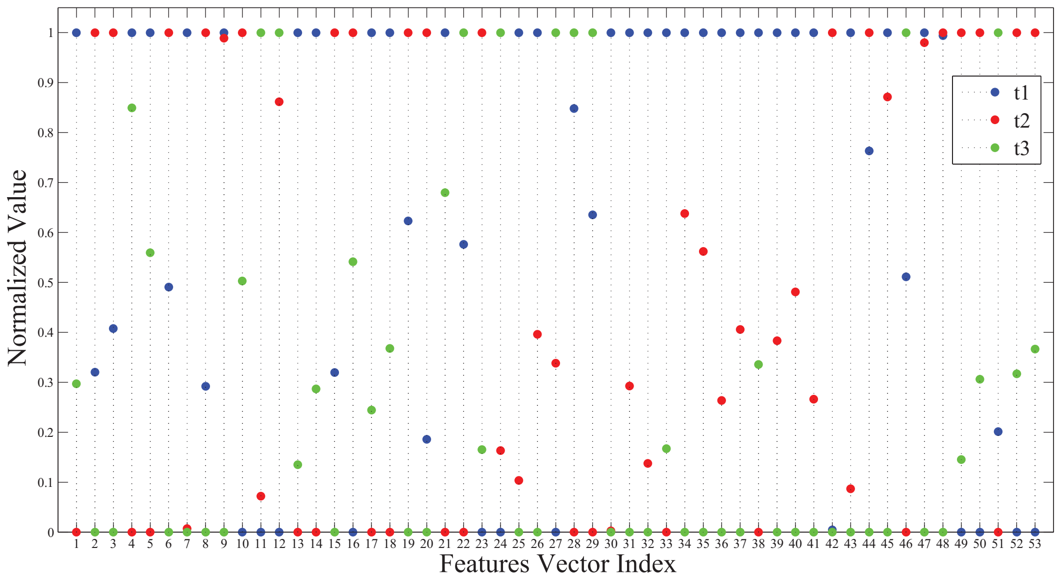


Figure 5.5. All the features averaged across the executions. The features are indexed as: 1 - RMS, 2 - Spectral Centroid, 3 - Spectral Spread, 4 - Spectral Skewness, 5 - Spectral Kurtosis, 6 - Flatness, 7 - Roughness, 8 - Spectral Entropy, 9 - Inharmonicity, 10 - Brightness, 11 - ZCR, 12 - Attack Time, 13 - Attack Slope, 14 - Attack Leap, 15 - Spectral Rolloff, 16 - Spectral Irregularity, 17..29: MFCC, 30..53 - Spectral Contrast. All the features are presented for each steps of the varnishing process: t1 - violin in white, t2 - violin with ground coat, t3 - violin varnished.

the figure under analysis. On the y-axis are showed the normalized values related to each feature. As we explained before, each feature value is obtained after the calculation of the average across all the performances values.

The aim of this first representation of the features is to give an overall idea of the trend of each feature for the three steps of the varnishing process, although very often will be necessary to focus separately on each single performance. Indeed, in section 2.2.2 we asserted that the spectrum of the violin varies in a significant way from note to note.

Łukasik shows in [22] that considering a single feature at a time, the related values for each of the four strings can be completely different. For this reason, it is not always possible to get considerations looking at all the notes and scales/songs at once. In the rest of the chapter we will focus on the single performance or on the all performances, regarding to the kind of analysis to be performed.

5.2.1 Energy

In this section we analyze the energy through the audio feature RMS, presented in sec. 3.2.1. For this analysis it is possible to get an overall idea of the differences between the steps looking at all the played notes at once.

The Figure 5.6 shows the RMS results compared for the three steps of varnishing. The ground coat and the varnish make up a new layer of material that limits the vibrations of the wood. For this reason an overall damping effect can be expected. This layer was identified and quantified by the PLM analysis and the related results are presented in section 5.1. Even if the thickness of the layer resulted in a very small amount, in Figure 5.6 it is possible to note how the ground coat entails a significant reduction of the energy. On the other hand, no significant differences in terms of energy come out between the ground coat step and the varnish step. The RMS values of the last two steps are on average equal in most of the notes.

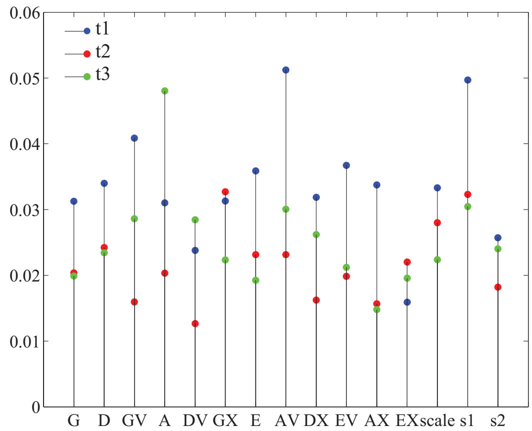


Figure 5.6. The RMS values computed for each note and scale/song compared in the three steps. The notes are ordered by frequency. The feature is presented for each steps of the varnishing process: t1 - violin in white, t2 - violin with ground coat, t3 - violin varnished. The two songs played are indicated by s1 and s2.

According to our results we can affirm that, although the maker has applied a very small quantity of ground coat, this produces a significant reduction of the energy emitted. Instead, the varnish has minor effect in this regarding.

5.2.2 Frequency Responses Analysis

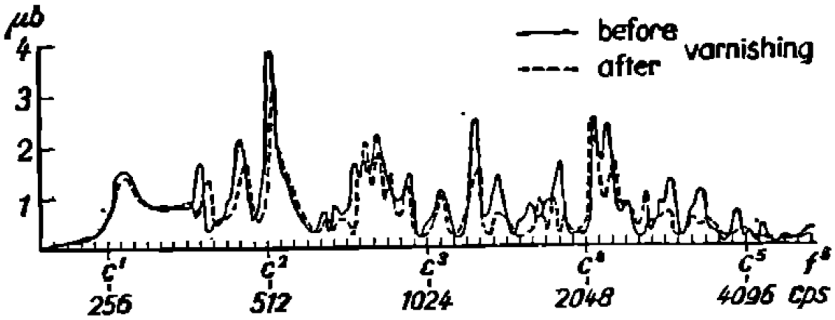
In this section we perform an analysis extracted from several frequency responses computed on our recordings.

We previously mentioned the study conducted by Meinel [2], in which he investigated the acoustical effects of the varnish on a violin. The frequency response that he obtained is showed in the Figure 5.7(a). It can be noticed that, after the varnishing (dashed line) the low frequency are carried forward and the high frequencies are damped.

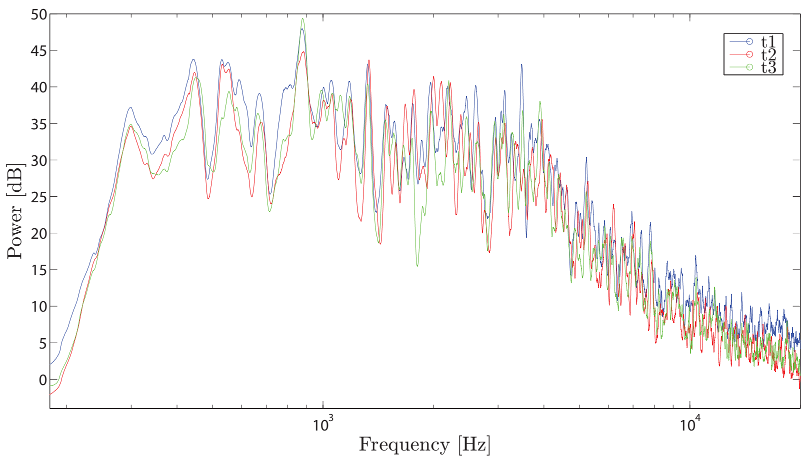
We want to compare these effects with our results and to do this we computed the frequency responses of the major scale and the two songs played by the musician. Afterwards, we averaged the resulting three audio spectrums in order to get a more complete response. The Figure 5.7(b) shows the resulting responses computed for each varnishing step.

The same observations made for the Meinel's study can be done for our results. Comparing the violin in white (blu line) and the violin after last varnishing step (green line) is possible to note a forward shifting of the low frequencies and a damping at the high frequencies.

In addition, we can do several observations for the spectrum computed after the ground coat application. It resulted in a damped low frequency range, but without



(a)



(b)

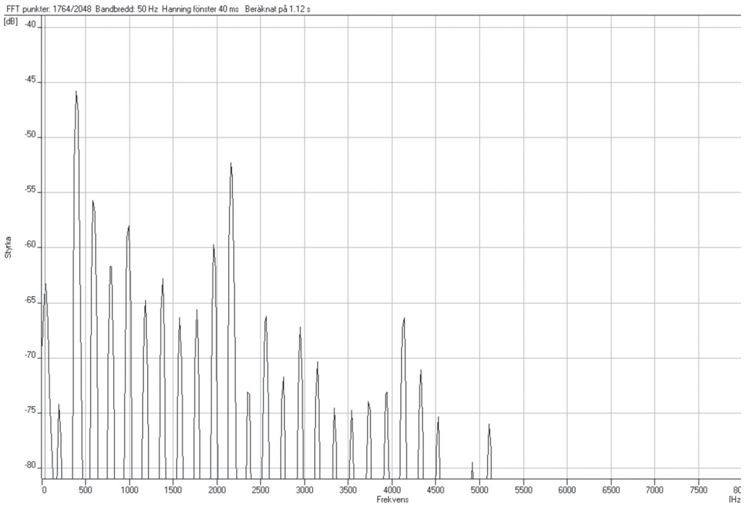
Figure 5.7. The Fig. 5.7(a) shows the frequency spectrum measured by Meinel in [2] before and after the varnish application. The Fig. 5.7(b) shows the frequency spectrum resulted from our analysis: t1 - the violin in white, t2 - the violin with the ground coat, t3 - the violin varnished.

pushing it forward as for the varnish. Finally, the ground coat introduced a lot of additional peaks over the frequency ranges 1800-3000 Hz and 5000-8000 Hz.

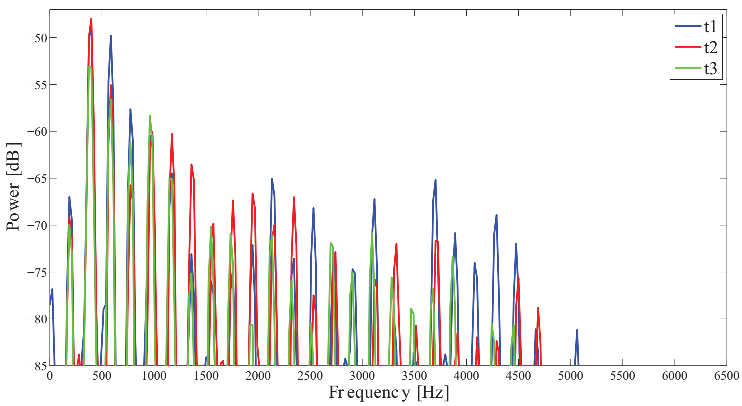
At this point we can do a general consideration that will be confirmed after. The ground coat enhances the high frequencies and the varnish damps it. Indeed, we discussed in section 5.1 about the hardness of the ground coat that introduces higher frequencies and about the elastic properties of the varnish that result in a light damping effect, which is incisive on the high frequencies. Finally, the MFCC coefficients are not presented. We anticipated in section 3.2 that the reduction process of the spectrum components entails the loss of spectral information. Looking at our results this reduction has been too much large for our purposes.

The former range includes the bridge-hill, discussed in chapter 2. For this reason is very important for the timbre quality perception. This statement is confirmed by Jansson in [3]. First of all, he identified the air resonance (the frequency prominence around 300 Hz related to the Helmholtz resonance, see section 2.1.2) and the wood resonances (the frequency prominence around 500 Hz related to the Helmholtz resonance, see section 2.1.2) for the frequency response computed on a Guadagnini 1778. In the same way, in Figure 5.7(b) is possible to recognize the air resonance represented by a prominence slightly below 300 Hz and the two prominences called wood resonances before and after 500 Hz. The Figure 5.8(a) shows the spectra of the sound produced by the Guadagnini. It can be compared with the Figure 5.8(b) that illustrates the frequency response obtained from our results for the same note.

In our case we miss the second maximum present in the Jansson figure just above 2 kHz, due to the different resonant characteristics of the wood used in this work. The rest of the spectral content has the similar characteristic for the two figures. The ground coat enhanced the frequency over the range 1200 - 2500 Hz, including the bridge-hill. From this analysis come out also that the varnish shifted backward and damped the harmonics over the range 1500-4000 Hz, due to the elastic properties discussed in section 5.1. The last two effects operate on bands classified by Jansson with maximum influence on the quality evaluation.



(a)



(b)

Figure 5.8. In Fig. 5.8(a) the frequency spectrum of the open string G (196 Hz) produced by the 1778 G.B. Guadagnini in [3]. In Fig. 5.8(b) the spectrum of the open string G (196 Hz) produced by the violin used in this work, plotted for each step: t1 - violin in white, t2 - violin with the ground coat, t3 - violin varnished.

Looking at the Meinel and Jansson studies compared with our results, we realized that the ground coat has a considerable effect over the bridge-hill frequency region. The varnish acts on the low and high frequencies. The former are moderately shifted forward ($\sim$ 10 Hz) and the latter are damped and shifted backward. In addition, the varnish shifted backward the bridge-hill harmonic content. We will reconfirm these observations later, when we will focus on the feature based analysis.

5.2.3 Feature-based analysis: Reduced Frequency Spectrum

In [4], the authors reduced the frequency response in 1/12 octave bands and identified eight frequency ranges of interest. Each range represents a *formant* region or a *gap*. The formant is identified by prominences in the frequency spectrum and the gap by the space that separates them. The table 5.5 shows the labels assigned to the formants and the gaps, and illustrates them in terms of frequency and number of 1/12 octave bands. The formants and the gaps are identified by the same label names used in the Bažant et al. work.

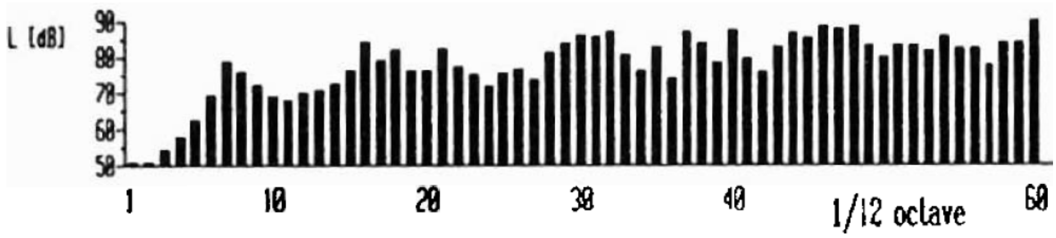
Band No.	Freq. bands (Hz)	1/12 Octave No.	Formant	Gap
1.	190 - 600	1 - 21	I1E1	
2.	600 - 700	22 - 23		G1
3.	700 - 1200	24 - 32	A	
4.	1200 - 1800	33 - 39		G2
5.	1800 - 3000	40 - 48	E2	
6.	3000 - 3300	49		G3
7.	3300 - 4000	50 - 53	I2	
8.	4000 -	54 -		G4

Table 5.5. The eight frequency bands identified in [4].

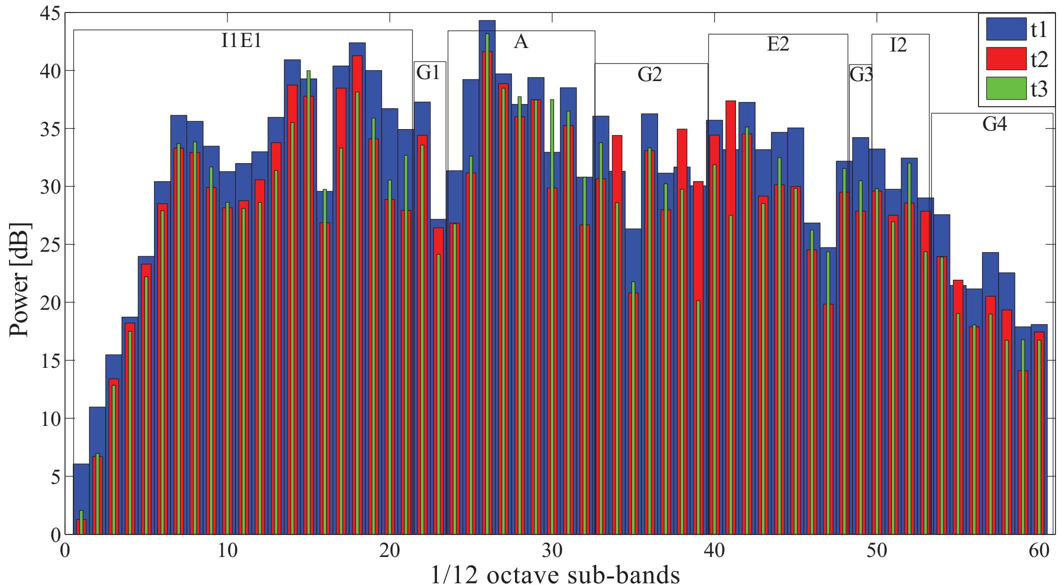
Bažant et al. come out with several rules from their study that should be respected by a violin of good quality: the first air resonance (see section 2.1.2) should be lower than 280 Hz, the valley above 300 Hz shouldn't be very deep, the levels in bands IlEl should be quite stable as well as the levels in bands A, E2 and I2 and otherwise, levels in gaps G1, G2 and G3 should be as small as possible. The first two of these statements can be compared with our results in Fig. 5.7(b). The first air resonance is located more or less at 290 Hz for the first step (blu line). It has been worsened by the application of the ground coat (red line), shifting the resonance 10 Hz forward. The varnish hasn't introduced any effect in this regard. For what concerns the valley above 300 Hz, in the first step it is located at 340 Hz. Also in this case the ground coat has resulted in worsening bringing down the valley of 3 dB, whereas the varnish hasn't any effect. Therefore, for these two first rules asserted by Bažant et al. we affirm that the ground coat has a worsening effect and, on the other hand, the varnish hasn't any effect.

At this point, we analyze the above statements regarding the levels of the formant and gap regions. This is done taking into account the reduced frequency response that comes out from the Bažant et al. study, showed in Fig. 5.9(a).

The feature that corresponds to the reduction performed in Fig. 5.9(a) is the Audio Spectral Envelope (ASE) (see section 3.2.3) with r equals to 1/12. We



(a) Reduced frequency response



(b) ASE

Figure 5.9. In Fig. 5.9(a) the frequency response reduced in 1/12 bands computed by Bažant et al in [4]. In Fig. 5.9(b) the ASE computed on the averaged spectrum of the scale and the two songs. The ASE is showed for each of the varnishing step and for a better visualization, the width of the bars was diversified for each of them.

computed the ASE coefficients using the spectrum illustrated in Figure 5.7(b). In this way we obtain the spectrum as complete as possible, since a considerable number of notes were played along the whole neck. The Fig. 5.9(b) shows our results.

In general, all the formants and the gaps are recognized also in our results according to the table 5.5, excepted for the $G3$ (bin n. 49)region that doesn't represent a gap in our spectrum, even though we can identify another gap region just few bins backward (bins n. $46-47$).

We can observe that the violin in the white sounds louder in almost each frequency as we already stated for the RMS. On the other hand, comparing the ground coat and the varnish steps is possible to identify significant differences. In the formant A the varnish has a pleasant effect since contributes to enhance the frequencies with respect to the ground coat which applied a reduction. The $E2$ formant includes the bridge-hill prominence, so is a significant sub-band for the reasons discussed previously. In this formant the varnish has again a pleasant effect as for the A region, but in this case the ground coat has a stronger damping effect. The opposite situation comes out looking at $G2$ and $G4$ regions. The former gains benefit from the varnish according to the above rules that want this gap as deep as possible, the latter exhibit the known effects of the varnish [2] [11], i.e. a reduction of the high frequencies.

Another feature that performs the spectral reduction in sub-bands is the Spectral Contrast (see section 3.2.3). A significant variation between the violin in the white and the ground coat application can be identified. The interesting result is that this variation concerns only the fifth Spectral Contrast band. In order to validate this observation, we computed the variation percentage for each band between the first and the second step. These percentages are illustrated in table 5.6.

Band no.	Percentage variation
1.	-1.94
2.	-0.70
3.	0.33
4.	1.35
5.	11.5
6.	-0.08
7.	5.37
8.	3.94

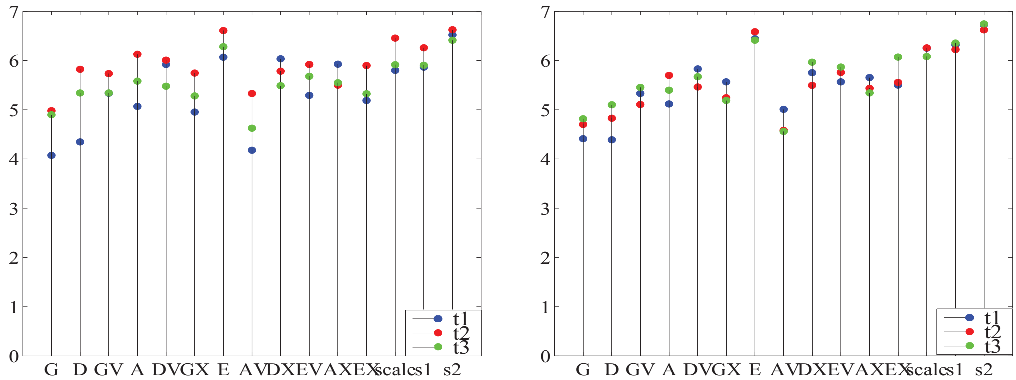
Table 5.6. Percentages of variation computed on the Spectral Contrast values rsulting from the violin in the white analysis and ground coat analysis.

The table shows that the fifth sub-band has a significant variation of 11.5%. The other percentages are all below the half of this value. This means that the ground coat makes more spiky the frequency region between 1600 Hz and 3200 Hz, due to the effect of hardening produced by the calcium (see section 5.1). This region matches the first of two indicated in section 5.2.2, where the same observation has been done analyzing the Figure 5.7(b). The Figure 5.10 shows the results for the Spectral Contrast band number five and, for the sake of example, also for the number six.

In conclusion, we confirmed in this section what has been stated previously about the energy. That is, the ground coat reduces the sound power compared to the untreated violin. Then, we have seen that the ground coat damps both the formant A and the formant E2, respectively in the ranges 700-1200 Hz and 1800-3000 Hz. At the same time, in the latter range (including the bridge-hill) the ground coat introduces additional spikes compared to the untreated violin. On the other hand, the varnish "restore" the frequency damped by the ground coat in the formant A. Finally, the varnish damps the gap G2 (1200-1800 Hz) and the very high frequencies ($>$ 4000 Hz).

5.2.4 Feature-based analysis: Spectral Features

Among the spectral features presented in sec. 3.2.3, we identified three of them which are related to the high frequencies: Spectral Centroid, Spectral Rolloff and Brightness. Looking at the Figure 5.5, it is possible to note that generally, these features have the same trend between the three steps of varnishing. So, we can affirm



(a) Spectral Contrast - Band 5

(b) Spectral Contrast - Band 6

Figure 5.10. The Fig. 5.10(a) shows the Spectral Contrast computed over the fifth band (1600 - 3200 Hz). The Fig. 5.10(b) shows the Spectral Contrast computed over the sixth band (3200 - 6400 Hz). The feature values are illustrated for each varnishing step: t1 - violin in white, t2 - violin with the ground coat, t3 - violin varnished. The two songs played are indicated by s1 and s2.

that the ground coat enhances the high frequencies. This observation is confirmed by the Fig. 5.11, that shows the values for each performance. The results obtained from the major scale and the songs, don't indicate any difference among the three steps. Conversely, analyzing each note at a time we can confirm our observation.

The ground coat has the effect of producing more high frequencies after its application and this is confirmed by the analysis on the materials. In section 5.1, we affirmed that the presence of the calcium results in a hardening of the body, contributing to the enhancement of the high frequencies.

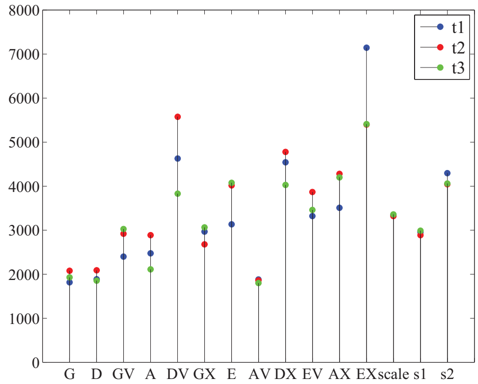
5.2.5 Feature-based analysis: Harmonic Features

Another important aspect that comes out from our results are the features related to the harmonicity and in particular for what concerns the dissonance. In Fig. 5.12 are showed the Roughness and the Inharmonicity.

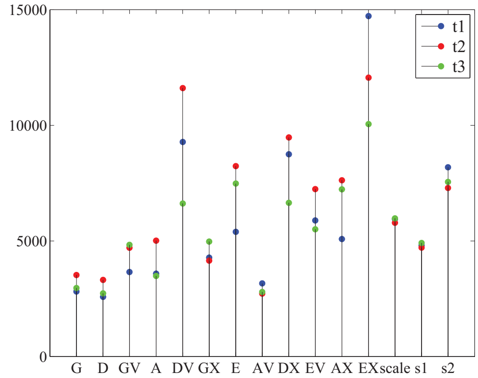
The violin in the white presents very higher values of Roughness rather then the ground coat step. This led back to the effect of constriction that the calcium has on the body, as explained in section 5.2.4. On the other hand, the values of Roughness for the ground coat and the varnish are almost the same.

Looking at the Inharmonicity, the most important effect is that of the final varnish. Indeed, the varnish has in almost all performances the lowest values. This can be immediately noticed in the Figure 5.5. The elastic properties of the varnish give a lighter damping effect rather than ground coat. This means that the strongest frequencies like the fundamental and its harmonics suffer from little reduction in energy. Conversely, the weaker frequencies outside these regions are damped. The consequence is a less inharmonic sound after varnish application. Of course, we must point out that the very high frequencies are included in the weakest. Therefore, the master maker needs to pay attention during the varnish application.

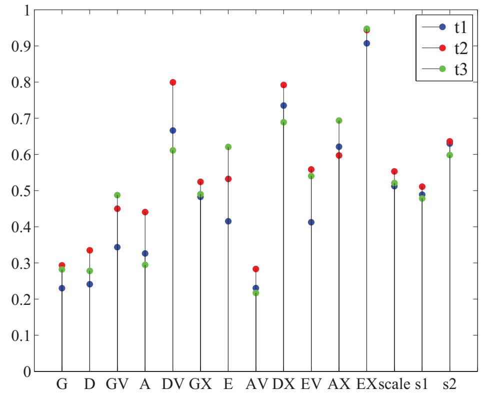
The inharmonic effect introduced by the ground coat and the opposite effect



(a) Spectral Centroid



(b) Spectral Rolloff



(c) Brightness

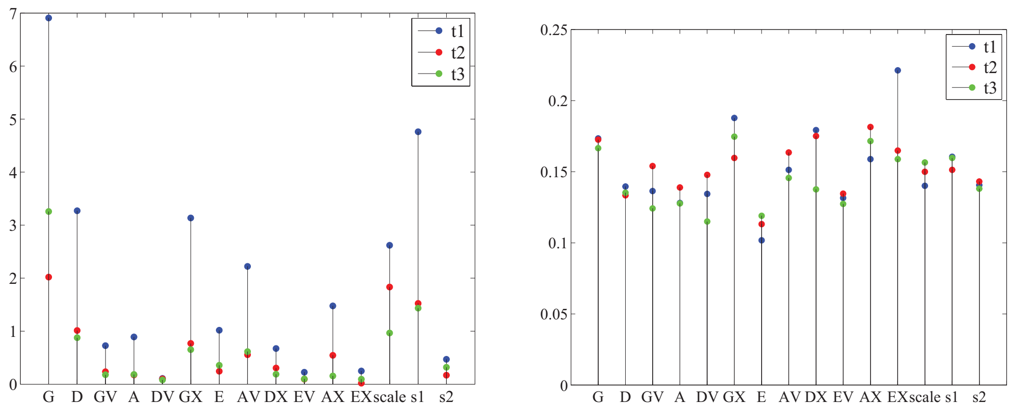
Figure 5.11. The three Spectral Features computed for each note and scale/song compared in the three steps. The notes are ordered by pitch. The values are illustrated for each varnishing step: t1 - violin in white, t2 - violin with the ground coat, t3 - violin varnished. The two songs played are indicated by s1 and s2.

introduced by the varnish are confirmed by the Spectral Kurtosis showed in Figure 5.13. In this case we analyze the feature looking at the single notes played at once. In particular, we focus on the four open strings played: G, D, A and E. After the ground coat application the Kurtosis values decrease, since several side peaks are introduced around the harmonics. Conversely, the varnish results in a larger values of Kurtosis, since it damps the smallest peaks and enhances the harmonics.

In conclusion, we can affirm that the level of dissonance can be managed by finding a trade-off between the ground coat and the varnish application. This is one of the most difficult challenges for a violin maker.

5.3 Results on string tension

In this section we discuss the results on the correlation between the tension and the violin family sound quality. These results have been presented in [66] which is presently under review. We introduced the physical role of the string tension related to the frequency in section 2.1.3. The importance of this parameter for its common



(a) Roughness **(b)** Inharmonicity

Figure 5.12. The two dissonance Harmonic Features. The notes are ordered by pitch. The values are illustrated for each varnishing step: t1 - violin in white, t2 - violin with the ground coat, t3 - violin varnished. The two songs played are indicated by s1 and s2.

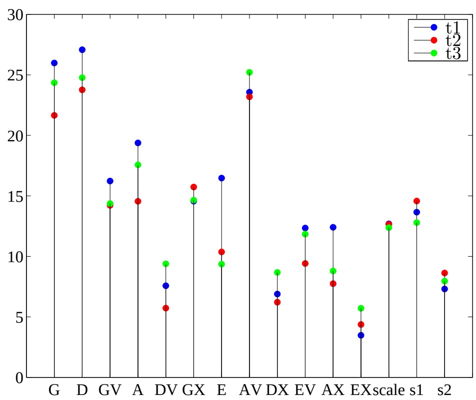


Figure 5.13. The Spectral Kurtosis. The notes are ordered by pitch. The values are illustrated for each varnishing step: t1 - violin in white, t2 - violin with the ground coat, t3 - violin varnished. The two songs played are indicated by s1 and s2.

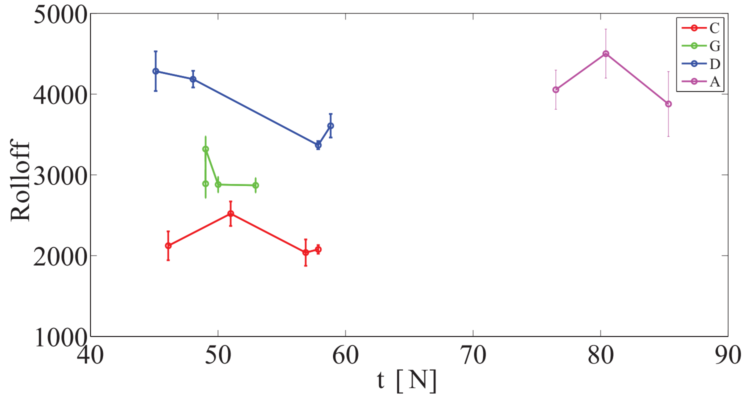
usage on the market and the methodology used for its analysis are explained in section 4.2.

The Figure 5.14, for the sake of example, shows the relation existing between Spectral Rolloff and Inharmonicity and the tension of the four strings.

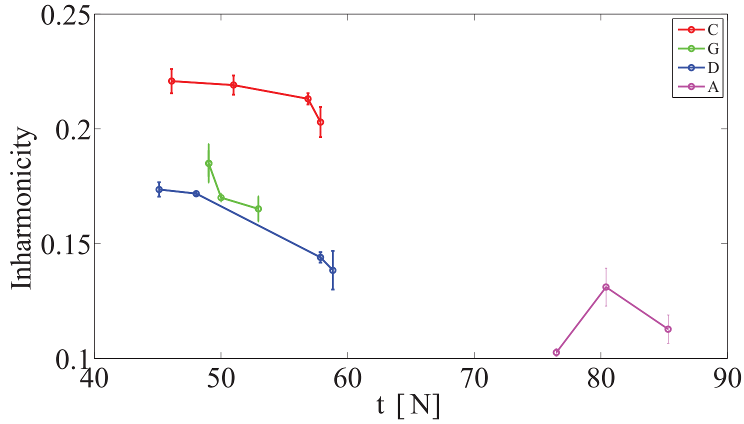
Notice that as the tension increases the Spectral Rolloff decreases. This is a well-known behavior [67]: in fact low-frequency harmonics are enhanced with higher values of the tension. The same behavior is observed for the Inharmonicity: in fact, by increasing the tension, the sound tends to become more harmonic.

In order to improve this concept, we take into account other two features on which we can make the dual observation stated for the Rolloff and the Inharmonicity: the Spectral Centroid and the Spectral Flatness.

In section 2.3.4 we provided several indications on the tonal qualities that differ between the three type of strings. Gut strings are tipically known for having a lot of harmonics and a lot of low frequencies. On the other hand, synthetic strings tend to



(a) Spectral Rolloff



(b) Inharmonicity

Figure 5.14. Spectral Rolloff and Inharmonicity as function of the tension of the string. The vertical lines indicate the error bars.

a warm sound, but having less harmonic content.

These statements are confirmed by our results. This time we will concentrate on the G string since it is the unique case for which we have the same tension for both the synthetic string (*Larsen*®) and the gut string (*Pirastro Eudoxa*®). The Figure 5.15(a) shows the Spectral Centroid computed for the G strings.

First of all, we can observe that increasing the tension the Spectral Centroid decreases, as for the Spectral Rolloff. Both features provide an indication about the high frequencies content, which results in a reduction. Afterwards, focusing on the *Larsen*® and the *Pirastro Eudoxa*® strings, we can note that for the same value of tension the Centroid computed for the gut string is lower rather than for the synthetic string.

The Figure 5.15(b) shows the Spectral Centroid computed for the G strings. The trend is the same as for the previous features analyzed. As the tension increases, the sound tends to become more harmonic and the spectrum becomes less similar to a white noise. Looking once again at the *Larsen*® and the *Pirastro Eudoxa*® strings, the gut string has more harmonic content rather than the synthetic string.

The Figure 5.16 shows the two spectrums of the *Larsen*® and *Pirastro Eudoxa*®

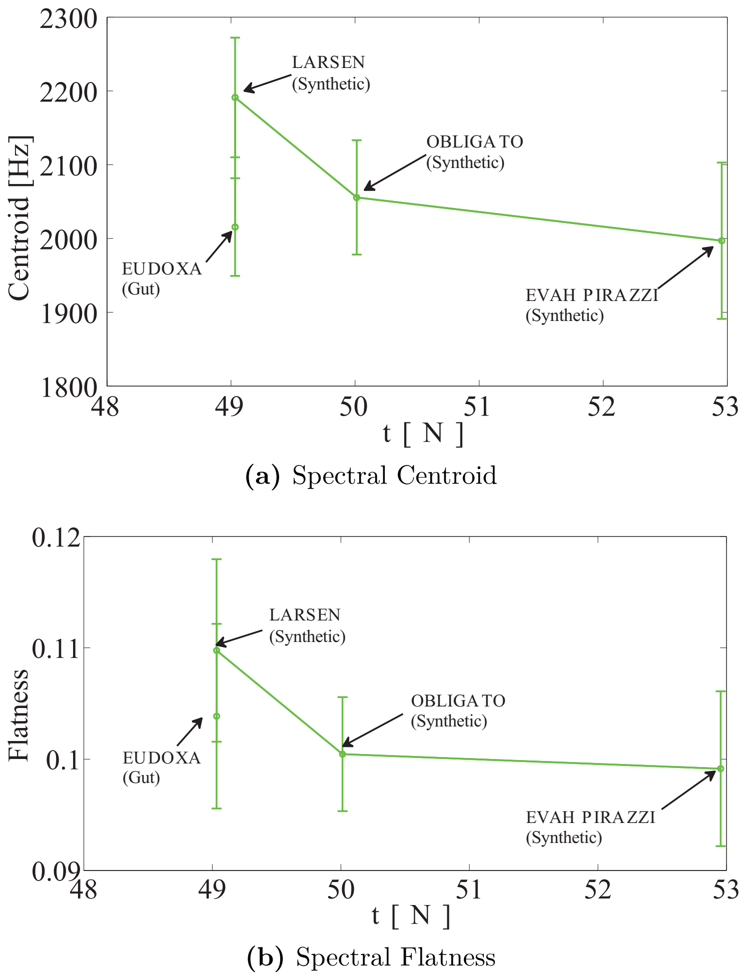
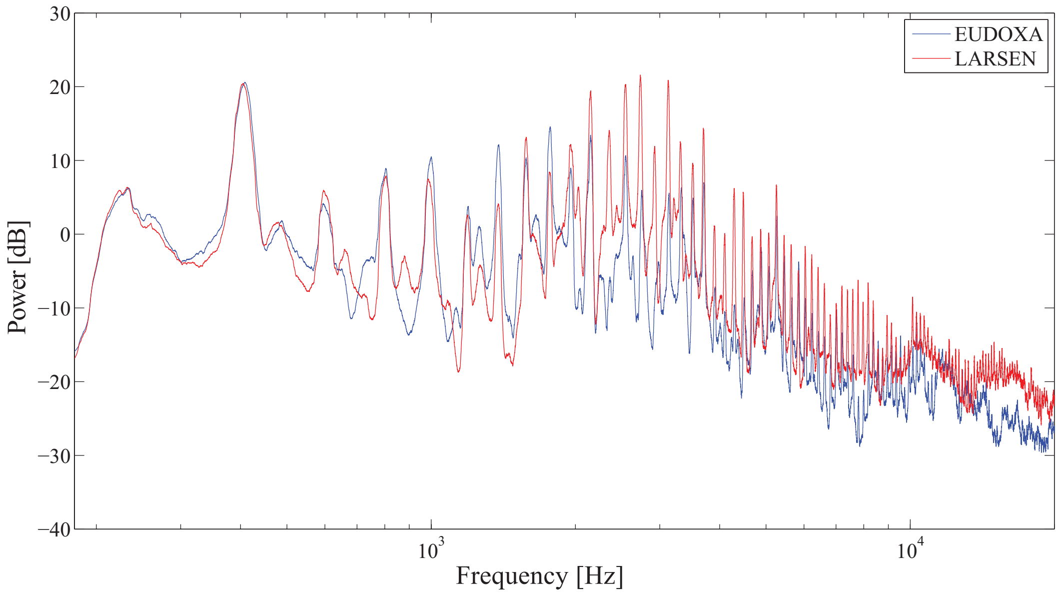


Figure 5.15. In Fig. 5.15(a) the Spectral Centroid as function of the tension of the string. In Fig. 5.15(b) the Spectral Flatness as function of the tension of the string.

G strings. The gut string shows a richer harmonic content, but presents less high frequencies. In order to better appreciate these observations, the Figure 5.16 shows a magnified version of the two spectrums, in the range from 2 kHz to 20 kHz.

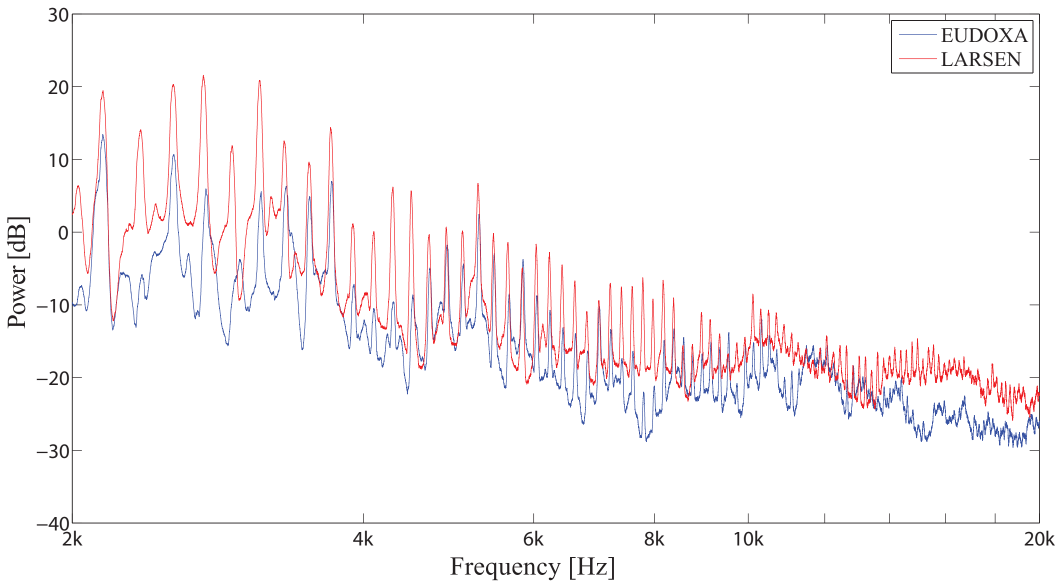
The features presented so far, are among those most correlated with the tension, regarding our analysis. In Table 5.7 we show the eight most correlated features, sorted by the Distance Correlation coefficient (see section 4.2.3).

Figure 5.18 shows the weights obtained with the RreliefF algorithm for the G string. In particular, the number above each bar represents the index of the feature specified in the caption. Notice that if we sort by the weights of the RreliefF, the absolute value of the Pearson Correlation or Distance Correlation coefficients, the ranking would remain almost unaltered, thus confirming the validity of the analysis. Even if just a few tension values are available, therefore, the Pearson and Distance Correlation coefficient provide consistent results for the considered experiment.



(a) Frequency response

Figure 5.16. Frequency response of the *Larsen*® and *Pirastro Eudoxa*® G strings.



(a) Frequency response

Figure 5.17. Frequency response of the *Larsen*® and *Pirastro Eudoxa*® G strings, magnified in the range 2 - 20 kHz.

C	sc22	flatness	spread	sc21	centroid	skewness	kurtosis	sc13
P_i	0.997	-0.984	-0.978	0.950	-0.945	0.923	0.924	0.944
D_i	0.998	0.983	0.976	0.964	0.957	0.947	0.946	0.940
G	sc23	centroid	sc17	kurtosis	sc2	sc24	flatness	mfcc10
P_i	0.9626	-0.958	-0.947	0.930	-0.931	0.894	-0.906	-0.884
D_i	0.966	0.959	0.954	0.948	0.939	0.919	0.913	0.900
D	entropy	skewness	kurtosis	inharm.	spread	flatness	centroid	mfcc11
P_i	0.700	0.994	0.991	-0.989	-0.985	-0.980	-0.911	-0.925
D_i	0.999	0.997	0.994	0.993	0.991	0.988	0.972	0.947
A	sc12	sc16	sc9	sc4	sc13	sc23	entropy	mfcc1
P_i	-0.998	-0.986	-0.977	0.966	0.966	-0.958	0.901	-0.856
D_i	0.999	0.988	0.979	0.971	0.971	0.965	0.914	0.920

Table 5.7. Pearson and Distance Correlation coefficients.

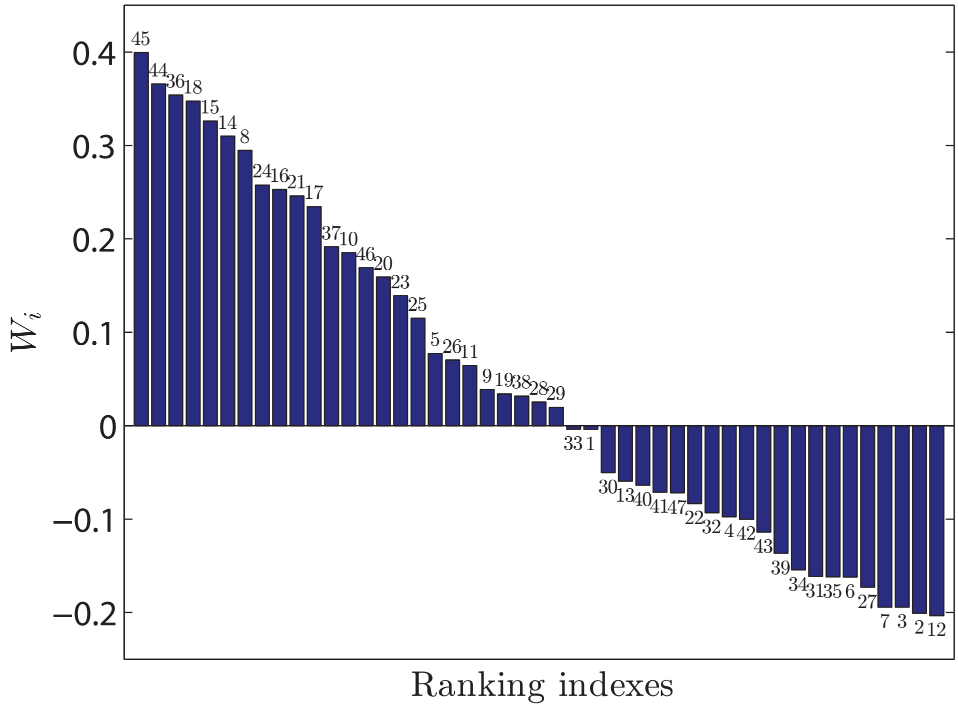


Figure 5.18. Weights computed with the RreliefF algorithm on the features extracted from the microphone placed on the body. The correspondences between bar indexes and features are: 1..13: MFCC, 14 - Spectral Centroid, 15 - Spectral Spread, 16 - Spectral Skewness, 17 - Spectral Kurtosis, 18 - Flatness, 19 - Spectral Entropy, 20 - Inharmonicity, 21 - ZCR, 22 - Spectral Rolloff, 23 - Spectral Irregularity, 24..47 - Spectral Contrast

Chapter 6

Conclusions and Future Works

In this thesis we proposed a new approach to estimate the violin sound quality through the use of a modern analysis tool: Audio Feature Extraction. The main purpose is to create an experimental setup that aims to evaluate the sound of the violin varying the constructive parameters through a novel approach.

In the first part of the work we wanted to inspect the effect that the varnishing process has on the final "voice" of the violin. In the same way, in the second part we analyzed the role of the strings related to the timbral qualities.

Several studies on the varnishing effects have been proposed in the literature. On one hand, most of them are strictly related to the changes of the physical/mechanical properties disregarding the human perceptive aspects. On the other hand, other studies inspected the sound quality according to judgments done by skilled people, disregarding the constructive characteristics. Whereas, at the best of our knowledge, no studies have been conducted on the timbral effects of the strings.

Our system exploits the audio feature extraction, which is a method that gives a perceptive-oriented approach. In addition, the audio features provide more compact and less redundant information rather than tools used in prior studies, like the frequency response. Thanks to these properties, the values obtained from the features extraction can be well used to correlate them with the material values.

For our purposes, the audio needs to be acquired with the minimal environmental influences. This can be obtained taking the performances in a semi-anechoic room flanked by a high-quality recording system. Once the data is acquired, the audio feature extraction can be performed and it provide three categories of information: Energy, Temporal and Spectral. Combining all these three aspects we obtained information on the timbre characteristic. Our system provides the features more correlated to the variation of the varnish or to different kind of strings.

The information obtained from our system revealed several new indications on the varnish influences. In addition to prior studies, we divided the varnish process in two different steps. The first is related to the preparation of the violin for the varnishing, through the application of the ground coat. The second is considered after the drying of the last hand of varnish. This information can be used to help the violin maker during the varnishing process, in which the choice of the substances and the methodology of their application makes the difference.

The results obtained revealed that the varnishing process affects in a significant way the violin tone quality. Moreover, the idea to separate the ground coat step from the varnish one has proved successful for recognizing different effects, that often compensate each other.

In the first step, when the violin is in white, the sound is very loud. On the other hand, the body is completely free to vibrate and due to the strong vibrations the sound is rough and not very harmonious.

The ground coat has a great role in the energy reduction, even if applied for small quantities. Indeed, the calcium contained in its recipe applies a constriction on the body once is dried. Furthermore, the calcium renders the wood surface more hard. This causes the introduction of a lot of harmonics in the bridge-hill region and enhances the highest frequencies. Finally, the ground coat damps a little bit the low frequency.

The varnish enhances the harmonics generated so far through the damping effect on their side lobes, reducing the inharmonicity. This is due to its plastic and elastic properties, which, however, at the same time dampen the highest frequencies.

Given these considerations, we can affirm that the maker could model his violin according to the effects presented, finding the right trade-off between the ground coat and the varnish application.

Our analysis on the strings revealed that their contribute to the variation of the violin tone quality is correlated to their tension. Strings with higher tension sounds more brilliant and more harmonic. In addition, we demonstrated that for the same tension value a string with a gut core sounds more harmonic and warm rather than a string with a synthetic core.

These results can supports the musicians for the choice of the set of strings or even better the mixing of different sets, in order to obtain a customized "quartet" of strings.

6.1 Future works

This work of thesis is part of a new research project started by Politecnico di Milano and Università di Pavia, which have their headquarter in Cremona at the Museo del Violino. It is the first study performed on the topics presented at the beginning of this chapter and has been developed with the early equipment available at the research center.

The results obtained by our system indicate that such a methodology offers a good potential. Certainly, the system needs be improved and can be extended to other materials involved. In particular, different recipes for the ground coat and the varnish could be employed. The varnishing process could be performed on different instruments of the violin family to validate even better our considerations. Finally, a larger number of strings could analyzed. Taking into account these improvements we identified four main directions that can inspire possible future works.

The first interesting project is to create a common description language for violin experts and makers: if we can define a common dictionary to illustrate the properties of a violin, certainly the audio features would be correlated. This process uses the audio features extracted, in order to categorize and standardize the acoustical and timbral properties of the violin, according to the common defined language: the sound quality can be defined with precise adjectives like "loudly", or qualitative features like "warm", which both can be used to describe the quality of a specific violin model. This strategy could be widely adopted among violin experts and not only.

Another inspiring research area concerns the integration of vibrational measurements to our acoustic and timbral estimates: as discussed in chapter 2, the dynamical coefficient and the internal friction of the violin wood plates is strictly associated to the varnish. It could be interesting to build a method to directly link vibrational and varnish measurements, in order to obtain a more accurate definition of the properties of the instrument body.

The third stimulating future objective is to create a predictive model using the values of the tension of the strings. This could results in a useful application for musician in order to choose their strings without the necessity of mounting them and wait for the settling time.

The final interesting research can be performed on the soundpost and the bridge settings. Several studies have been done in this directions, but also in this case could be profitable the introduction of the features extraction, in order to directly correlate the setting effects with the perceived violin tone quality.

Appendix A

Technical Specifications

Loudspeakers

Table A.1. Experimental setup: loudspeakers specifications.

Model	**Empire M2**
Total Power Output	RMS 3 W $\times$ 2
Total Harmonic Distortion	10%
Signal-to-Noise ratio	$\geq$ 90 dBA
Input Impedance	15 kΩ
Driver	$2''$, 8 Ω

Microphone

Table A.2. Experimental setup: microphone *Beyerdynamic MM1* specifications.

Model	**Beyerdynamic MM1**
Transducer type:	condenser (back electret)
Polar pattern	omnidirectional
Frequency range:	20 Hz to 20 kHz
Sensitivity:	15 mV/Pa (-36.5 dBV)
Max SPL for 1% THD:	128 dB
Signal-to-Noise ratio (rel. to $1Pa$):	57 dB
Nominal impedance:	330 Ω
Powering:	12 V to 48 V phantom power
Connector:	3-pin XLR
Dimensions (diameter):	9 mm(*head*)
Dimensions (length):	133 mm

Microphone Preamplifier

Table A.3. Experimental setup: microphone preamplifier *Aphex 188* specifications.

Model	**Aphex 188**
Gain:	26 dB to +65 dB
Input Impedance:	Variable
EIN:	−125 dBu
Frequency Response:	Not specified by manufacturer

A/D and D/A Converter

Table A.4. Experimental setup: *Symphony I/O* specifications (Analog I/0).

Model	**Symphony I/O**
Type:	32 inputs and 32 outputs Balanced through Apogee's proprietary Perfect Symmetry Circuitry
Level:	+4 dBu nominal / −24 dBu max. or −10 dBV nominal / +6 dBV max
Input Impedance:	10 kΩ
Output Impedance:	25 Ω

Table A.5. Experimental setup: *Symphony I/O* specifications (Analog to Digital).

Model	**Symphony I/O**
Frequency response:	1 Hz to 20 kHz, +0/ − 0.05 dB
Dynamic range:	120 dB (A-weighted)
THD+N @ +20 dBu	−113 dB (0.00024%)

Table A.6. Experimental setup: *Symphony I/O* specifications (Analog Out performance).

Model	**Symphony I/O**
Frequency response:	*dc* to 20 kHz, +0/ − 0.05 dB
Dynamic range:	129 dB (A-weighted)
THD+N @ 20 dBu	−117 dB (0.00014%)

Table A.7. Experimental setup: *Symphony I/O* specifications (Connections).

Model	**Symphony I/O**
Digital I/O Ports:	25-pin female D-sub connectors
	Port A: channels 1-8 I/O
	Port B: channels 9-16 I/O
	Yamaha pinout standard
Analog I/O Ports:	25-pin female D-sub connectors
	Analog In 1-8, Analog In 9-16
	Analog Out 1-8, Analog Out 9-16
	Tascam pinout standard
External Clock	75 Ω BNC word clock input and output
MIDI	1 In and 1 Out, 5-pin female DIN connectors

Appendix B

Database design

Requirements Analysis

The research center hosted in the Museo del Violino started a new project joining the competencies of Politecnico di Milano and Università di Pavia. The topics covered by Università di Pavia mainly concern the analysis of materials involved in the violin making process, while Politecnico di Milano aims to inspect the acoustics and timbre properties of musical instruments.

The aim of this databse is to support this project, gathering several types of information related to the violin. Initially, each violin is identified by: its name, a general description and a set of high quality photos.

Information on the materials involved are needed. In particular, we want to take trace of the species of wood and the varnish recipes used for the manufacturing.

A structural analysis is performed on each violin analyzed: a laser scanner is used to get a 3D-model stored in a cad file; an X-rays camera in order to get a related X-rays image. In addition, the design parameters are associated.

Through an acoustics analysis is possible to compute and preserve the acoustic pattern and the spectrums.

The acoustics analysis are done by recording several kinds of performances realized by one of the musicians collaborating for the project. A set of low-level feature will be extracted from the audio files acquired. Each feature is represented by a name and the correspondent value computed which for a subset of feature could be binary.

In addition, the performance considers also a vibrational analysis on which, similarly for the audio analysis, a set of vibrational features are extracted. The vibrational features have the same characteristics of the audio features.

In order to keep all the performance recordings, audio and vibrational takes are preserved. During the performances several kinds of audio takes are recorded in parallel: mono near and far field, array and soundfield. All kinds of the audio takes use a single microphone, apart the array take. In the last case, a set of microphones (typically 32 or 64) are used and they can be placed in different geometry assets, located in x-y-z coordinates. Indeed, the array are used for 3D audio purposes. For the same reason, video takes are recorded during each performance, using classical 2D and 3D techniques (e.g. Microsoft Kinect depth rendering). Finally, a spatial tracking of the violin is acquired and saved into a file.

After the performance is completed, the musician is asked to express its personal opinion about the violin played. Whenever the violin is a modern instrument, also the maker is asked to express his opinion on the quality of the instrument. Before start their collaboration, all the musicians and violin makers are registered in the database with their name and email address.

In the next section will be illustrated the ER model designed taking into account the requirements presented.

ER model

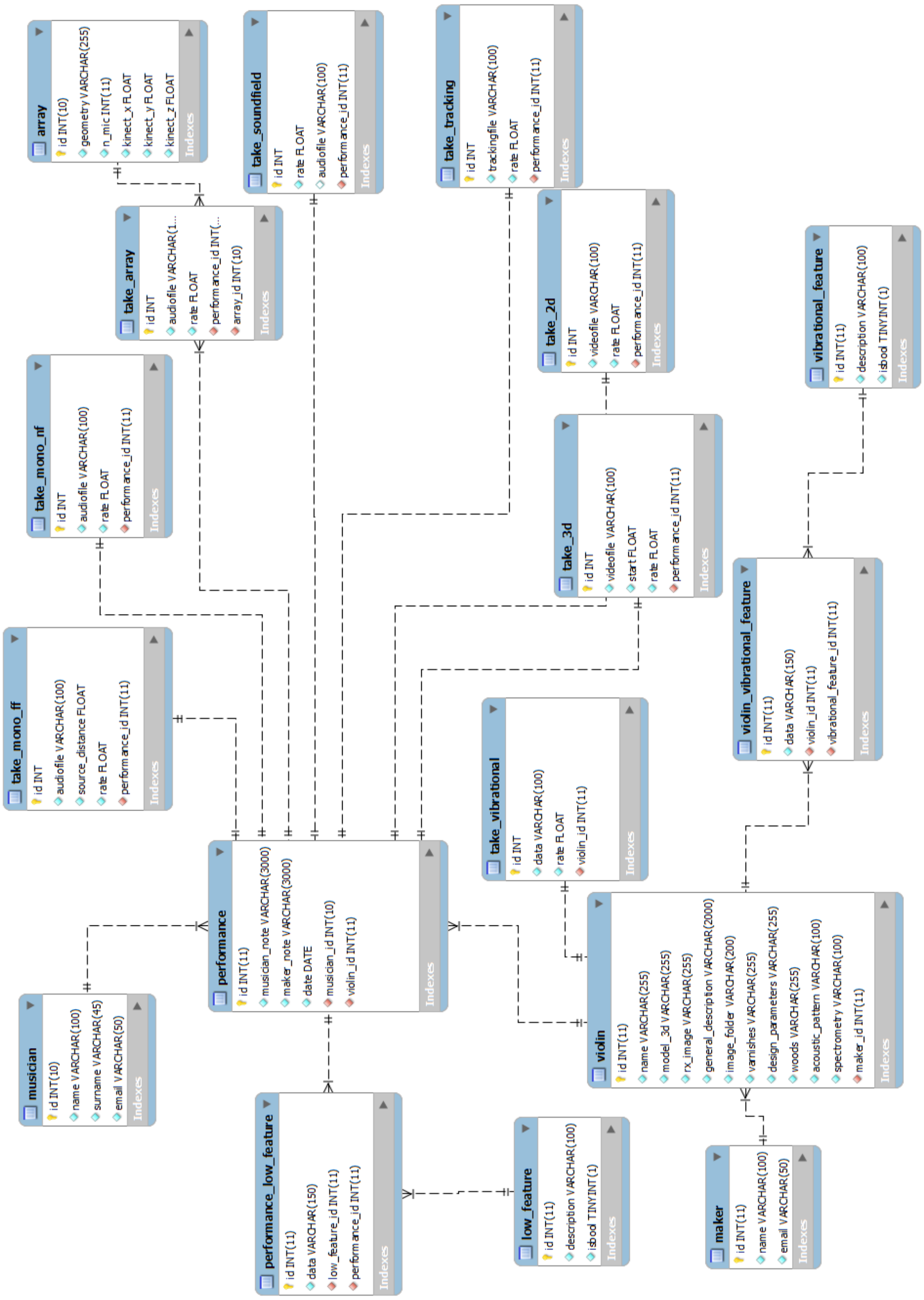


Figure B.1. The ER model designed with MySQL Workbench 5.

The Entity and Relationships identified are listed below:

- VIOLIN is an entity.
- MAKER is an entity. The maker can make different violins, so he has a 1:N relationship with the VIOLIN.
- VIBRATIONAL FEATURE is an entity. It describes a vibrational feature class.
- VIOLIN VIBRATIONAL FEATURE is a relationship. A violin has one or more vibrational features.
- PERFORMANCE is an entity. Each violin can belong to many performances. So the VIOLIN has a 1:N relationship with the PERFORMANCE.
- MUSICIAN is an entity. He has a 1:N relationship with the PERFORMANCE, since one musician can make one or more performances.
- LOW FEATURE is an entity. It describes a low-level audio feature class.
- PERFORMANCE LOW FEATURE is a relationship. A performance has one or more low level audio feature.
- TAKE MONO FAR FIELD is an entity. It has a 1:1 relationship with a performance.
- TAKE MONO NEAR FIELD is an entity. It has a 1:1 relationship with a performance.
- TAKE SOUND FIELD is an entity. It has a 1:1 relationship with a performance.
- ARRAY is an entity. It describes the geometry, dimension and location of the microphone array.
- TAKE ARRAY is a relationship. A performance has one or more arrays and each array is made of multiple microphones.
- TAKE TRACING is an entity. It has a 1:1 relationship with a performance.
- TAKE 2D is an entity. It has a 1:1 relationship with a performance.
- TAKE 3D is an entity. It has a 1:1 relationship with a performance.
- TAKE VIBRATIONAL is an entity. It has a 1:1 relationship with a violin.

All the related attributes are listed in the Figure B.1.

Implementation

The implementation of the DB has been made using the MySQL database server. A user interface and an administration panel were created with the support of the Yii Framework and all the web pages are implemented in the PHP server-side scripting language.

Presently, the database is under experimental use and will be improved during the evolution of the research project.

Bibliography

[1] C. Gough. The violin: Chladni patterns, plates, shells and sounds. *The European Physical Journal Special Topics*, 145(1):77–101, 2007.

[2] H. Meinel. Regarding the sound quality of violins and a scientific basis for violin construction. *The Journal Of The Acoustical Society Of America*, 29(7):817–822, 1957.

[3] E. V. Jansson. On the prominence of the violin bridge hill in notes of played music. *Journal of the Violin Society of America*, 22(1):169–176, 2009.

[4] P. Bazant, J. Stepanek, and A. Melka. Predicting sound quality of violin from its frequency response. In *Publications- Royal Swedish Academy Of Music*, volume 79, pages 351–354, 1994.

[5] H. L. F. von Helmholtz. *On the Sensations of Tone*. Courier Dover Publications, 1954.

[6] E. V. Jansson. Violin frequency response – bridge mobility and bridge feet distance. *Elsevier Applied Acoustics*, 65(12):1197–1205, 2004.

[7] J. E. McLennan. The soundpost in the violin. part i: The effect of soundpost stiffness on peak resonance and sound output. *AAMIM*, XX(1):24–35, 2001.

[8] C. M. Hutchins. The acoustics of violin plates. *Scientific American*, 245(4):170, 1981.

[9] J. Woodhouse. On the bridge hill of the violin. *Acta Acustica United With Acustica*, 91:155–165, 2005.

[10] Teruaki Ono. Effects of varnishing on acoustical characteristics of wood used for musical instrument soundboards. *J. Acoust. Soc. Jpn.*, 14(6):397–407, 1993.

[11] J. C. Schelleng. Acoustical effects of violin varnish. *The journal of the Acoustical Society of America*, 44(5):1175–1183, 1968.

[12] Zhouyu Fu, Guojun Lu, Kai Ming Ting, and Dengsheng Zhang. A survey of audio-based music classification and annotation. *IEEE Transactions on Multimedia*, 13(2):303–319, 2011.

[13] Paolo Bestagini, Massimiliano Zanoni, Augusto Sarti, and Stefano Tubaro. Feature-based classification for audio bootlegs detection. In *IEEE International Workshop on Information Forensics and Security - WIFS13*, 2013.

[14] Luca Chiarandini, Massimiliano Zanoni, and Augusto Sarti. A system for dynamic playlist generation driven by multimodal control signals and descriptors. In *Proceedings of the IEEE International Workshop on Multimedia Signal Processing, MMSP*, 2011.

[15] F.D. Bloss. *An Introduction to the Methods Of Optical Crystallography*. Holt, Rinehart and Winston, New York, 1961.

[16] Joseph Goldstein, Dale E. Newbury, David C. Joy, Charles E. Lyman, Patrick Echlin, Eric Lifshin, Linda Sawyer, and J.R. Michael. *Scanning Electron Microscopy and X-ray Microanalysis*. Springer, 2007.

[17] J. Woodhouse and P. M. Galluzzo. The bowed string as we know it today. *Acta Acustica United with Acustica*, 90(4):579–589, 2004.

[18] M. E. McIntyre and J. Woodhouse. The acoustics of stringed musical instruments. *Interdisciplinary Science Reviews*, 3(2):157–173, 1978.

[19] M. Mersenne. *Harmonie Universelle*. 1636.

[20] J. W. Strutt Baron Rayleigh. *The Theory of Sound*, volume 2. Courier Dover Publications, 1945.

[21] F. Durup and E.V. Jansson. The quest of the violin bridge-hill. *Acta Acoustica United With Acustica*, 91(2):206–213, 2005.

[22] Ewa Łukasik. Matching violins in terms of timbral features. *Archives of Acoustics*, 31(4):227–233, 2006.

[23] N. H. Fletcher and T.D. Rossing. *The Physics of Musical Instruments*. New York: Springer Verlag, 1998.

[24] E. V. Jansson, N-E. Molin, and H. Sundin. On vibration modes of a violin top plate. a study by hologram interferometry. *STL-QPSR*, 11(2-3):50–54, 1970.

[25] A. Moral and E. V. Jansson. Eigenmodes, inadmittance, and the function of the violin. *STL-QPSR*, 22(1):58–86, 1981.

[26] T. D. Rossing. Chladni's law for vibrating plates. *American Journal of Physics*, 50(3):271–274, 1982.

[27] J. C. Schelleng. The action of the soundpost. *Catgut Acoustical Society*, 19(16):11–15, 1971.

[28] C. M. Hutchins. A note on the function of the soundpost. *Catgut Acoustical Society*, 21:27–28, 1974.

[29] E. V. Jansson, B. Niewczyk, and L. Fryden. Timbre and properties of a violin. In *SMAC 93*, pages 383–386, 1994.

[30] Charalampos Saitis, Bruno L Giordano, Claudia Fritz, and Gary P Scavone. Perceptual evaluation of violins: A quantitative analysis of preference judgments by experienced players. *The Journal of the Acoustical Society of America*, 132(6):4002–12, 2012.

[31] Jan Štěpánek. Musical sound timbre: Verbal description and dimensions. In *Int. Conference on Digital Audio Effects (DAFx-06), Montreal, Canada*, 2006.

[32] Ewa Łukasik. *Digital Media Processing For Multimedia Interactive Services*, chapter Timbre dissimilarity of violins: specific case of musical instruments identification, pages 324–327. World Scientific, 2003.

[33] B. Kostek. *Soft Computing in Acoustic.* Physics Verlag, 1999.

[34] J.B. Kruskal and M. Wish. *Multidimensional Scaling.* Sage Publications, 1978.

[35] Hyoung-Gook Kim, Nicolas Moreau, and Thomas Sikora. *MPEG-7 Audio and Beyond: Audio Content Indexing and Retrieval.* John Wiley & Sons, 2005.

[36] M. Zanoni, D. Ciminieri, A. Sarti, and S. Tubaro. Searching for dominant high-level features for music information retrieval. In *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*, 2012.

[37] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex. *Math. Comput.*, 19:297–301, 1965.

[38] L. Gerosa, G. Valenzise, F. Antonacci, M. Tagliasacchi, and A. Sarti. Scream and gunshot detection in noisy environments. In *European Signal Processing Conference (EUSIPCO)*, pages 1–5, 2007.

[39] L. Lu, J. Zhang, and H. Jiang. Content analysis for audio classification and segmentation. *IEEE Tra*, 10(7):504–516, 2002.

[40] J. Saunders. Real-time discrimination of broadcast speech/music. In *Proc. of Acoustics, Speech, and Signal Processing*, 1996.

[41] Sang-Heon Oh, Youn-Ho Cho, Won-Jung Yoon, and Kang-Kue Lee. Mrtb framework: a robust content-based music retrieval and browsing. In *Proc. of International Conference on Consumer Electronics (ICCE)*, 2005.

[42] Jonghyung Lee and Min-Uk Kim. Music recommendation system using emotion triggering low-level features. *IEEE Transactions on Consumer Electronics*, 59(2):612– 618, 2012.

[43] G. Tzeanetakis and P. Cook. Musical genre classification of audio signal. *Audio, Speech, and Language Processing, IEEE Transactions on*, 10(5):293–302, 2002.

[44] Yi-Hsuan Yang and Home H. Chen. *Music Emotion Recognition.* CRC Press, 2011.

[45] G. Prandi, A. Sarti, and S. Tubaro. Music genre visualization and classification exploiting a small set of high-level semantic features. In *Proc. of the 12th International Conference on Digital Audio Effects (DAFx09), Como, Italy*, 2009.

[46] O. Lartillot and P. Toiviainen. MIR in matlab (ii): A toolbox for musical feature extraction from audio. In *2007 International Society for Music Information Retrieval conference (ISMIR)*, 2007.

[47] D. Cabrera. Psysound: A computer program for psychi-aciustical analysis. In *Proc. of Australian Acoustic Society Conference*, 1999.

[48] E. Pampalk. A matlab toolbox to compute music similarity from audio. In *Prof. of International Conference Music Information Retrieval*, 2004.

[49] C. Landone C. Cannam and M. Sandler. Sonic visualiser: An open source application for viewing, analysing, and annotating music audio files. In *Open Source Software Competition, ACM Multimedia*, 2010.

[50] D. N. Jiang, L. Lu, H. J. Zhang, J. H. Tao, and L. H. Cai. Music type classification by spectral contrast features. In *Proc. of IEEE Internationa Conference Multimedia Expo*, 2002.

[51] Y. Wang, Z. Liu, and J.C. Huang. Multimedia content analysis using both audio and visual cues. *IEEE Signal Processing Magazine*, 17(6):12–36, 200.

[52] S. McAdams. Perspectives on the contribution of timbre to music structure. *Computer Music Journal*, 23(3):85–102, 1999.

[53] J. M. Grey. *An Exploration of Musical Timbre*. PhD thesis, Dept. of Psychology, Stanford University, 1975.

[54] D. Zolzer. *DAFX: Digital Audio Effects*. Wiley, second edition edition edition, 2011.

[55] J.W. Beauchamp. Synthesys by spectral amplitude and brightness mathcing analyzed musical sounds. *Journal of Audio Engineering Society*, 30(6):396–406, 1982.

[56] S.Z. Li. Content-based audio classification and retrieval using the nearest feature line method. *IEEE Transactions on Acoustics Speech and Signal Processing*, 10(7):619–625, 2000.

[57] Koen Tanghe, Sven Degroeve, and Bernard De Baets. An algorithm for detecting and labeling drum events in polyphonic music. In *In Proc. of First Annual Music Information Retrieval Evaluation eXchange*, pages 11–15, 2005.

[58] T. Pohle, E. Pampalk, and G. Widmer. Evaluation of frequently audio feaures for classification of music into precptual categories. In *Workshop on Content-Based Multimedia Indexing*, 2005.

[59] E. Schmidt, D. Turnbull, and Y. Kim. Feature selection for content-based, time-varying musical emotion regression. In *in Proc. of the 11th ACM SIGMM International Conference on Multimedia Information Retrieval (MIR)*, 2010.

[60] S. Davis and P. Mermelstein. Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences. *IEEE Trans. Acoustics, Speech & Signal Processing*, 28(4):357–366, 1980.

[61] D. Liu, L. Lu, and H.J Zhang. Automatic mood detection from acoustic music data. In *Proc. of the International Symposium on Music Information Retrieval (ISMIR 2003)*, Baltimore, Maryland (USA), October 2003.

[62] K. Jensen. Timbre models of musical sounds, tech. rep. rapport 99/7. Technical report, University of Copenhagen, 1999.

[63] R. Plomp and W. J. M. Levelt. Tonal consonance and critical bandwidth. *Journal of the Acoustical Society of America*, 38:548–560, 1965.

[64] W. A. Sethares. *Tuning, Timbre, Spectrum, Scale.* Springer, 1998.

[65] M. Robnik-Ŝikonja and I. Kononeno. An adaptation of relief for attribute estimation in regression. In *Fourteenth International Conference on Machine Learning (ICML'97)*, pages 296–304, 1997.

[66] Massimiliano Zanoni, Luca Trapasso, Fabio Antonacci, and Augusto Sarti. Feature-based analysis of the characteristics of strings in musical instruments. Under review for the conference EUSIPCO 2014, February 2014.

[67] E. V. Jansson. Properties of the violin and the guitar. chapter iv: Properties of the violin and the guitar string. Technical report, Dept of Speech, Music and Hearing, School for Computer Science and Communication, KTH, 2002.