THE VIOLIN ONTOLOGY

Massimiliano Zanoni, Francesco Setragno, Augusto Sarti[1]
[1] *Dipartimento di Elettronica e Informazione - Politecnico di Milano University*
Correspondence should be addressed to:massimiliano.zanoni@polimi.it

Abstract: **Bowed musical instruments have been the subject of scientific investigations for centuries. Yet, the physical phenomena that are behind their timbral quality are still far from being fully understood. This is one of the reasons why the art of violin making is still so strongly tied to tradition. This manuscript describes early results in a study of the relations that exist between timbral and acoustic characteristics of such instruments and their high-level descriptors. In particular, we propose a suitable ontology for a timbral characterization of violins, where every resource is connected and provided with formally defined semantics. Semantic web technologies have taught us how ontologies can become a powerful tool for gathering and managing knowledge in specific areas of interest, where resources are connected and described with formally defined semantics. This, in fact, represents a crucial step for building applications that reason over Web data. In this paper we present an ontology for knowledge representation of violins, as part of a wider ontology of bowed instruments. With this ontology we capture timbral and acoustic aspects of violins as well as violin making and properties of the materials used for their production. We collected and organized semantic descriptors used by numerous violin makers (from natural language) to describe sound proprieties of musical instruments. We also developed an initial model of the relation between semantic descriptors and low-level audio features. The ontology that we present in this study formalizes the semantics of the high-level descriptors and investigates the relation with low-level features. The terminology has been collected through a series of interviews with violin makers in the city of Cremona (Italy), world heritage site for the practice of violin makers. Through listening tests and a feature extraction we study the correlation between high-level descriptors and objective properties of sound.**

1. INTRODUCTION

The art of violin making is the result of five centuries of evolution, which started in Cremona, Italy, with its patriarch Antonio Amati, followed by the renowned families of Stradivari and Guarnieri. Today Cremona is still the heart of this tradition, as confirmed by UNESCO, which in 2012 made Cremona a World Heritage Site for the art of lutherie. Cremona is currently home to over 150 violin makers, and thousands more have studied there, all following techniques that have been passed forward from generations to generations for centuries. As the practice of violin making evolved and consolidated, violins have been a constant subject of scientific investigations. Yet, the physical phenomena that are behind their timbral quality are still far from being fully understood. Because of this inherent difficulty in capturing what physical phenomena make such instruments sound the way they do, violin making has remained an art that that is solidly tied to tradition.

In the past few decades, however, there has been a renewed frenzy in research, aimed at pushing the boundaries of our physical understanding of the quality of violin tone. This recently prompted the city of Cremona to start new research projects with the *Politecnico di Milano* (for aspects of musical acoustics) and the *University of Pavia* (for aspects of material analysis), aimed at exploring new directions in contemporary lutherie. Among the many goals of such projects is that of understanding the links that exist between objective and semantic descriptors related to such instruments. The former are geometric, vibro-acoustic, acoustic and timbral features; physical and chemical properties of materials, etc. The latter are the terms of natural language that are customarily used for describing qualities of the instrument. Identifying such relations would allow us to achieve multiple goals: improving the quality (breadth and span) of information gathering through interviews and questionnaires; improving the effectiveness of training-based algorithms for classification/recognition applications; translating scientific descriptors/features in more intuitive terms for the communitiy of instrument makers and that of musicians; and more.

Several studies from different disciplines are already present in the literature and the presence of several forums for instrument makers on the Web stresses the needs of sharing information in this particular area. However, to the best of our knowledge, an expressive representation of the knowledge and an excellent organization of information, able to provide powerful tools for searching and sharing, are still needed.

Numerous data knowledge paradigms have been developed for a variety of applications, often based on XML [1, 2]. Such paradigms, however, are not expressive enough for defining semantic relationships and constraints over information [3]. The lack of semantics, in fact, makes interoperability difficult, and often hinders the automatic reasoning over data [4]. For this reason, in order to capture, formalize and organize the knowledge we need to take advantage of Semantic Web methodologies [5, 6].

The development of tools for Semantic Web has become particularly popular in past few years. Semantic Web supports the collection and the linkage of structured formalization of knowledge. In this context, ontologies are a very useful tool for expressing knowledge in a specific area of interest. In ontologies, resources are equipped with formally defined semantics. Each resource can be also connected to others in order to express semantic relations in such a way that it becomes possible to build applications that reason over Web data. More recently, Semantic Web has been successfully applied to the area of music information management [7, 8, 9]. One of the main contributions in this area is the OMRAS2 project[1], which aims at building a framework to share information about music (performances, productions and so on) in the Semantic Web format. For this purpose several ontologies have been created. Some examples are the Music Ontology [8], the Audio Effect Ontology [10] and the Audio Feature Ontology[2]. For all these projects, the use of Semantic Web turned out to be a better choice compared to other knowledge management methods [4].

In order to investigate, collect, and conceptualize the traditional and modern knowledge concerning violin instruments, in this study we propose the Violin Ontology[3]. This Ontology is designed to capture the knowledge-related different aspect of these instruments, from the materials to the construction methods, to the description of the sound proprieties. The Violin Ontology is part of the Bowed Instrument Ontology, which is our next goal.

The collection of information related to materials and to the construction methods is a very complex and time-consuming process. In fact, although there exist some general techniques that violin makers usually refer to, they tend to build instruments according to their own personal styles. Moreover, for historical instruments, information is often based on old and often very limited written descriptions. Consequently additional investigations based on advance material analysis (x-ray, micro- and UV imaging, spectrography, and chemical analysis) are often in order, when it comes to tracing the construction history of such instruments. For these and other reasons, in this manuscript we only present a preliminary version of this Section of the ontology.

The characterization of the violin sound should consider timbral and acoustic proprieties as well as the perceptual aspects of the sound, where perceptual aspects refer to how the sound is perceived and

[1] http://www.omras2.org/
[2] http://motools.sourceforge.net/doc/audio_features.html
[3] http://purl.org/net/bowedinsturmentsontology

how it is described and what the relation between sound perception and description is.

As far as sound description is concerned, people use terms from the natural language, such as *Sweet* or *Bright*. These terms are not directly related to mathematical formulas and they are characterized by a high-level of abstraction (*Semantic Descriptors* or *High-Level Features*). Try to formalize the semantics of these terms (conceptualization) is a tough task because of the subjectivity in their use and the lack of a precise and shared meaning. Because of that, in the past decades, several studies have been presented in the literature [11, 12, 13, 14]. The main purpose of these studies is to collect and arrange descriptors in a multi-dimensional perceptual space where distance between concepts can be defined. Similar approaches have been adopted also for the semantic description of the violin timbre [15, 16, 17]). Nevertheless, this approach showed some limitations: the positions of the terms in the space tend to have a large variance because their use is subjective and depends on the context. Moreover, these studies mainly focus on the organization of terms in spaces. They do not provide any formal definition of their semantics in the context of application, which is mandatory in the characterization of the knowledge. In [18] the authors highlight the potential of semantic audio description. However, they present the limitation of the used description paradigms. They stressed the importance of a superior representation able to express the semantic of the descriptors and the complex relation among. We assert that Semantic Web technologies represent a superior choice.

Though our way of describing sounds is based on subjective Semantic Descriptors, there exist a strong connection between sound description, sound perception and physics. Our brain, in fact, processes stimuli from the auditory system in order to formulate a proper description. It is very hard to understand what aspects of the sound influence our perception [19]. However, several studies have been to address such issues. In particular, in Music Information Retrieval, an interdisciplinary research area that studies the extraction of useful information from musical content, the relation between objective timbral and acoustic proprieties (Low-Level Features - LLFs) and subjective semantic descriptors (High-Level Features - HLFs) is studied by means of feature analysis methodologies [19, 20, 21]. Also in the context of bowed instruments, some work has been presented in the literature [22, 23].

The conceptualization of LLFs and HLFs are included in our ontology in order to provide the violin sound characterization. As far as HLFs is concerned, we collected a set of Semantic Descriptors through several interviews to violin makers. In order to investigate the relation between LLFs and HLFs, we recorded musical performance from eleven violins. We then extracted a set of Low-Level Features from the recordings in order to provide a low-level audio characterization for each violin. The recordings were then annotated by five expert violin makers using the collected terms. The correlation between Low-Level Features and semantic annotations were then explored through two statistical correlation indexes: Distance Correlation Index and RreliefF algorithm.

2. A Brief Overview of Semantic Web Technologies

Semantic Web [24] aims at bringing intelligence to the Web by assigning the Web resources a formally defined semantics and forming a very large web of data that can be interpreted by the machine. In order to represent information in a machine interpretable format, a proper language is needed. The Resource Description Framework (RDF) is a data model for describing statements about Web resources using subject, predicate, object triples. When combined, object triples form a RDF graph. If the elements in the RDF graph are identified by a uniform resource identifiers (URI) they build the Web of Linked Data, which is interlinked and globally distributed. In order to represent the knowledge on the entities described in the graph and, hence, to enable reasoning over linked data, a further, more complex representation is needed. The one adopted in this study is the ontology.

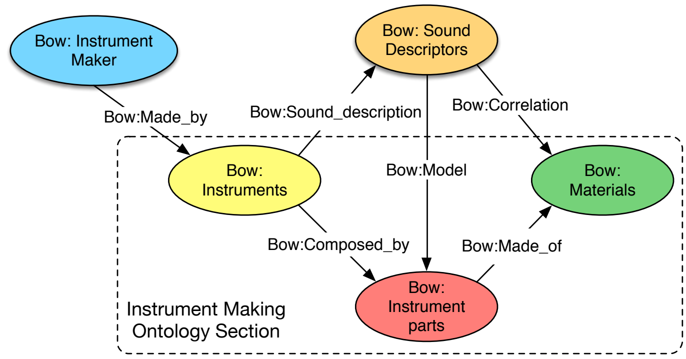

Figure 1: Overall scheme of the Violin Ontology. Different colors represent the different sections of the ontology.

2.1. Ontologies

An ontology is a semantic formalization of a domain of interest. In other words, it is a representation for expressing knowledge in a certain area in a formal (machine-readable) way. In particular, it provides and shares conceptualization of the world of interest by characterizing entities in terms of their meaning and relationships. Ontologies are the core of the world of Semantic Web and their main components are:

- Classes: they represent categories or concepts. A class can have subclasses that have a more specific meaning.
- Individuals: individuals are instances of classes and represent concrete objects. An individual can belong to more than one class.
- Properties: a property is a relation defined between two classes and applied on their individuals.

Classes and properties can be organized in hierarchies with some constraints over their use. Several languages for ontologies exist. In this study, we adopt the well-known Web Ontology Language (OWL) [25].

2.2. OWL

Web Ontology Language is a very popular language endorsed by the World Wide Web Consortium (W3C). Its latest version is OWL2[4].

OWL2 is based on a Description Logic, a family of formalisms used to represent knowledge in a domain. Description Logic is what makes ontologies differ from *taxonomies*. Taxonomies are, in fact, a list of terms and definitions. Whereas, Description Logic allows the construction of intelligent applications (knowledge based systems) able to infer implicit consequences from the explicit information that has been represented. This is done by means of the reasoner.

For better understand the reasoner we provide here an example. Given the class *ItalianViolin* that collects violins built in Italy and an its individual *Stradivari Soil* built in the year 1714; given the class *ViolinMaker* and an its individual *Antonio Stradivari*. Through the reasoner we can in infer that Antonio Stradivari lived Italy at that time. We could not infer such an assumptions without building an ontology.

3. The Violin Ontology

The overall ontology is composed of three sections: *Instrument Maker Ontology Section*, *Instrument Making Ontology Section*, *Sound Description Ontology Section*. Figure 1 shows the overall scheme of the Violin Ontology.

Currently, the Instrument Makers Ontology Section is composed of a unique class *Bow:Instrument Maker* that collects information about the makers of specific instruments, such as historical period and used techniques. The Instrument Making Ontology Section aims to capture the knowledge related to the violin making process

[4]http://www.w3.org/TR/owl2-syntax/

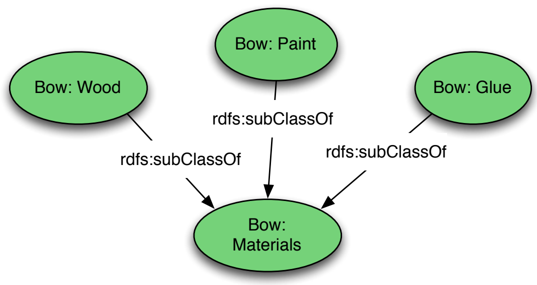

Figure 2: Ontology scheme for the Violin Making Ontology section.

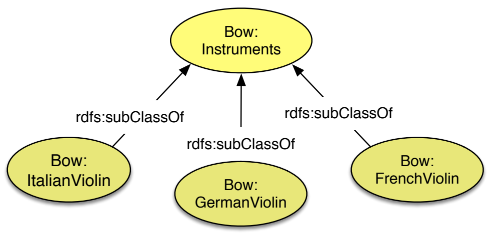

Figure 3: Ontology Scheme for the instrument concept.

and includes information on instrument parts and material. Finally, the Sound Description Ontology Section conceptualize the timbral and acoustic proprieties of the instrument sound.

3.1. Instrument Making Ontology Section

The Instrument Making Ontology Section is composed of three classes (figure 1): *Bow:Instruments*, *Bow:Instrument parts* and *Bow: Materials*. *Bow:Instruments* class provides the characterization of the violins instrument (figure 3). Along the history, different making process procedures has been developed. It is possible to narrow three main schools: Italian, German and French. For this reason, the *Bow:Instruments* concept has three subclasses: *Bow:ItalianViolin*, *Bow:GermanViolin*, *Bow:FrenchViolin*.

The *Bow:Instrument Maker* class is linked through the propriety *Bow:Made_by* to the *Bow:Instruments* class, since a violin maker produces and repair violins. Instruments are composed of numerous components, each one with specific characteristics that has been refined along the centuries. For this reason, the *Bow:Instruments* class is linked to the *Bow:Instrument parts* class through the propriety *Bow:Composed_by*. The *Bow:Instrument parts* class defines a conceptualization of the description of single component that compose the instruments. Each part of the instruments is made of specific materials, whose choice has been a process refined along the centuries. The *Bow:Made_of* propriety describes the relation between the *Bow:Instrument parts* class and the *Bow: Materials*.

Bow:Materials (figure 2) represents the generic concept of material and has three subclasses: *Bow:Wood*, *Bow:Paint*, *Bow:Glue*, which are the predominant materials used in violin liuthery. The class *Bow:Wood*, which individuals are the types of wood, such as maple and spruce, represents the description of the typically used type of wood. Internal and external parts of the instruments are generally painted using a mixture of pigments, copals, white coat and other essences. The class *Bow:Paint* conceptualize the paint mixtures. Various components of the instruments are typically joint using different types of glues. For this reason, we also added the class *Bow:Glue*.

3.2. Sound Description Ontology Section

In the Violin Ontology sound proprieties are described through the *Bow: Sound Descriptors* concept, which is linked to the *Bow:Instruments* class through the *Bow:Sound_description* propriety. As previously mentioned, the quality of sound is strictly related to the intrinsic composition of the materials and to the characteristics of the violin parts. For this reason, we link *Bow: Sound Descriptors* to the class *Bow: Materials* through the *Bow:Correlation* propriety and to *Bow: Instrument parts* through the *Bow:Model* propriety.

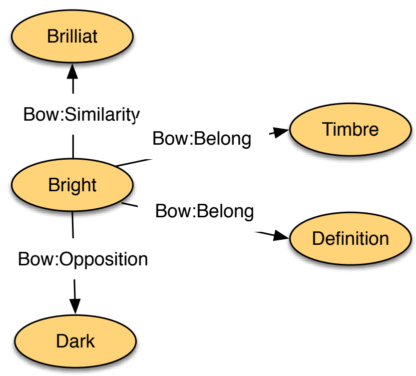

Figure 4: Example of ontology usage for the description of the term Bright

In order to represent the knowledge related to both Low- and High-Level Features, *Bow: Sound Descriptors* class has the *Audio: Low-Level Features* and *Bow: High-Level Features* subclasses. The overall scheme of this section is depicted in figure 5.

As far as HLFs are concerned, there is a common vocabulary that violin makers use to describe the sound proprieties of instruments. Terms in the vocabulary express different shades of sounds. Our intent is to find the most used and relevant terms that are the terms whose meaning (and self-confidence) reached a consensus among makers. We populate the vocabulary of the semantic descriptors through a series of interviews with 20 violin makers in Cremona and Milano, Italy. Starting from a seminal vocabulary populated with terms retrieved by other studies in the literature, in each interview we asked to propose further terms to add. After that, the subjects were asked to outline the meaning of each term, their usage, the relations between them in terms of similarity and their nuance (positive or negative, when possible).

Trying to generalize the human description of sound is a very hard task. This is due to different reasons: different reactions to physical stimuli, different musical backgrounds, different tastes and different way to describe the sounds. For this reason, terms which semantic did not converge to an agreed definition, were discarded. The residuals compose the set of individuals of the *Bow: High-Level Features* concept and the nuances (positive, negative, neutral) are attributes.

The terms might be related to each others in terms of similarity or opposition through the two proprieties: *Bow: Similarity* and *Bow: Opposition*. This two proprieties have been declared symmetric, so that if a term A is linked to the term B, then B will be linked to A with the same property. The vocabulary is listed in Table 2. Terms in the same row and different columns have been annotated as opposite. Whereas terms in same row and same column are annotated as similar.

Figure 4 shows an example of the formalization of the term *Bright* using the Violin Ontology. In the example, *Dark* and *Bright* are opposite, whereas *Bright* and *Brilliant* are similar.

Terms in the vocabulary tend to capture specific aspects of the sound (timbre, propagation, instrument response, overall sensations). This effect is formalized in the Ontology by introducing the class *Bow: SemanticContext* and the propriety *Bow: Belong*. Each term can belong to one or more areas. Terms *Timbre* and *Definition* are two semantic context (figure 4). The semantic contexts included in our Ontology are:

- *Timbre*: this class includes terms that are mainly related to the

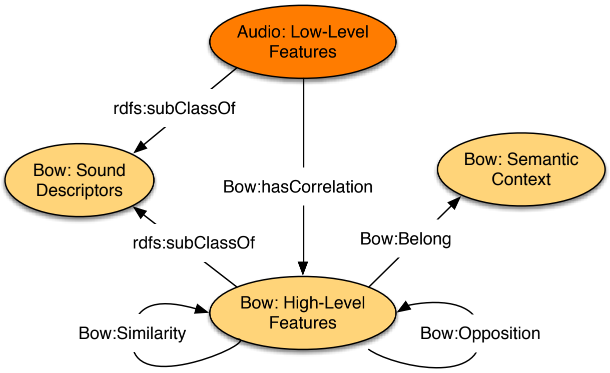

Figure 5: Ontology scheme for the sound description part. *Audio:Low-Level Features* class refers to the Audio Features Ontology developed in the OMRAS project.

1: Root Mean Sqaure energy	21: Zero Crossing Rate
2: Spectral Centroid	22: Spectral Centroid
3: Spectral Skewness	23: Spectral Kurtosis
4: Spectral Roll-off	24: Spectral Flux
5: Spectral Irregularity	25: Harmonic Ratio
6: Attack Slope	26: Attack Time
7-20: MFCC	

Table 1: List of Low-Level Features used for the correlation study

spectral shape of the sound.

- *Definition*: this area groups terms that describe how the sound is defined and focused (e.g. *blurred*, *sharp*).
- *Propagation*: it is related to the irradiation of the sound in the space, including information about volume and directionality.
- *Ease of play*: this area represents the sensations felts in playing the instrument.

As far as the LLFs are concerned, we exploit the Audio Feature Ontology [5] that is one of the ontologies developed within the OMRAS2 project[6]. It is a representation of the knowledge concerning the audio feature characterization of sounds. We include the Audio Feature Ontology as part of the Violin Ontology.

As introduced so far, there is a relation between Low-Level and High-Level descriptors. In order to conceptualize this relation the classes (*Audio:Low-Level Features*) and *Bow:High-Level Features* are linked by the propriety *Bow:hasCorrelation*. This relation will be explored in Section 4.

4. Automatic LLF-HLF Relation Investigation

In order to investigate the relation between LLFs and HLFs we exploit feature analysis and semantic description modeling techniques [19]. The overall procedure is shown in figure 6. It is first needed to collect both low- and high-level audio descriptors for a set of instruments. We collected recordings for eleven modern violins played by a professional musician. The recordings were realized with a high-quality recording system in a semi-anechoic room. In order to capture different nuances of the instrument sound and in order to attenuate the impact of possible noises, interfering signals, and performance-dependent features, the musician was asked to play single notes along the neck for each string, a major scale and two short compositions.

As far as the HLFs are concerned, we collected the annotations by asking to five professional violin makers to provide a description for each violin using the semantic descriptors in the ontology. The testers had a disposal all the recordings for each instrument. Each descriptors were presented along with its opposite. For those that no opposite term exists in the ontology, they were presented with their nominal opposites (e.g. *nasal-not nasal*). The testers were asked to assign a graded annotation ranging from -1 to 1. -1 assigns a strong preference to the first term, whereas 1 assigns a strong preference to the second term.

As far as the LLFs are concerned we extracted a set of audio cues able to capture timbral, harmonic and temporal information. Feature extraction has been performed on the music pieces and on the major scale. The used set of features are listed in Table 1.

Exploiting the low- and high-level sound characterization for each violin, we investigate the correlation between LLFs and HLFs through two well-known correlation methods. The first method is the Distance Correlation Index [26]. Given X_i the vector that collects all the values of the $i-th$ LLF within the corpus and Y_j the vector that collects all the values of the $j-th$ HLF, the Distance Correlation Index is defined as:

$$D = \frac{dCov(X_i, Y_j)}{\sqrt{dVar(X_i)dVar(Y_j)}}, \tag{1}$$

where $dCov(X_i, Y_j)$ is the distance covariance that is defined as

$$dCov(X_i, Y_j) = \sqrt{\frac{1}{L^2}\sum_{l1=1}^{L}\sum_{l2=1}^{L}||X_{l1} - X_{l2}||||Y_{l1} - Y_{l2}||}, \tag{2}$$

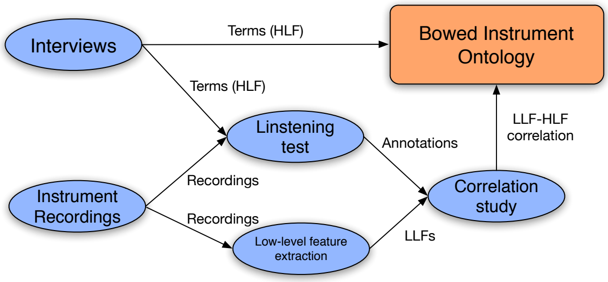

Figure 6: Overall procedure for the Low- and High-Level Features correlation study.

Dark	Bright, Brilliant
Strident, Metallic	Warm
Harsh	Sweet
Unfocused	Focused
Hard	Soft
Hoarse	Clean
Balanced, Even	Unbalanced, Uneven
Weak	Strong
Thin	Full
Deep	
Nasal	
Rounded	
Big, Loud	Quiet
Responsive, Sensitive	Inert
Open	Closed, Covered, Boxy
Projecting	
Edgy	
Sparkly	
Colourful	

Table 2: This table shows the relations among descriptors. Terms in the same row but in different columns are considered as opposite. Terms in the same row and in the same columns are considered as similar.

[5]http://motools.sourceforge.net/doc/audio_ features.html
[6]http://www.omras2.org/

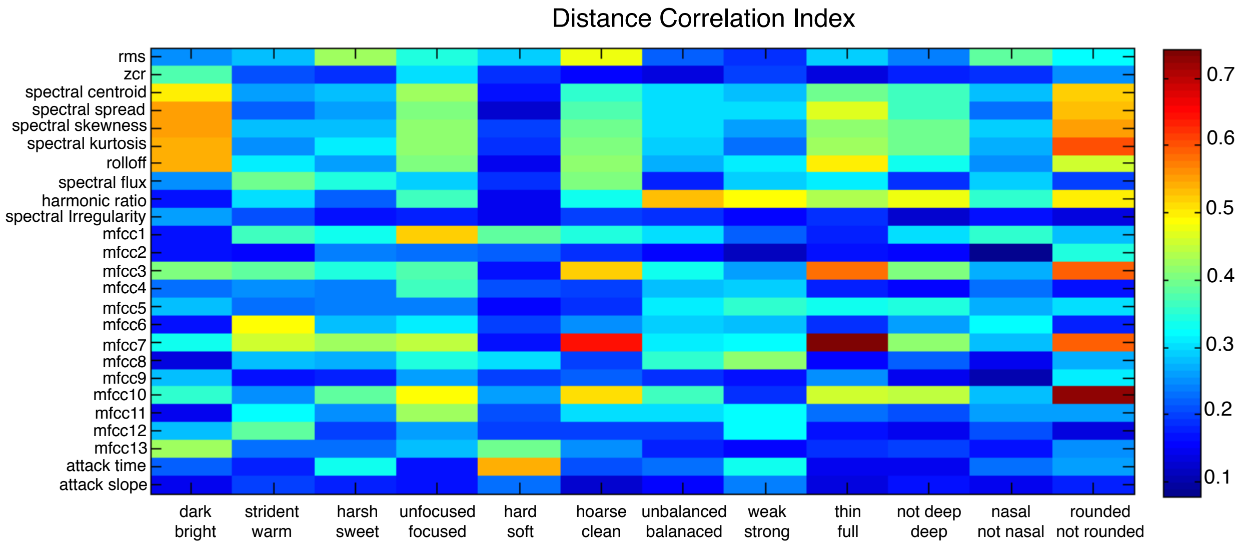

Figure 7: Distance Correlation Index computed on Low-Level Features extracted from 11 violin recordings and semantic descriptors used for the annotations.

L is the number of observations and $dVar$ is the distance variance which is given by

$$dVar(X_i) = dCov(X_i, X_i). \quad (3)$$

The Distance Correlation Index ranges from 0 to 1, where 0 means no correlation (independence).

The second method is a regression-based analysis using the RreliefF algorithm [27]. Given X the observations-by-features matrix and Y_j the collection of the $j-th$ HLF associated to each observation, the algorithm performs a regression and returns a weighted ranking of the Low-Level Features. The higher is the weight for X_i, the higher is the correlation between X_i and the semantic descriptor Y_j.

The limited space of the paper do not allow us to show all the results of the study. For this reason, we present here only some representative examples. Figure 7 shows the Distance Correlation Index for the LLFs and a subset of the semantic descriptors. It can be noticed that each semantic term has a strong correlation with the only a subset of audio cues. For example, *Dark* and *Bright* are intuitively mainly related to the predominance of low and high frequencies. In fact, their definition collected in our ontology are:

- Dark: *A dark sound has a predominance of low frequencies and it is slightly blurred. It is opposed to the concept of bright and sharp. The exaggeration is Unfocused, Covered. Full of harmonics. This feature is related to the wood. Low notes are enhanced.*
- Bright: *This term is often used both in timbre and definition contexts. It indicates a clear, direct and defined sound, with a predominance of high frequencies. It is a quality that characterizes violin from other instruments (viola, cello).*

This effect is confirmed by the Distance Correlation Index analysis. The figure 7, indeed, shows that *Dark/Bright* are mainly correlated with features that capture the spectral shape, such as *Spectral Spread*, *Spectral Skewness* and *Spectral Kurtosis*. Particularly important is also the high correlation with the *mfcc3* and the *mfcc13* features that are respectively related to low and high frequencies.

The results for the *Dark/Bright* terms are confirmed by the RreliefF analysis. In figure 8 it can be noticed that *mfcc13*, *mfcc3*, Spectral Spread, Spectral Skewness and Spectra Kurtosis are at the top of ranking.

Intuitively, semantic descriptors that share a semantic similarity, should also presents some similarity in the distribution of the correlation values with the LLFs. With the Distance Correlation Index analysis it is possible to capture this effect. For example,

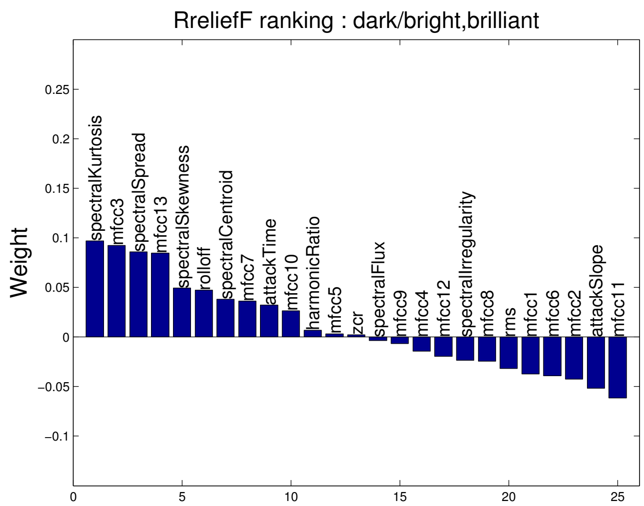

Figure 8: Output of the RreliefF algorithm for the pair Dark/Bright,Brilliant. Feature are ranked based on the value of the correlation. The higher is the histogram, the higher is the correlation.

figure 7 shows how the terms *Thin/Full* and *Rounded/Non Rounded* presents a similar correlation with *Spectral Centroid*, *Spectral Spread*, *Spectral Skewness*, *Spectral Kurtosis*, *Spectral Rolloff*, *mfcc10* and *mfcc13* features.

5. Conclusion and Future Works

In this study, we presented a Semantic Web Ontology covering the domain of the knowledge about violin instruments. The Violin Ontology aims at collecting and sharing knowledge concerning materials, violin parts, as well as Low-Level and High-Level audio features. In particular, the ontology allows to conceptualize the relation between these elements, which are naturally correlated. At this purpose in this study we collected the vocabulary and the semantic definition of descriptors used to describe instrument sound proprieties through a set of surveys. We then exploited the recording of eleven violins to study the relation with Low-Level features.

Feature works provide for extensively examine the dependencies between material proprieties, violin parts proprieties and the sounds proprieties.

References

[1] D. Beenham, P. Schmidt, , and G. Sylvester-Bradley: *XML based dictionaries for MXF/AAF applications*. Technical report, Sony Broadcast and Professional Research Laboratories, UK, 2000.

[2] J. Smith and P. Schirling: *Metadata standards roundup*. In *MultiMedia, IEEE*, volume 13(2):84–88, 2006.

[3] M. Klein, D. Fensel, F. van Harmelen, and I. Horrocks: *The relation between ontologies and XML schemas*. In *LINKÖPING ELECTRONIC ARTICLES IN COMPUTER AND INFORMATION SCIENCE*, page 2001. 2001.

[4] G. Fazekas and M. Sandler: *Novel methods in information management for advanced audio workflows*. In *In proceedings of 12th International Conference on Digital Audio Effects*. Como, Italy, 2009.

[5] G. Fazekas, Y. Raimond, K. Jacobson, and M. Sandler: *An Overview of Semantic Web Activities in the OMRAS2 Project*. In *Journal of New Music Research*, volume 39(4):295–311, 2010.

[6] P. Borst, H. Akkermans, and J. Top: *Engineering Ontologies*. In *International Journal of Human-Computer Studies*, volume 46(2-3):365–406, 1997.

[7] G. Fazekas and M. Sandler: *The Studio Ontology Framework*. In *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR2011)*. 2011.

[8] Y. Raimond, S. Abdallah, M. Sandler, and F. Giasson: *The Music Ontology*. In *Proceedings of the International Conference on Music Information Retrieval (ISMIR 2007)*. 2007.

[9] K. Barkati, A. Bonardi, A. Vincent, and F. Rousseaux: *GAMELAN: A Knowledge Management Approach for Digital Audio Production Workflows*. In *Proceedings of Artificial Intelligence for Knowledge Management Workshop (A14KM) of ECAI*. 2012.

[10] T. Wilmering, G. Fazekas, and M. B. Sandler: *The Audio Effects Ontology*. In *Proceedings of the 14th International Society for Music Information Retrieval Conference*. 2013.

[11] A. Zacharakis, K. Pastiadis, G. Papadelis, and J. D. Reiss: *An Investigation of Musical Timbre: Uncovering Salient Semantic Descriptors and Perceptual Dimensions*. In *Proceedings of the 12th International Society for Music Information Retrieval Conference (ISMIR2011)*. 2011.

[12] J.M.Grey: *Multidimensional perceptual scaling of musical timbres*. In *Journal of the Acoustical Society of America*, 1977.

[13] G. Bismarck: *Timbre of steady tones: A factorial investigation of its verbal attributes*. In *Acustica*, 1974.

[14] A. C. Disley, D. M. Howard, , and A. D. Hunt: *Timbral description of musical instruments*. In *Proceedings of the 9th International Conference of Music Perception and Cognition*. 2006.

[15] J. Štěpánek: *Musical Sound Timbre: Verbal Description and Dimensions*. In *Proceedings of the 9th Int. Conference on Digital Audio Effects (DAFx-06), Montreal, Canada*. 2006.

[16] C. Fritz, A. F. Blackwell, I. Cross, B. C. J. Moore, and J. Woodhouse: *Investigating English Violin Timbre Descriptors*. In *Proceedings of the 10th International Conference on Music Perception & Cognition (ICMPC 10)*. 2008.

[17] C. Saitis, B. L. Giordano, C. Fritz, and G. P. Scavone: *Perceptual evaluation of violins: a quantitative analysis of preference judgments by experienced players*. In *The Journal of the Acoustical Society of America*, volume 132(6):4002–12, 2012.

[18] P. Herrera and X. Serra: *Audio descriptors and descriptor schemes in the context of mpeg-7*. In *Proceedings of the International Computer Music Conference (ICMC 1999)*. Beijing, China, 1999.

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

[20] Y. E. Kim, E. M. Schmidt, R. Migneco, O. G. Morton, P. Richardson, J. Scott, J. A. Speck, and D. Turnbull: *Emotion Recognition: a State of the Art Review*. In *In Proceedings of the 11th International Society for Music Information and Retrieval Conference (ISMIR2010)*. 2010.

[21] D. Turnbull, L. Barrington, D. Torres, and G. Lanckriet: *Towards Musical Query-by-semantic-description Using the CAL500 Data Set*. In *Proceedings of the 30th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '07, pages 439–446. ACM, New York, NY, USA, 2007.

[22] J. Štěpánek and Z. Otčenášek: *Acoustical correlates of the main features of violin timbre perception*. In *Actes du Colloque interdisciplinaire de musicologie (CIM05)*, pages 10–12. Québec, Canada, 2005.

[23] E. Lukasik: *Towards timbre-driven semantic retrieval of violins*. In *Proceedings of the Fifth International Conference on Intelligent Systems Design and Applications (ISDA 2005)*. 2005.

[24] T. Berners-Lee, J. Hendler, and O. Lassila: *The Semantic Web*. In *Scientific American*, 2001.

[25] G. Antoniou and F. van Harmelen: *Web Ontology Language: OWL*. In S. Staab and R. Studer (eds.), *Handbook on Ontologies*, International Handbooks on Information Systems, pages 67–92. Springer Berlin Heidelberg, 2004.

[26] G. Székely, M. Rizzo, and N. Bakirov: *Measuring and testing independence by correlation of distances*. In *Annals of Statistics*, 2007.

[27] M. Robnik-Ŝikonja and I. Kononeno: *An adaptation of relief for attribute estimation in regression*. In *Proceedings of the fourteenth International Conference on Machine Learning (ICML1997)*. 1997.