In the paper the problem of matching instruments according to their timbral similarity is considered for a collection of 53 high quality violins. The fact that the instruments are of the same type makes the subject particularly challenging. The starting point of the analysis is calculation of a set of features based on harmonic analysis, e.g. odd-to-even harmonics amplitudes ratio, intensity of the first and higher harmonics etc., and a set of linguistic descriptors of violin timbre related to these features. The semantically disjoint categories of timbre characteristics are proposed. The result of the analysis is the annotation of the instruments to those semantic categories and creation of a map of similarly sounding instruments. The outcome of the research provides supportive cue for further analysis. It attempts to fill in the existing semantic gap between human and machine levels of sound understanding and to provide the enriching aspect to the design of methodology of inferring preference models from the objective characteristics of musical sounds.

**Key words** musical acoustics, timbre descriptors, semantic retrieval, violin sound.

1. **Introduction**

For many years musical acoustics has been based on physical measurements of sound radiated from musical instruments. Later psychoacoustics have played supportive role in understanding how humans hear. A huge number of listening tests led to construction of reliable models of human auditory systems. They were applied in such domains as sound compression, hearing aids, or sound classification and recognition. The challenge was to discover mechanisms that resemble the way humans make decisions based on what they perceive. Nowadays with the development of computer systems we expect from machines to resemble the way humans make decisions based on how they interpret musical sound. The efforts are focused on adding knowledge to the systems enabling its more ‘intelligent’ processing. Traditionally, information retrieval methods propose the bottom-up approach from acoustical measurements to human-like understanding of musical content where similarity is based on the lowest level analysis (measurements). The reverse, top down process is not mature yet and even if high-level similarity is recognized, while the process goes down it stops at certain point and is able to proceed only with the additional information from the user. Hybrid approaches going from the bottom to the top and in the opposite way are foundation of the critical paradigm of “bridging the semantic gap” [1] between low level description and high level semantic interpretation.

The possible procedure of performing such analysis using fuzzy classification
methods and rough sets has been proposed by Kostek in [5]. She also proposed the inclusion of the idea of “computing with words” to the domain of musical acoustics, e.g. to the process of assessment of musical instruments timbre quality. Zadeh’s computing with words [14] is inspired by the human capability to perform a wide variety of physical and mental tasks without any measurements and any computations. The key role plays manipulation of perceptions (instead of manipulation of measurements) and computing with words provides a foundation for a computational theory of perceptions.

This paper deals with the sound of violin and attempts to make a step towards adding semantics to its analysis. Timbral features may be treated at a low level as some measured parameters responsible for particular sensations of the human listener. Timbre may also get descriptive notations – they would be timbral features at higher, semantic level, e.g. such descriptive concepts as bright, dark, clear, soft, nasal etc. The relationship of both levels in most cases cannot be crisply defined and depend on the listener’s individual preferences. It has been concluded from former research of the author et al. [3] that including some lexical description to the process of inducing expert’s preferences in terms of acoustic features would improve the analysis and results.

The experiments described in the literature on one hand try to find the physical parameters representing given attributes of timbre (some of them have been standardized within the MPEG-7 standard) and on the other try to establish whether any timbral adjectives commonly used in the context of a given instrument have demonstrable universal understanding. From the free descriptions of participants who take part in the experiments usually the most common adjectives are extracted and used as an obligatory vocabulary. Certainly these adjectives may have slightly different meaning for different listeners and in various languages, but such ambiguity has to be accepted, at least at present.

A short presentation of the research on the semantic description of the violin sound has been made in [9]. Also the initial results of matching violins via timbral semantic and physical features have been published there. The instruments tested belonged to a collection of violins contesting during the 10th International Henryk Wieniawski Violinmaking Competition in Poznań (corpus of AMATI database [8]).

This paper aims to extend previous results for larger set of sounds while retaining the methodology. We first take the high level timbre descriptors and their correspondence to the appropriate low-level parameters. This low-high levels match has been proposed some years ago by Harajda and Kwiek [2][6]. It was based on the experiments performed by violinists, acousticians and musicologists. Then we calculate values of these parameters for a set of instruments and try to find the relationship to the preferences of the jury.

The paper is structured as follows. Section 2 discusses problems related to violin timbre, its semantic descriptors and related physical features. Section 3 present results of experiments and their discussion, Section 4 concludes the paper.

2. Violin timbre, its semantic descriptors and related physical features

Timbre is one of the main acoustic features of musical instruments existing along, and often strongly correlated with musical scale, dynamics, time envelope of the sound and sound radiation characteristics [5]. The most used is the ASA definition of timbre (1960): "Timbre is that attribute of auditory sensation in terms of which a listener can judge that two sounds similarly presented and having the same loudness and pitch are dissimilar”. Thus, timbre should be regarded also as a subjective quality that is assessed by listeners.
Such subjective assessment is usually not crispy and often depends on listeners’ individual preferences not possible to be simply correlated with particular measured sound parameter.

Harajda and Kwiek proposed some years ago a taxonomy of violin timbre [2][6] based on the extensive experiments performed by violinists, acousticians and musicologists. We follow an approach as in [12] that decompose observations of music audio into semantically significant dimensions, where each resultant dimension corresponds to the perceived meaning of the audio, and only the most significant meanings (those which are most effective in describing music audio) are kept. The related physical features are commonly used in the literature, e.g. in [4] and their adapted form has been recalled in [9]. Below we give closer characteristics of categories used in this paper.

**Oddness and evenness of harmonics amplitudes**

The descriptors that are related to the intensity of odd or even harmonics amplitudes are given below. We have introduced the parameter odd-to-even amplitudes ratio (O_Ev).

- tense (oboe like) if $O_{Ev} < 1$,
- suppressed (clarinet like) if $O_{Ev} > 1$,
- equalized if $O_{Ev} = 1$.

**Intensity of harmonics**

The descriptors that are related to the intensity of harmonics amplitudes are following:

- deep (strong first harmonic and a big number of other harmonics),
- full (strong first harmonic and a small number of other harmonics),
- flat (weak first harmonic and a big number of other harmonics),
- empty (weak first harmonic and a small number of other harmonics).

The suitable parameters expressing the relationship to the amplitude of particular harmonics is Tristimulus 1 (relative contribution of the first harmonic in harmonic spectrum) and the Tristimulus 3 (relative contribution of fifth and higher harmonics in harmonic spectrum).

**Centre of gravity of harmonic spectrum**

Centre of gravity of the spectrum is responsible for the impression of brightness - one of the best recognized attributes of musical timbre. We have two descriptors expressing it: bright (sharp) and dark. According to Stepanek [11] the boundary between two categories is for the frequency 1200-1400 Hz, however certain pitch-related scaling should be applied due to the pitch dependency in perceiving timbre.

Subjective perception of timbre characteristics changes with the increase of pitch. Kostek [5] proposes exemplary division of classes to be tested to pitch (22 classes) For four violin open strings G3, D4, A4 and E6 we have the span of 15 classes:

- string G3 - pitch range F#3 – G#3 class no. 7
- string D4 - pitch range C4 – D4 class no. 9
- string A4 - pitch range A4 – B4 class no. 12
- string E6 - pitch range D#6 – F6 class no. 18

There are more different categories that have been used for expressing timbre, e.g. those related to tendency of harmonics fall (normal, strained, light), formant placement and character (soprano-alto and sonorous-hollow). Interesting is the concept of Lottermoser [7] who back in fifties analysed the intensity of harmonics and their relationship with timbre from the point of view of musical intervals - consonance or dissonance. This idea has been recalled recently by Wrzeciono in his MSc thesis supervised by the author [13].
3. Discussion of experiment results and their relation to jury ranking

The experiments have been carried out on the collection of master quality violins competing during the 10th International Henryk Wieniawski Violinmakers Competition in Poznań, Poland, 2001. The parameters of 53 instruments, gathered in the AMATI database [8], have been analysed from the point of view of their semantic description related to the physical features. The bottom-up process has been applied: we try to infer instruments semantic features from the appropriate low-level features. Worth noting is, that initial top-down process resulting in matching low-level to high-level features has been already done by the predecessors. The experiments have shown, that on one hand feature values for a set of violins are spanned almost linearly over given range, and on the other, the scattering of values for individual violin strings is not similar form instrument to instrument. These individual values have been visualized in the consecutive figures: 3.1, 3.2, and 3.3. The boundaries inserted between semantic categories are not expected to be very exact, and have only approximate character. Second figure in each pair represents the parameters of the instruments, that have got the highest score for timbre during the competition.

The majority of examined violin sounds are “equalized”, i.e. the odd and even harmonics are equally intensive. The best instruments are also “suppressed”. It was interesting to discover, that the instruments examined - contemporary concert violins from all over the world - are not “deep”, according to [2] – i.e. do not have strong first and strong higher harmonics (high values of Tristimulus1 and .Tristimulus3). Some authors call the sound with many intensive higher harmonics - “squeaky”. Brightness is the feature the most recognized for instruments. The difference of the location of the centre of gravity for each string is different for various instruments. It is roughly expressed by the values of cross-correlation factor for four strings. This factor is different for each individual instrument, but generally, for a whole set of violins, its value is rather small. Estimated coefficients for strings are given below. The medium cross-correlation is observed for strings G and D:

<table>
<thead>
<tr>
<th></th>
<th>D</th>
<th>A</th>
<th>E</th>
</tr>
</thead>
<tbody>
<tr>
<td>G</td>
<td>0,423</td>
<td>0,116</td>
<td>0,083</td>
</tr>
<tr>
<td>D</td>
<td>0,232</td>
<td>0,316</td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>0,092</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

To visualize the feature-spaces for four strings the multidimensional scaling (MDS) has been applied to show it in two dimensions. No visible clusters of similarly sounding instruments have been obtained. Exemplary MDS chart for brightness has been shown in the figure 3.3, on the right. Also we do not observe much similarity between the best instruments. Adding the loudness of the strings to MDS analysis does not change much the instruments’ location in the space. This is in concordance with the results published in [10]. Therefore we tend to conclude, that in the considered timbral auditory dimensions weak similarity would be observed. However it is worth noting, that the whole timbral space is not extensive indeed we examine the timbre of the same type of instrument.

4. Conclusions

In the paper a framework for matching signal processing derived features to the higher-level (linguistic) description of violin timbre has been presented. 53 master quality violins have been analysed from the point of view of low level features responsible for such descriptions as tense, suppressed, full, empty, deep, equalized, bright dark and others.
3.1. Odd-to-Even harmonic amplitudes ratio for individual open strings (left) calculated for 53 violins. The data are ordered according to G-string. Graph on the right: O_Ev ratio for instruments that got the highest scores. There is no clear similarity between the best instruments.

3.2. Intensity of the first versus higher harmonics expressed as Tristimulus1(T1) vs. Tristimulus3 (T3) for individual open strings (left) calculated for 53 violins (4 repetitions of each sound). Graph on the right represents T1 vs. T3 for the instruments ranked as the best.

3.3. Brightness calculated as the centre of gravity of the spectrum of individual open strings (left). Results are ordered according to A-string. The approximate boundary of bright (upper part) and dark (lower part) is inserted. Right: 2-D (reduced using MDS) space of brightness calculated for four strings. Bigger circles - violins with the highest scores.
The experiments have shown, that the low-level features are diverse for each individual instrument and this attitude may be shifted to the impression of timbre distinctiveness in auditory perception. Experiments have been performed on open strings. To give characteristics of “violin voice” a method for analysing the melodies are needed. This is the initial research stage aiming at computer manipulation of perceptions or “bridging the semantic gap” between measurements and semantic description of musical sound. We also see some promise of improving models of musician experts’ preferences in assessing violin voices.

Dealing with violin is on one hand rather narrow, but on the other important for the special interest groups (e.g. violinmakers). The creation of widely approved lexicon is needed, so that each research group would not need to create his own from scratch. The experiments performed on the AMATI database gave characteristics of contemporary concert violins and this is additional extra value of the research and provides supportive cue for further analysis.

References