Spectral sources of basic perceptual dimensions of violin timbre

Jan Štěpánek

Sound Studio of the Faculty of Music, Academy of Performing Arts Prague, Malostranské nám. 13, 118 00 Praha 1, Czech Republic, Email: stepanek@hamu.cz

Introduction

Various kinds of listening tests and statistical procedures suitable for evaluation of test results are used in timbre research. The main goals are usually to recognise the main features or dimensions of human perception and to identify their spectral causality.

Five sets of violin tone recordings (pitches B3, F#4, C5, G5, D6) were used in our studies. Attack and decay transients were unified to weaken their influence on judgements. A set of tones for each pitch were listened to in headphones and judged by a group of experienced listeners (violin players), who assessed the timbre of the violin tones they heard.

Factor analysis was appropriate when a set of variables to describe selected features of violin timbre was used, for example, using the VARR method (Verbal Attribute Ranking and Rating, modified VAME – Verbal Attribute Magnitude Estimation) for a set of verbal descriptions of timbre [1, 2]. Spectral sources of these verbal descriptions were also described [3, 4].

A multidimensional scaling method – latent class approach (CLASCAL) [5] – was applied to dissimilarities in timbre acquired from the pair tests. This method yielded a perceptual space of common dimensions shared by all listeners, and (latent) classes of listeners based on the similarity of individual perceptual models (a posteriori defined groups of listeners); the class models differed in weights of common dimensions.

Two models were used: the weighted Euclidean model (where each stimulus is described by its coordinates in common dimensions and each class of listeners has its own weight for each dimension) and the extended weighted Euclidean model (where each stimulus is additionally described by its specificity value, which indicates the existence of a stimulus feature not shared with other stimuli; a set of all specificities is weighted separately for each class of listeners). The two listener classes for all five pitches fit best [6, 7]. The best models for classes without outstanding listeners were also sought [6, 7] and are summarised in Table 1.

Table 1: The most appropriate models for both classes; Di ... number of dimensions is i, S0 ... without specificities, S1 ... with specificities.

<table>
<thead>
<tr>
<th>Tone</th>
<th>B3</th>
<th>F#4</th>
<th>C5</th>
<th>G5</th>
<th>D6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class1</td>
<td>D1S1</td>
<td>D2S1</td>
<td>D2S1</td>
<td>D2S0</td>
<td>D2S1</td>
</tr>
<tr>
<td>Class2</td>
<td>D1S1</td>
<td>D1S0</td>
<td>D2S0</td>
<td>D2S1</td>
<td>D2S1</td>
</tr>
</tbody>
</table>

Table 2: Significant correlations (level α = 5% in italics, 1% normal, 0.1% bold) of the dimension 1 coordinates with values of spectral centre of gravity (fcg), with levels in harmonics (H1, H2) and in critical bands (B14–B24). Each band containing no harmonic is marked by the symbol ▒.

Table 2:

<table>
<thead>
<tr>
<th>Tone</th>
<th>B3</th>
<th>F#4</th>
<th>C5</th>
<th>G5</th>
<th>D6</th>
</tr>
</thead>
<tbody>
<tr>
<td>Class1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Class2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Methods and Results

Seventeen violin tone recordings were used in the test (listening in headphones) for each pitch. Twenty listeners noted timbre difference (dissimilarity) of steady state parts of sounds.

The perceptual dimension coordinates of these tones in CLASCAL best class models were compared with spectral characteristics calculated from the time-averaged power spectrum of the steady state part of the sound:

a) The overall spectrum level \( L_i \) (in dB) and levels of individual harmonics \( L_{Hi} \) (harmonic spectrum, \( i \) from one up to the value dependent on the fundamental frequency used, levels in dB).

b) Spectral centre of gravity \( f_{cg} \), sometimes referred to as the spectral centroid (defined for example in [8], expressed in Hz).

c) Levels in critical bands (barks) in dB and denoted as \( L_{Bi} \) (bark spectrum, \( i \) from 1 to 24).

Pearson’s correlation coefficient between perceptual dimension coordinates and spectral characteristics was used to establish spectral sources of the main perceptual dimensions. Results for dimension one and two (existing in all the most appropriate class models) and for selected characteristics are summarised in Tables 2 and 3.
Discussion

The first dimension correlated significantly with fCG (positively), with levels in high frequency bands (positively, but in frequency regions that differ for different pitches), and with LH1 (negatively) for pitches B3 (in both listener classes), F#4 (second class, the dimension one of the first class is apparently different) and C5 (both classes). For both classes of the pitch G5 the first dimension correlated with fCG (positively) and with LH1 (negatively), but levels of high frequency bands correlated with the second dimension together with LH2 (the same sign). For both classes of the pitch D6 the first dimension correlated with fCG and with levels of high frequency bands (the same sign) starting from 5 kHz, but LH1 correlated with dimension two. Dimension two correlated significantly with LH2, except in D6.

The spectral energy distribution (represented by fCG) and the level of the first harmonic have a dominant influence on the perception of violin tones: common (but opposite) in dimension one in pitches B3, F#4, C5 and G5, but independent in D6 (where the first harmonic pushes out the second one from dimension two). This corresponds with results indicated in [3, 4] on 'sharpness', which is usually associated with fCG. The behaviour of fCG and LH1 is documented in Figure 1.

This change in behaviour may be connected with the change of the shape of spectral envelope with increasing pitch. This means not only the change of the fundamental frequency but also includes the slope of the envelope and position and number of formants inside (see Figure 2).

Acknowledgement

The research was supported partly by the Ministry of Education and Youth of the Czech Republic (Project No. 511100001) and by the Grant Agency of the Czech Republic (Project No. 202/02/1370).

References