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.

Tone	B3	F#4	C5	G5	D6
$Class_1$	D_2S_1	D_2S_1	D_2S_1	D_2S_0	D_2S_1
$Class_2$	D_3S_1	D_6S_0	D_5S_0	D_2S_1	D_2S_1

Table 1: The most appropriate models for both classes; Di ... number of dimensions is i, S_0 ... without specificities, S1 ... with specificities.

Perceptual spaces of classes exhibit stable dimensions. The second class has higher number of dimensions in pitches B3, F#4 and C5.

This article describes the search for spectral sources of the main perceptual dimensions of the best class models.

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 (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 fcg, 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.

Tone	B3		F#4		C5		G5		D6	
Class	1	2	1	2	1	2	1	2	1	2
fcg	**+.86**	**+.84**		**+.79**	**+.90**	**-.91**	**+.74**	**+.75**	**+.88**	**+.83**
L_{H1}	**-.86**	**-.86**		**-.91**	-.62	**+.78**	**-.94**	**-.94**		
L_{H2}										
L_{B14}				*+.50*	*+.51*	*-.52*	▒	▒	▒	▒
L_{B15}			-.53		*+.58*					
L_{B16}	*+.53*						*+.52*		▒	▒
L_{B17}	**+.75**	+.68								
L_{B18}	*+.51*	*+.58*			*+.51*	*-.52*			▒	▒
L_{B19}	*+.53*		*+.53*		*+.52*					
L_{B20}	+.72	+.72							+.69	+.61
L_{B21}	+.67	+.71				*-.53*			*+.54*	*+.55*
L_{B22}	*+.53*	*+.51*	+.70		*+.50*	-.69			**+.86**	**+.89**
L_{B23}	*+.55*	*+.54*	*+.52*	*+.57*		**-.74**			**+.87**	**+.90**
L_{B24}			*+.53*	+.62	*+.57*	-.72			**+.77**	**+.81**

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 ▒.

Tone	B3		F#4		C5		G5		D6	
Class	1	2	1	2	1	2	1	2	1	2
fcg										
L_{H1}									**+.91**	**-.92**
L_{H2}	*+.58*	*-.51*	+.72	**+.95**	-.63	+.68	-.65	**-.75**		
L_{B14}			*+.49*				▒	▒	▒	▒
L_{B15}										
L_{B16}					**+.79**				▒	▒
L_{B17}							+.65	+.59		
L_{B18}	*-.50*					-.50			▒	▒
L_{B19}									*-.50*	*+.55*
L_{B20}										
L_{B21}										
L_{B22}								*-.58*		
L_{B23}			+.66				-.72	-.70		
L_{B24}			*+.50*				-.71	-.71		

Table 3: Significant correlations (level α = *5% in italics*, 1% normal, **0.1% bold**) of the dimension 2 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 ▒.

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 L_{H1} (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 L_{H1} (negatively), but levels of high frequency bands correlated with the second dimension together with L_{H2} (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 L_{H1} correlated with dimension two. Dimension two correlated significantly with L_{H2}, 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 L_{H1} is documented in Figure 1.

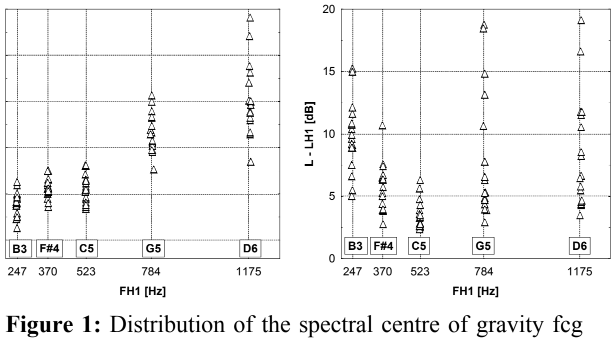

Figure 1: Distribution of the spectral centre of gravity fcg and levels of the first harmonic LH1 (relative to overall spectrum level L) in studied pitches.

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).

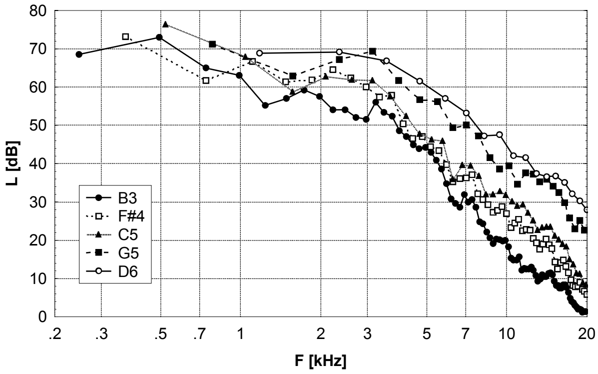

Figure 2: Spectral envelopes of studied violin pitches represented by the average harmonic spectrum.

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

[1] Štěpánek, J. (2002): Evaluation of timbre of violin tones according to selected verbal attributes, 32nd International Acoustical Conference, EAA Symposium "ACOUSTICS BANSKÁ ŠTIAVNICA 2002", 129-132.

[2] Štěpánek, J. (2002): The Study of Violin Timbre Using Spontaneous Verbal Description and Verbal Attribute Rating, Forum Acusticum Sevilla 2002, MUS-06-008.

[3] Štěpánek, J., Otčenášek, Z.: Spectral sources of selected features of violin timbre, 6 CFA, Lille, 695-698, 2002.

[4] Štěpánek, J., Otčenášek, Z. (2004): Interpretation of Violin Spectrum Using Psychoacoustic Experiments, Proceedings of the International Symposium on Musical Acoustics (ISMA2004), Nara, Japan, in print.

[5] Winsberg, S., De Soete, G. (1993): A latent class approach to fitting the weighted Euclidean model, CLASCAL. Psychometrika, **58**, 315-330.

[6] Štěpánek, J. (2003): Latent listener classes and class models in violin timbre, DAGA 2003, Aachen (ISBN 3-9808659-0-8), 532-533.

[7] Štěpánek, J., Otčenášek, Z. (2003): Listener common and group perceptual dimensions in violin timbre, Proceedings of the Stockholm Music Acoustics Conference (SMAC 03), Stockholm, 663-666.

[8] Štěpánek, J., Otčenášek, Z., Melka, A., Syrový, V. (1997): Violin Tones Spectra and their Relationship to Perceived Sound Quality, In *Proceedings of the Institute of Acoustics ISMA´97,* **19** (5), Edinburgh, 125-130.