Latent listener classes and class models in violin timbre

Jan Štěpánek
Sound Studio of the Faculty of Music, Academy of Performing Arts Prague,
Malostranské nám. 13, 11800 Praha 1, Czech Republic; Email: stepanek@hamu.cz

Introduction

Usually the aim of a listening test is a detection of distinctive features of specific sound context and the search for their acoustic correlates. A question arises whether the listeners are in higher or lower agreement in their judgements or even groups of listeneres with different evaluation models exist.

In this contribution the results of listening tests of violin tones are studied according to listeners and their perceptual models. Five sets of violin tone recordings (pitches B3, F#4, C5, G5, D6) were used in the study [1]. Attack and decay transients were unified to weaken their influence on judgements. Seventeen tones for each pitch were listened in headphones and judged. Twenty experienced listeners – violin players (Academy professors and students) assessed dissimilarity in timbre in pairs of violin tones.

The results of five listening tests (individual dissimilarity matrixes) were separately processed using latent class approach applied on weighted Euclidean model (CLASCAL) [2] and extended CLASCAL model (owerview of models see in [3]). This yields to the construction of perceptual spaces of common dimensions shared with all listeners and classification of listeners in classes (groups); the groups differ in dimension weights.

Method

In classical multidimensional scaling method (MDS) dissimilarities among stimuli are modelled to fit distances in Euclidean space of low dimensionality. Latent class approach (CLASCAL MDS) [2,4] solves two optimisation tasks:
1. Fitting of stimuli dissimilarities in distances of Euclidean space of appropriate low dimension.
2. Establishing of appropriate number of (latent) classes of listeners and adding each individual listener to one of the classes.
Optimal model selection is an iterative process in which alternately the first and the second task solution is improved.

New version of CLASCAL program [4] enables an application of latent class approach with both weighted Euclidean model (each stimulus is described by coordinates in common dimensions; each class of listeners has its own weight of every dimension) and extended weighted Euclidean model (each stimulus is described moreover by specificity value; specificity indicates the existence of stimulus feature not shared with other stimuli; set of all specificities is weighted separately for each class of listeners).

CLASCAL use will be illustrated on the dissimilarity data of pitch F#4. Initial number of classes was established using Hopes Monte Carlo significance test [2]; two classes (C2) were chosen. The models with two to five dimensions (D2 – D5), without specificities (S0) and with specificities (S1) were estimated. For the choice of appropriate dimensionality a minimum value of information criterion BIC [3] was used in both cases (see Table 1). For the selection which model (C2D3S0: three-dimensional model without specificities and C2D3S1: three-dimensional model with specificities) is better Hopes test was used; model C2D3S0 revealed as better.

Number of dimensions D	Models S0 BIC	Models S1 BIC
2	6010	6000
3	**5704**	**5595**
4	5754	5645
5	5795	5762

Table 1: Information criterion BIC values for two-class models of different dimensions (D2 – D5) without (S0) and with (S1) specificities; minimal values are bold. Solutions for pitch F#4.

For the verification of number of classes (Cn), comparison of a selected model with models having the same dimensionality but n+1 and n–1 classes is proposed in [4]. Two criteria are available in CLASCAL program: Hopes test and latent class (LC) bootstrap (description see in [4]). LC bootstrap defines bootstrap samples (generated by sampling of individual dissimilarity matrixes with replacement), calculates latent class analyses one for each bootstrap sample, and establishes dissimilarity for each pair of listeners (based on the agreement – disagreement of their classification in individual bootstrap sample solutions). The listeners dissimilarity matrix can be analysed with clustering algorithm. Clustering trees for Cn-1, Cn and Cn+1 classes enable to decide the number of classes.

When the number of classes differs from the initial choice, new number of classes is accepted and the iterative process continues with the selection for the appropriate dimensionality until no change in number of classes and dimensions is suggested.

Hopes criterion verified model C2D3S0 for pitch F#4 as optimal. In LC bootstrap one hundred bootstrap samples were used for the calculation of listeners dissimilarity matrix. Cluster analyses showed two classes (Figure 1), so C2D3S0 was also verified.

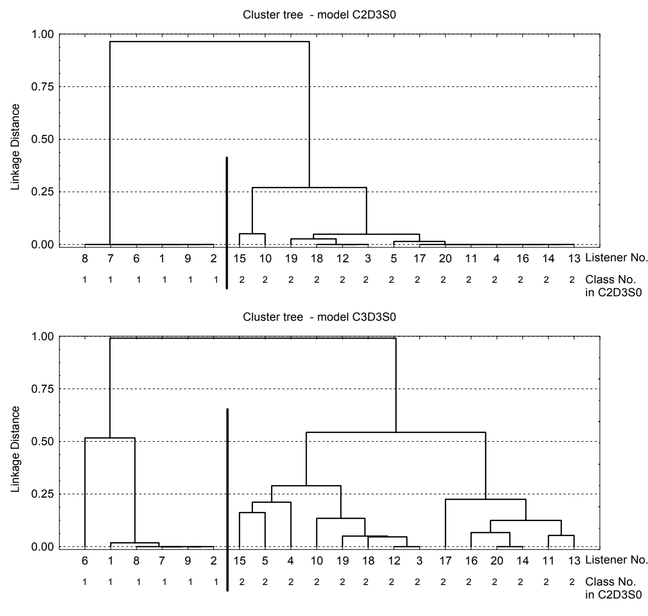

Figure 1: Cluster trees for models C2D3S0 and C3D3S0 for pitch F#4. Unweighted pair-group average amalgamation method was used.

CLASCAL MDS procedure in common yields to a model CiDjSk, which best fits dissimilarity data. Model parameters are as follows:
- number of listener classes; each listener belongs to a certain class,
- number of common dimensions,
- stimulus coordinates according to dimensions, defining its position in common perceptual space,
- stimulus specificity value (eventually), one for each stimulus,
- weight, one for each common dimension (eventually for a set of all specificities) and each class.

Results

Latent class approach was applied on the dissimilarity data of all five pitches. Table 2 illustrates models obtained as best fits.

Pitch	B3	F#4	C5	G5	D6
Model	C2D3S0	C2D3S0	C2D2S1	C2D2S1	C2D2S1
No. of listeners in class 1 / 2	10 / 10	6 / 14	6 / 14	9 / 11	12 / 8

Table 2: Selected latent class models (CiDjSk, i ... number of classes, j ... number of dimensions, k ... specificities: 0=No, 1=Yes) and number of listeners in individual classes.

More detailed behaviour of individual listeners according to their classification is summarised in Table 3.

No.	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
B3	■	■	■	■	■	■	■	■	■	■	□	□	□	□	□	□	□	□	□	□
F#4	■	■	□	□	□	■	■	■	■	□	□	□	□	□	□	□	□	□	□	□
C5	■	■	■	■	■	□	□	□	■	□	□	□	□	□	□	□	□	□	□	□
G5	■	■	■	■	■	■	■	■	□	■	□	□	□	□	□	□	□	□	□	□
D6	■	■	■	■	■	■	■	■	■	■	■	■	□	□	□	□	□	□	□	□
c1:	5	5	4	4	4	4	4	4	4	3	1	1	0	0	0	0	0	0	0	0
c2:	0	0	1	1	1	1	1	1	1	2	4	4	5	5	5	5	5	5	5	5

Table 3: Individual listeners belonging to classes in the studied pitches; ■ - class 1, □ - class 2.

All listener classes share common dimensions (and specificities), differences in class models consist in weights of dimensions and specificities. Summary of weights is illustrated in Figure 2.

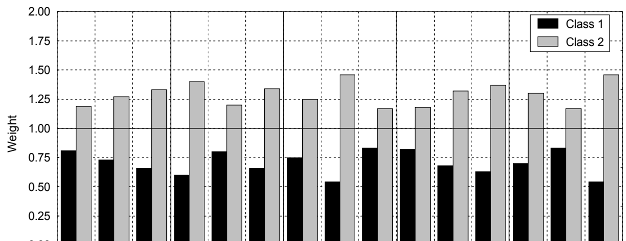

Figure 2: Weights of dimensions and specificities in individual classes.

Discussion

The selected most appropriate models (Table 2) created logical sequence according to the increasing pitch. In the lower pitch tones revealed three common dimensions without stimuli specificities. In the higher pitches only two common dimensions appeared, specificities induce existence of small number of stimuli with individual features, which were not joined into one common dimension.

Division of the listeners into classes in different pitches is very stable (Table 3). Class one has lower weights in all dimensions/ specificities (Figure 2). These facts suggests an idea of two different timbre judgement strategies which are independent on pitch. To prove this hypothesis, mean dissimilarity values judged in each listener were calculated and averaged in each class (see Figure 3).

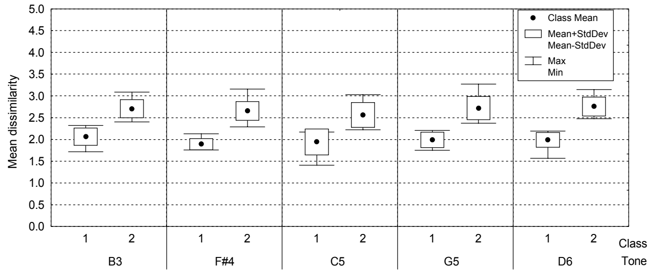

Figure 3: Means of dissimilarities in classes of listeners together with standard deviations and minimal and maximal values of individual mean dissimilarities. Mean dissimilarity scale agrees with the values used during the listening tests (0, 0.5, 1, ..., 5).

Statistical t-test for independent samples verified significant mean differences between classes in all five pitches, differencies in standard deviations were not significant. So one of the differences between classes is the measure of the exploitation of dissimilarity scale by individual listener but this can not fully explain different ratios of model dimension weights (CLASCAL MDS method differentiates listeners into classes just on the base of ratios of the dimension weights).

Conclusion

Further investigation of test results is necessary, namely the interpretation of common dimensions according to spectral features (preliminary calculation revealed significant correlation of the first dimension coordinates with spectral centre of gravity in all five pitches). This interpretation can also help for the search of further resasons of different ratios of dimension weights. An attempt to compare spectral correlates in different pitches could lead to better understanding of judgement strategies in perception of violin timbre.

Acknowledgement

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

[1]| 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.

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

[3]| McAdams, S., Winsberg, S., Donnadieu, S., De Soete, G., Krimphoff, J. (1995): Perceptual scaling of synthesized musical timbres: common dimensions, specificities, and latent subject classes. *Psychological Research*, 58, 177-192.

[4]| Winsberg, S., De Soete, G. (2002): A bootstrap procedure for mixture models: applied to multidimensional scaling latent class models. Applied Stochastic Models for Business and Industry, 18, 391-406.