# DIRECTIONAL TIMBRE SPACES OF VIOLIN SOUNDS

**Zdeněk Otčenášek, Jan Štěpánek**

Sound Studio of the Faculty of Music, Academy of Performing Arts Prague,
Malostranské nám. 13, 11800 Praha 1, Czech Republic,
e-mail: otcenasek@hamu.cz

**Abstract**

The goal of this study is to describe timbre dependence of sound in geometrical space surrounding a musical instrument (violin). Sounds of different violin tones played *detache*, *naturale*, and *non-vibrato* were recorded using the 98-microphone measurement method described in [1]. Listening to these recordings revealed distinguishable changes in loudness and timbre depending on direction (microphone position). Various listening tests (sorting, dissimilarity in pairs of representatives from sorting) were used for lowering the time necessary for the recognition of the span and description of changes in timbre. Results for the G3 violin tone revealed that timbres appearing around the instrument have a high-dimensional nature and are unique in different smaller or larger areas.

## INTRODUCTION

Timbre of musical instrument sound is determined by the amount of radiated energy in individual spectral components, which are captured by the auditory organ and perceived by the listener.

Timbre and loudness of a steady-state tone played on a musical instrument varies according to listening place around the instrument. The cause for this variability is differences in energy distribution radiated from the source in various directions and frequencies.

As it is not possible to pre-determine width and gradient changes in sound for a specific instrument, it is necessary when measuring to divide the whole sphere surrounding the instrument into a large number of small segments with one microphone in each, and to subsequently search for very similar sounds by listening to recordings from the spherical areas. Considerable similarity of recorded sounds requires application of sensitive methods of listening tests used for subjective assessment of these recordings (e.g. dissimilarity pair test [3]). However, the enormous number of recordings eliminates their direct use. For example, 98 recordings generated 4753 pairs, which cannot be processed within a reasonable time. In order to reach the goal – to describe timbre dependence of sound in geometrical space surrounding a musical instrument (violin) – it is necessary to adapt existing or develop new methods to handle huge amount of recordings in listening tests with an acceptable number of assessments.

## METHODS AND RESULTS

The 98-microphone method was used to map changes in sound surrounding the instrument played by a musician. In this method, the sounds were recorded by microphones positioned on a closed cylinder with a spatial angle offset approximately $2\pi/16$. A François Gand (1825) violin was played in an anechoic room; the tones were played the violinist in *mezzoforte*, *detache*, *naturale*, *non-vibrato*. The player's variability in loudness was detected and corrected. The sounds in all prepared listening tests were listened in closed Sennheiser HD 250 linear II headphones, and the SPL was equal to SPL during recording, corrected to unify distance from the instrument.

### Data reduction listening test

The recordings from all 98 microphones were used in the first listening test. The goal of this test was to reduce the number of sounds without significant loss of timbre and loudness variability. The test was prepared in the MATLAB ® (The Mathworks, Inc.) environment on PC (see screen copy in Figure 1).

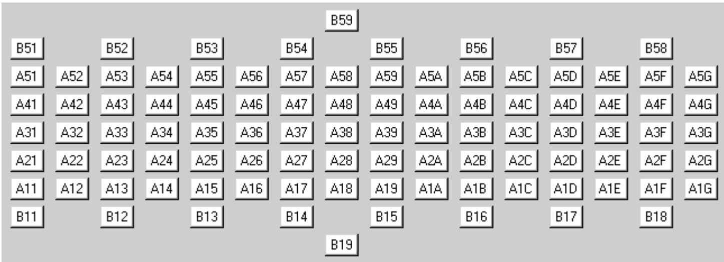


**Figure 1.** PC screen in sorting 98 signals. Each rectangle represents the signal of one microphone. The rectangles are organized in rows according to planes of microphone positions used in the measurement (bottom-up: B1, A1–A5, B5), one row displays an expanded circle of 16 or 9 microphones from a single plane.

Each sound was played by clicking on the selected button. The respondent's task was to determine the extent of dissimilarity in pairs of neighbouring sounds, which required listening to 196 sound pairs. This extent was marked by a dividing line between the buttons as follows:

| | |
|---|---|
| no difference was listened | without a line |
| very slight difference | ······· |
| small difference | —— |
| big difference | ▬▬ |

The results of this listening test are found in Figure 2. We can see a continuous area in which sounds change minimally, where their boundaries are represented by unbroken lines. At the same time, 19 label of microphone positions whose sound recordings represent a demarcated area are bold highlighted.

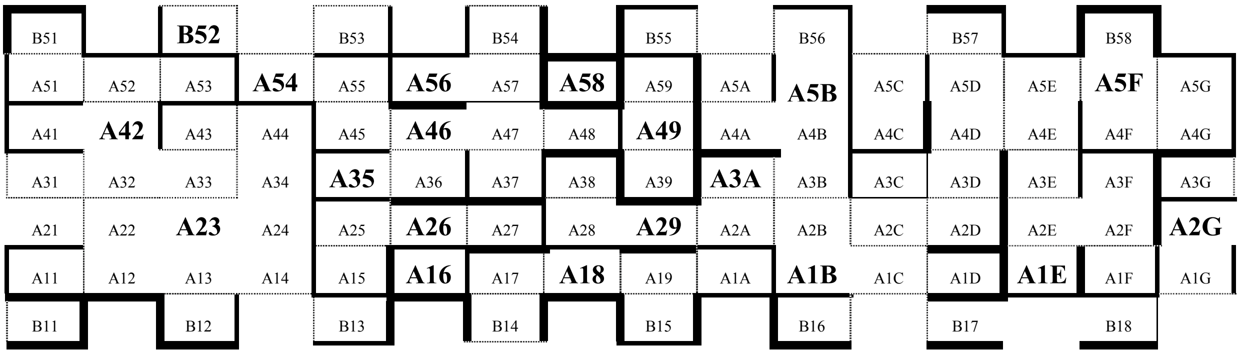


**Figure 2.** Results of sorting 98 signals. The dissimilarity measure of neigbouring signals found is illustrated:

These 19 "representatives" were used in the subsequent listening tests.

**Pair dissimilarity test of representative sounds**

The second listening test was a classical pair dissimilarity listening test on a set of representatives from the first test. Consequently, the measure of dissimilarity between sounds from non-adjacent microphones was determined. A scale from 0 (no difference) to 4 (largest dissimilarity) was used.

| | A42 | A23 | B53 | A54 | A35 | A56 | A46 | A26 | A16 | A58 | A18 | A29 | A49 | A3A | A5B | A1B | A1E | A5F | A2G |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a42 | 0.00 | | | | | | | | | | | | | | | | | | |
| a23 | 2.25 | 0.00 | | | | | | | | | | | | | | | | | |
| b53 | 1.25 | 1.25 | 0.00 | | | | | | | | | | | | | | | | |
| a54 | 3.00 | 3.00 | 3.25 | 0.00 | | | | | | | | | | | | | | | |
| a35 | 3.25 | 3.25 | 3.50 | 2.50 | 0.00 | | | | | | | | | | | | | | |
| a56 | 3.00 | 3.00 | 2.00 | 2.50 | 3.00 | 0.00 | | | | | | | | | | | | | |
| a46 | 2.70 | 2.75 | 1.75 | 1.50 | 2.25 | 1.75 | 0.00 | | | | | | | | | | | | |
| a26 | 2.70 | 1.75 | 3.50 | 3.00 | 1.50 | 3.25 | 1.50 | 0.00 | | | | | | | | | | | |
| a16 | 1.00 | 1.50 | 3.00 | 3.50 | 2.50 | 3.50 | 1.50 | 1.00 | 0.00 | | | | | | | | | | |
| a58 | 2.50 | 2.50 | 3.00 | 2.50 | 3.50 | 3.25 | 2.75 | 2.00 | 3.00 | 0.00 | | | | | | | | | |
| a18 | 3.50 | 2.00 | 3.00 | 3.50 | 3.75 | 3.50 | 2.00 | 3.00 | 1.00 | 3.00 | 0.00 | | | | | | | | |
| a29 | 2.00 | 1.00 | 3.00 | 3.50 | 3.00 | 2.50 | 3.50 | 3.00 | 3.00 | 3.00 | 1.00 | 0.00 | | | | | | | |
| a49 | 2.00 | 3.50 | 2.50 | 2.50 | 3.50 | 2.00 | 2.50 | 3.00 | 3.00 | 1.00 | 3.00 | 1.00 | 0.00 | | | | | | |
| a3a | 3.00 | 2.75 | 3.75 | 3.25 | 3.50 | 3.75 | 3.75 | 3.50 | 3.50 | 3.50 | 3.50 | 3.00 | 2.75 | 0.00 | | | | | |
| a5b | 2.75 | 2.75 | 3.25 | 2.75 | 3.75 | 3.00 | 3.00 | 3.50 | 3.00 | 2.00 | 2.00 | 2.50 | 1.25 | 2.50 | 0.00 | | | | |
| a1b | 1.50 | 1.00 | 2.50 | 2.75 | 3.50 | 3.50 | 2.75 | 2.75 | 2.00 | 2.00 | 2.25 | 1.25 | 1.00 | 1.50 | 1.25 | 0.00 | | | |
| a1e | 3.00 | 3.00 | 3.50 | 3.50 | 4.00 | 4.00 | 4.00 | 4.00 | 1.00 | 3.50 | 1.50 | 3.00 | 2.50 | 3.00 | 3.00 | 2.00 | 0.00 | | |
| a5f | 2.50 | 1.50 | 1.00 | 3.00 | 3.50 | 3.00 | 1.00 | 3.00 | 2.00 | 2.50 | 2.50 | 2.50 | 2.00 | 2.50 | 2.00 | 2.00 | 2.50 | 0.00 | |
| a2g | 2.50 | 2.00 | 2.50 | 3.00 | 3.50 | 3.50 | 2.50 | 3.00 | 2.00 | 2.50 | 2.50 | 2.50 | 2.00 | 2.50 | 2.00 | 1.50 | 2.50 | 1.00 | 0.00 |

**Table 1.** Resulting dissimilarity matrix from the listening pair test of representative sounds.

Results show a large amount of representative pairs had a large measure of dissimilarity (Table 1), and at least 12 sounds differed highly among themselves (dissimilarity equal or greater than 2.5). Thus, at least six dimensions are necessary for a successful fit of objects into constructed perceptual space reproducing their distances (dissimilarities). The Multidimensional Scaling Method (MDS) is usually used for the construction of perceptual space. The maximum number of acceptable dimensions D using S objects (stimuli) must fulfil the condition $D < (S-1)/4$ (see [4]), which gives maximum of four dimensions in the case of 19 stimuli. As a result, the MDS method could not be used and the space was not constructed.

**Spontaneous word description of representative sound properties**

The third listening test had two parts. In the first part the respondent listened to all representative sounds consecutively and spontaneously described their timbre in words. The respondent could listen to described sounds in contrast to any other representative sounds for a complete word description of each perceived sound.

The second part was the quantification of all properties (different words used in the first part of the test) on all the representative sounds.

| | DARK | BRIGHT | SOFT | BLURRED | MUTED | DELICATE | ROUGH | BUZZING | GLOSSY | GLOOMY | NASAL | GRAIN | MASSIVE | FULL | WIDE | RICH | NARROW | CLEAR |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| a42 | 2.00 | 0.00 | 0.00 | 1.60 | 0.00 | .50 | 0.00 | 1.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | .10 | 0.00 | .70 | .10 |
| a23 | 2.00 | 0.00 | .70 | .40 | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 | .50 | 0.00 | 0.00 | 1.00 | 1.00 | 0.00 | .80 | .20 | .10 |
| b53 | 1.00 | 0.00 | 1.00 | 1.60 | 0.00 | 1.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a54 | 0.00 | 2.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 2.00 | 1.00 | 0.00 | 2.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 0.00 |
| a35 | 0.00 | 0.00 | 0.00 | .10 | 0.00 | 0.00 | 0.00 | 3.00 | 2.00 | .50 | 0.00 | 0.00 | .20 | .50 | 2.00 | .80 | .10 | 1.00 |
| a56 | 0.00 | 1.00 | .50 | .80 | 0.00 | 1.00 | 0.00 | 2.00 | 0.00 | 0.00 | .30 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.20 | 0.00 |
| a46 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | .50 | .10 | .80 | .60 | .50 |
| a26 | 2.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 1.00 | .25 | 0.00 | 0.00 | 1.00 | .25 | 1.00 | 1.00 | 0.00 | 2.00 |
| a16 | 2.00 | 0.00 | 1.00 | 1.80 | 1.00 | 0.00 | 0.00 | .25 | .25 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.50 | 0.00 |
| a58 | 0.00 | 2.00 | .70 | 1.60 | 0.00 | 0.00 | 0.00 | 2.00 | .75 | 0.00 | .25 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a18 | 2.00 | 0.00 | 1.00 | 1.40 | .10 | 0.00 | .50 | 0.00 | 0.00 | 2.00 | 0.00 | .20 | 2.00 | 0.00 | 0.00 | .80 | .80 | .10 |
| a29 | 2.00 | 0.00 | .50 | 1.60 | 0.00 | 0.00 | 1.00 | .50 | 0.00 | .50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a49 | 1.00 | 0.00 | 1.00 | .20 | 0.00 | .50 | 0.00 | 1.00 | 0.00 | .50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a3a | 2.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | .50 | .25 | 0.00 | 0.00 | 1.00 | 1.00 | 2.00 | .50 | 1.50 | 0.00 | 2.00 |
| a5b | 2.00 | 0.00 | 1.00 | .80 | 0.00 | .50 | 0.00 | .50 | 0.00 | .25 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 2.00 | 0.00 |
| a1b | 2.00 | 0.00 | 1.00 | 1.80 | .50 | .50 | 0.00 | 1.00 | 0.00 | 1.00 | .50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.50 | 0.00 |
| a1e | 2.00 | 0.00 | .50 | 2.00 | 3.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.50 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a5f | 1.00 | 0.00 | 1.00 | 1.60 | .10 | .50 | 0.00 | 1.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |
| a2g | 0.00 | 0.00 | 1.00 | 1.50 | .50 | .20 | 0.00 | 1.00 | .25 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 | 1.00 | 0.00 |

**Table 2.** In the table, the values for individual representatives show the extent of a property for a sound. These properties are expressed here with a word description.

The Factor Analysis Method (FA) was used to determine the minimal number of factors in which representative sound differences were described, and to find words with a common appearance in the context of representatives. The results of FA are found in Table 3.

Factor Loadings (Varimax normalized) Extraction: Principal components (Marked loadings are > .700000)

| | Factor 1 | Factor 2 | Factor 3 | Factor 4 | Factor 5 | Factor 6 | Factor 7 | Factor 8 | Factor 9 |
|---|---|---|---|---|---|---|---|---|---|
| DARK | -.15 | .25 | .13 | .05 | -.16 | **.86** | -.11 | -.22 | .14 |
| BRIGHT | .10 | .04 | .18 | -.47 | .04 | **-.78** | .10 | .02 | .28 |
| SOFT | .40 | .52 | .15 | .09 | .04 | .07 | .03 | -.10 | **-.71** |
| BLURRED | .61 | .49 | .47 | .06 | .14 | .03 | -.25 | .06 | .08 |
| MUTTED | .14 | .14 | .14 | .06 | .12 | .13 | **-.94** | .02 | .03 |
| DELICATE | .11 | .22 | **-.88** | .10 | .21 | -.02 | .17 | .19 | .08 |
| ROUGH | -.01 | .17 | .16 | .12 | **-.92** | .15 | .11 | -.19 | .01 |
| BUZZING | .06 | **-.72** | -.08 | -.06 | .04 | -.56 | .28 | .14 | .08 |
| GLOSSY | .01 | **-.86** | .26 | -.04 | .13 | -.21 | .10 | .03 | .28 |
| GLOOMY | .18 | .09 | .18 | .03 | -.12 | .30 | -.54 | -.67 | -.18 |
| NASAL | .09 | -.03 | .04 | **-.95** | .05 | -.21 | .08 | .06 | .08 |
| GRAIN | **-.90** | .18 | .22 | .04 | .19 | .07 | .05 | -.08 | .05 |
| MASSIVE | -.44 | -.01 | .14 | .14 | -.20 | .13 | .14 | **-.82** | .01 |
| FULL | **-.95** | -.07 | .00 | .12 | -.16 | .05 | .03 | .05 | .04 |
| WIDE | -.26 | **-.92** | .04 | .18 | .08 | .02 | .00 | -.06 | -.03 |
| RICH | **-.79** | -.31 | -.13 | .20 | -.11 | .11 | .09 | -.39 | .11 |
| NARROW | .52 | .38 | .05 | -.68 | .18 | -.03 | -.06 | .16 | -.18 |
| CLEAR | **-.72** | -.49 | .02 | .16 | .19 | .17 | .12 | -.12 | .20 |
| Expl.Var | 4.04 | 3.30 | 1.32 | 1.76 | 1.17 | 1.94 | 1.44 | 1.49 | .83 |
| Prp.Totl | .22 | .18 | .07 | .10 | .07 | .11 | .08 | .08 | .05 |

**Table 3.** The results of Factor Analysis

The most suitable number of factors is five or nine (see Scree plot in Figure 3).

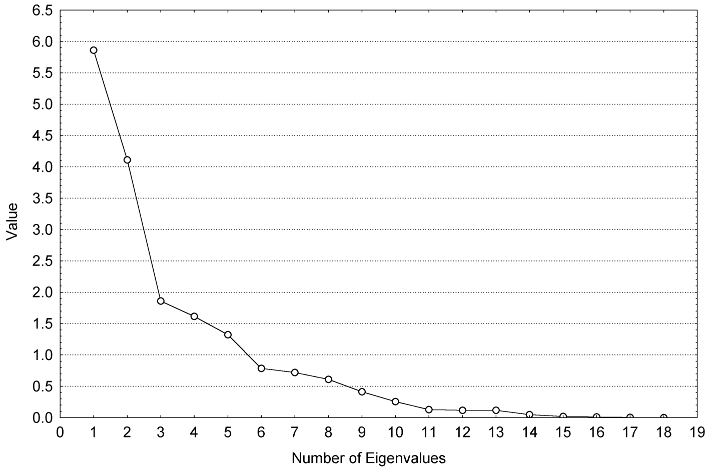


**Figure 3**. Scree plot of eigenvalues

## DISCUSSION AND CONCLUSIONS

In Figure 2. we can see one larger area of similar timbre in the G3 violin tone in the directions of typical violin listening (A42, A23) and a second behind the instrument (A5F). Remaining directions represent more frequent changes in timbre and loudness. From the second test it follows that the timbre varies markedly in different directions around the instrument without being repeated.

The third test results revealed that representative timbres can be described as a combination of at least nine factors. Sound properties are thus described in independent (dimensional, orthogonal) words. Although the words "glossy" and "buzzing" originate from common factor, it is only a contextual union (in this case, both properties appear for all but two sounds). In this test a respondent explained "glossy" as feeling of *sss* (high frequency components) and "buzzing" as feeling of *zzz* (lower frequencies); from this we can conclude "glossy" and "buzzing" are independent sound properties.

These demanding tests have to date only been carried out with two respondents; presented results have been processed for only one.

## ACKNOWLEDGMENT

The investigation was supported by the Ministry of Education and Youth (Project No. 511100001).

## REFERENCES

[1] Štěpánek, J., Otčenášek, Z., "Sound Directivity Spectral Spaces of Violins", Proc. ISMA'2001, Perugia, 2001.
[2] Dünnwald, H., *Die Qualitätsbestimmung von Violinen mit Hilfe eines objektiven Verfahrens*, Dissertation, TU Aachen, 1984.
[3] Guilford, J. P., *Psychometric methods*. McGraw Hill, New York, 1954.
[4] Coxon, A. P. M., *The User's Guide to Multidimensional Scaling*, Heinemann Educational Books, 1982.