

Towards the physical correlates of musical timbre(s)

Saleh Siddiq¹, Christoph Reuter¹, Isabella Czedik-Eysenberg^{1,2}, Denis Knauf²
¹Institute of Musicology, University of Vienna, Austria, ²Vienna University of Technology, Austria



Introduction

Assumptions

Reducing musical instruments to only one note, as implicitly suggested by the standard definition [1], is a misconception, because every instrument can produce a broad range of timbres. Studies with one single note per instrument (e.g. [2]) effectively study single sounds and yield results that perfectly apply to what they investigate: isolated sounds. Unfortunately the results are projected onto the entire instrument. Studies with several tones per instrument suggest that perceptual dimensions shift with respect to the newly included variable (e.g. [3]).

Research questions

- Which physical properties are best suited to explain the distribution of instrument sounds in a perceptual space, taking pitch and dynamics into account?
- Do any of these sound properties qualify as discriminator(s) of musical instruments?

Methods

Stimuli selection

Goal

Inclusion of a broad variety of excitation mechanisms.

- Bassoon (BS)
- Clarinet (CL)
- Flute (FL)
- Trombone (TM)
- Violoncello (VC)

Restrictions

Stimuli number limited because of listening task duration.

Selection

Five nonpercussive instruments with different excitation mechanisms.

instrument	excitation	octaves of pitch range				
		2	3	4	5	6
Flute (FL)	vortex at edge			E	E	E
Clarinet (CL)	reed		E		E	E
Cello (VC)	bowed string	E		E		E
Bassoon (BS)	double reed	E	E		E	
Trombone (TM)	player's lips	E	E	E		

Each note was tested in three different dynamic levels: *pianissimo* (*pp*), *mezzoforte* (*mf*), *fortissimo* (*ff*).

All sounds were taken from the Vienna Symphonic Library (VSL)

Listening task

pairwise comparison

$n \cdot (n-1)/2$ pairs ($n = 45 \rightarrow 990$ pairs)
 \rightarrow AB-pairs rated, AA and BA omitted
 7-point rating scale (7 = max dissimilarity)

Participants

$n = 43$ ($f = 20, m = 22, \text{other} = 1$)
 age: 13–75 yrs. ($\bar{X} = 32.86, SD = 14.93$)
 musical activity: 32 active, 8 formerly, 2 never
 experience: 2–60 yrs. ($\bar{X} = 20.66, SD = 14.08$)

MDS- Perceptual data

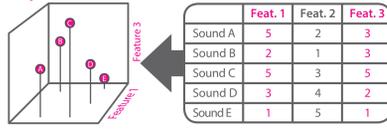
Multidimensional Scaling (MDS): spatial representation of perceived dissimilarities ordinal MDS (NMDS) via smacof in R

MIR- Physical data

Music Information Retrieval (MIR): computational audio feature extraction MIRtoolbox in Matlab

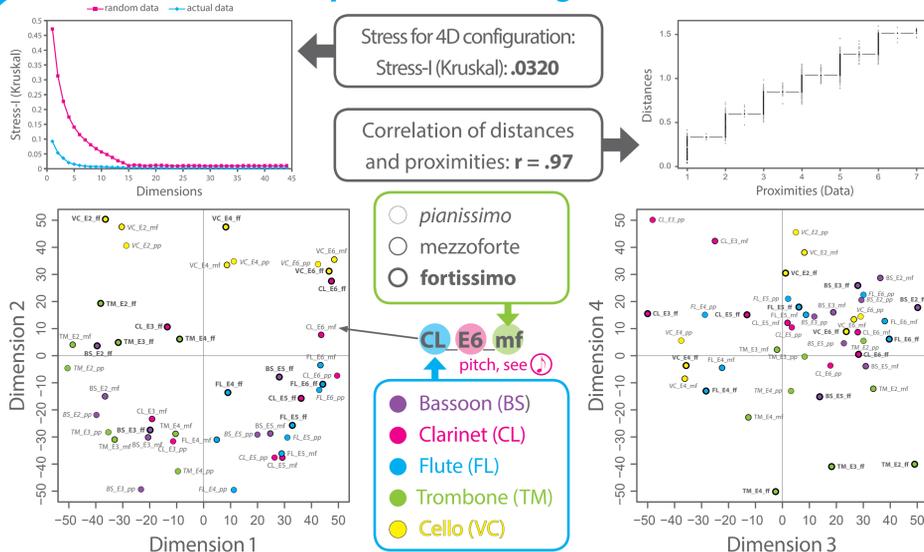
Correlation of perceptual and physical data

Correlations between feature values and coordinate vectors may explain the main axes of the spatial configuration calculated via MDS.



Results

The spatial MDS configuration



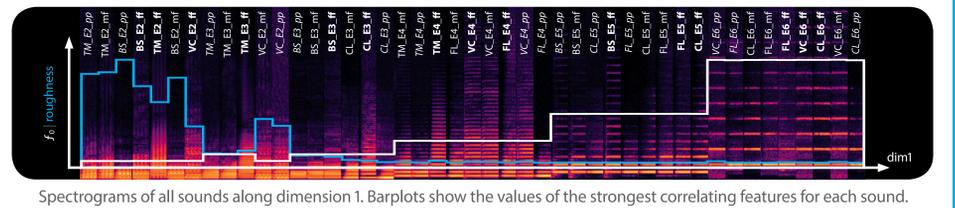
Three basic takeaways

- Except for the cello along dimension 2, there are no instrument clusters.
- Instruments strew widely across the perceptual space.
- Similarities (spatial proximity) do not indicate that tones stem from the same instrument.

The physical correlates of the perceptual dimensions

Dimension 1 correlates strongly with features that are very closely related to pitch.

f_0 : $r = .97$ ($p < .001$)
 roughness: $r = -.86$ ($p < .001$)



Dimension 2 correlates strongly with features that are very closely related to the spectral energy distribution.

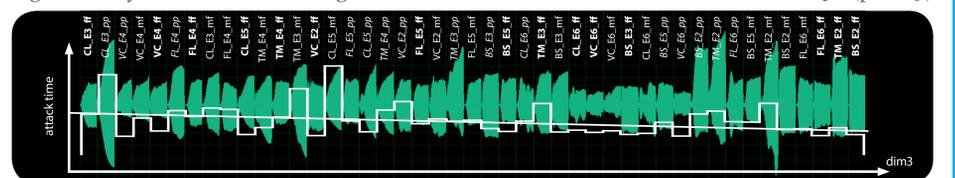
Brightness: $r = .86$ ($p < .001$)
 Spectral rolloff: $r = .72$ ($p < .001$)
 Spectral centroid: $r = .62$ ($p < .001$)



Above: Spectrograms of all sounds along dimension 2. Barplots show the values of brightness and spectral centroid for each sound.
 Left: Within one note (e.g. FL E4), sounds are ordered by dynamic level in almost every case. The cello is the odd exception.
Dimension 2 correlation with dynamic levels: $r = .56$ ($p < .001$)
 Average correlations of note and dynamics: **cello** $r = .38$ | **rest** $r = .88$

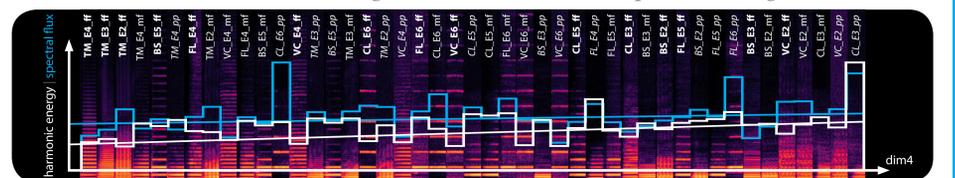
Dimension 3 correlates weakly but significantly with the decreasing attack time.

attack time: $r = -.32$ ($p < .05$)

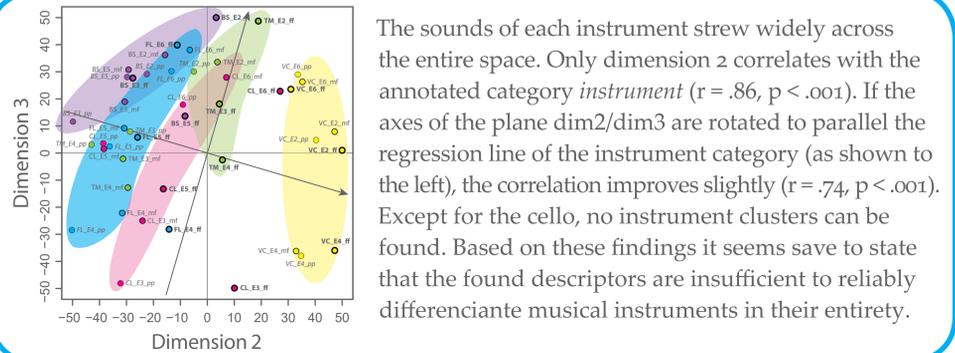


Please note: Attack time is only measured until the first peak. Thus, irregularities may force the algorithm to return lower values than there should be. With that in mind it can be seen from the waveforms that correlations should be even higher.

Dimension 4 shows no significant correlations. Examining the spectrograms and listening to the sounds indicates an increasing dominance of noise components along dimension 4.



Discrimination of musical instruments?



Takeaways Conclusions

- Unsurprisingly, the timbre of an instrument varies with pitch and dynamics changes.
- Consequently, timbral spectra of musical instruments are quite heterogeneous.
- Properties like pitch may even override similarities based on excitation mechanisms
- Musical instruments obviously cannot be appropriately tested on the basis of single sounds.
- The entity „musical instrument“ might not be a psychoacoustical but rather a psychological phenomenon after all.

References

- [1] American National Standards Institute (1960). *American Standard. Acoustical Terminology (Including Mechanical Shock and Vibration)*. New York, NY: ASA
- [2] Grey, J. M. (1975). *An exploration of musical timbre using computer based techniques for analysis, synthesis and perceptual scaling*. PhD dissertation, University of Stanford.
- [3] Handel, S. & Erickson, M. L. (2004). Sound Source Identification: The Possible Role of Timbre Transformations, *Music Perception* 21(4), 587–610.

Acknowledgements

Supported by funds of the Österreichische Nationalbank (OeNB, Anniversary Fund, project number: 16473). Special thanks to the Vienna Symphonic Library for supporting our work with two free „Vienna Super Package“ licenses.