Timbre Discrimination in Drum Tom-Toms

An approach based on audio descriptors and timbre space

Carlos Eduardo Leal de Castro[1,2] • **Marcelo Gomes de Queiroz**[2]

[1] Ibmec São Paulo, SP, Brazil | [2] Institute of Mathematics, Statistics and Computer Science, University of São Paulo (IME-USP), Brazil

carlos.castro@professores.ibmec.edu.br

1 Introduction & Motivation

The sonic identity of a drum arises from a complex network of physical and mechanical variables — shell material and thickness, reinforcement ring, drumheads, hoops, tuning and excitation. In tom-toms, even small construction changes can produce perceptually relevant timbral differences, which matters for both acoustics research and instrument design.

Research question. Can controlled *micro-variations* in real tom-tom construction be discriminated through audio descriptors and represented in compact *micro-timbre spaces*?

2 Controlled Dataset

Custom **12" × 10" double-headed tom-toms** fabricated with **Odery Drums** — same 45° bearing edge and consistent manufacturing, varying one factor at a time:

- **Wood species** — Araucaria, Copaiba, Teak
- **Shell thickness** — 5.5 mm / 8.5 mm
- **Reinforcement ring** — with / without
- **Batter drumhead** (Evans) — clear / coated
- **Hoop** — double-flanged steel / die-cast aluminum

A single drum size is kept throughout to avoid confounding construction effects with size-related pitch differences.

3 Recording & Excitation Control

- Dedicated room, close-miking with an omnidirectional Behringer ECM8000 (≈60°, 5 cm).
- Tension set with a Tama Tension Watch; four levels (60/70/80/85 t.m.u.) across batter & resonant heads → 16 tuning conditions.
- Standardized pendulum-based striking apparatus (Fig. 1): silicone-tipped 2B stick for shells; timpani mallet for assembled toms.
- A repeated-strike test confirmed excitation consistency.

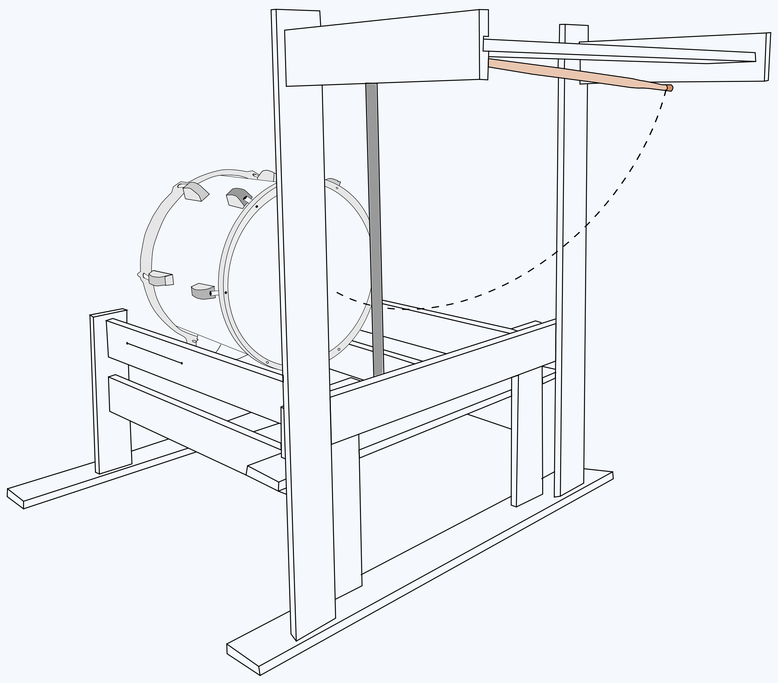


Figure 1. Pendulum striking apparatus. Left: shell vibration setup. Right: assembled tom-tom setup.

4 Descriptors & Feature Selection

Temporal, spectral and spectro-temporal descriptors — spectral centroid, flatness, entropy, slope, zero-crossing rate (ZCR), inharmonicity, log-attack-time, effective duration — computed with Librosa and Timbre Toolbox definitions (2048-sample Hann window, 50% overlap).

Four complementary methods were integrated, keeping descriptors chosen by multiple methods as *robust* indicators:

SVM **Random Forest**
ANOVA **Mutual Information**

5 Consensus Descriptors

Descriptors most consistently selected across the four methods for each construction / assembly factor (Table 1).

Factor	Main consensus descriptors
Wood type	Spectral Slope; Spectral Entropy; ZCR
Shell thickness	Effective Duration; Inharmonicity; Spectral Flatness
Reinforcement ring	Spectral Kurtosis; Effective Duration; RMS
Hoop	Spectral Entropy; ZCR; Effective Duration
Drumhead	Spectral Kurtosis; ZCR; Effective Duration

Table 1. *Effective Duration* recurs across most factors — sustain and damping are central to tom-tom timbral identity.

6 Micro-Timbre Spaces

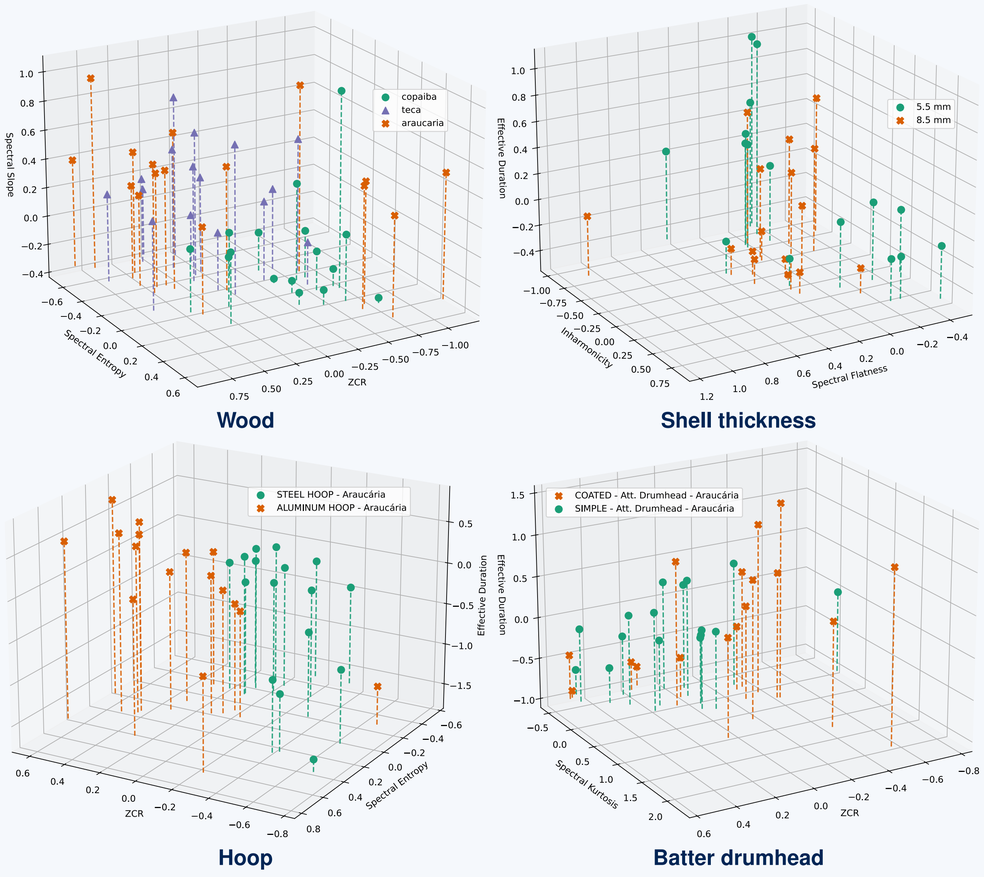


Figure 2. Per-factor 3-D micro-timbre spaces, each built from the descriptors most consistently selected for that construction / assembly factor.

Each factor is plotted in the space of *its own* most-selected descriptors. Single-factor construction changes are not scattered at random — they organize into a readable structure that maps back to audible properties of the sound:

- **Wood:** Copaiba separates clearly, while Teak & Araucaria overlap. Less internally damped woods radiate *brighter*, longer-decaying shells, which shifts Spectral Slope, Spectral Entropy and ZCR.
- **Shell thickness:** carried mainly by *Effective Duration* — selected by all four methods — a direct read-out of how fast spectral energy dissipates (the damping factor); Inharmonicity further pulls apart the higher-tuned heads.
- **Reinforcement ring:** Spectral Kurtosis + Effective Duration give an *almost total* separation regardless of tuning — ringed shells are stiffer and more damped, producing shorter, lower-energy, temporally compact responses.
- **Hoop:** dominated by Spectral Entropy and ZCR — swapping steel for die-cast aluminium changes the drumhead's vibration modes, and with them the spectral energy distribution and harmonic content.
- **Batter drumhead:** the clear head is *brighter* and less sustained than the coated head, separated along Spectral Kurtosis, ZCR and Effective Duration.

Across every factor, Effective Duration keeps re-appearing — sustain and damping are central to a tom-tom's timbral identity.

7 Conclusion

Subtle physical modifications in real tom-toms produce structured, perceptually organized variation in descriptor space. The proposed descriptor-based micro-timbre spaces are compact representations that remain meaningfully connected to audible aspects of drum sound — spectral complexity, brightness, sustain, damping and harmonic control. The contribution is therefore both *methodological* and *timbral*: controlled construction differences can be translated into interpretable, compact timbral representations.

References & Acknowledgements

Ibáñez-Arnal, Doménech-Ballester & Sánchez-López (2020). *Applied Acoustics*, 169, 107449.
La Cerra (2020). *The modern drum set, part 3: Toms*. Yamaha Hub.
McAdams, Winsberg, Donnadieu, De Soete & Krimphoff (1995). *Psychological Research*, 58, 177–192.
Peeters, Giordano, Susini, Misdariis & McAdams (2011). The Timbre Toolbox. *JASA*, 130(5), 2902–2916.
Siedenburg & McAdams (2017). *Frontiers in Psychology*, 8, 1747.

Tom-tom shells fabricated in partnership with **Odery Drums**. Research developed at **COMPMUS - IME, USP**. The author acknowledges financial support from **CAPES**.