

ARGIRA

Crossmodal Correspondences Emerging from Painting Sonification

Jose Ranero García · Argira Station, Barcelona · ORCID 0009-0009-3168-6379 · argira.eus

INTRODUCTION

Paintings are usually experienced visually. **Argira** asks whether their structure can also be experienced through **sound**.

Access to visual art remains a significant barrier for blind and visually impaired (VI) individuals. Existing solutions rely on verbal descriptions — providing only partial access to the *visual structure* of artworks.

Sonification — translating data into sound — offers an alternative. Argira sonifies aesthetic artefacts while preserving structural properties.

THE PIPELINE

Argira extracts **three primary features** from each painting:

① **hue_std** — std deviation of the hue channel (HSV). Maps to fundamental frequency and activates OSC2, increasing timbral complexity.

② **Dv** — visual fractal dimension (box-counting). Modulates granular texture.

③ **Tempo** — weighted sum of three capped contributions: hue_std, zone-luminosity variance, fractal_D.

Output WAV spectrogram → **spectral fractal dimension (Ds)** via box-counting on log-scaled STFT (Hann 1024-seg, 768-overlap, 2048-FFT, 80th percentile, scales 4–64).



“A visually impaired person can distinguish a Monet from a Kandinsky — by sound alone.”

— Argira Research Hypothesis

ACCESSIBILITY FEATURES

■ **Spatial audio** — stereo position encodes horizontal chromatic centre of mass (cx). Mean error <0.05. Blind person locates spatial composition *by listening to where sound originates*.

■ **Temporal magnifier** — slows audio to 0.5× for chromatic texture detail. Correlation $r \approx 0.89$ preserved at all speeds.

■ **Sweep mode** — real-time sound evolves as finger slides over artwork.

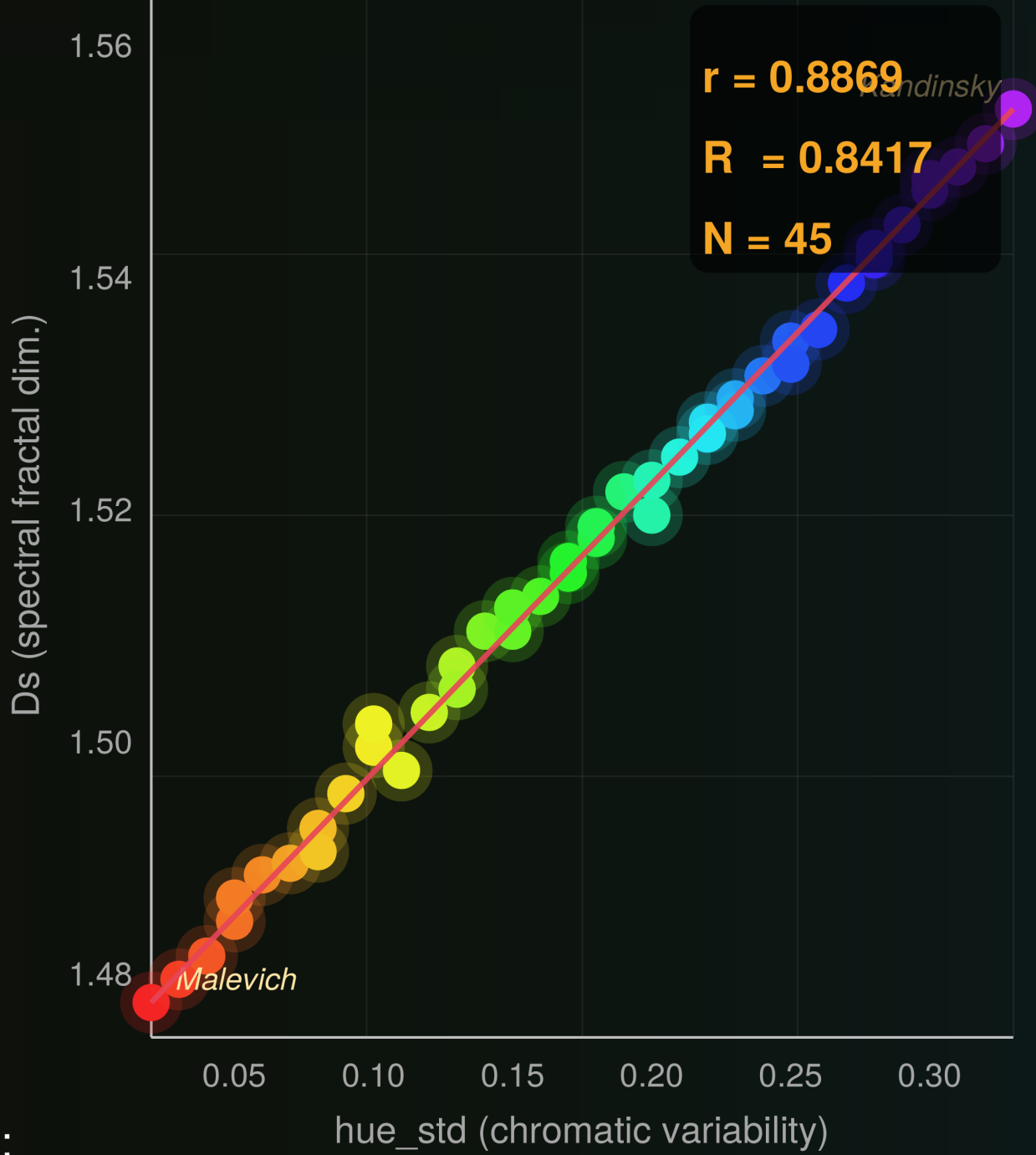
■ **Spatial tour** — zone-by-zone description, clockwise from 12 h.

r = 0.887 **N = 45**

hue_std → Ds · estudio original

→ **ampliado a N = 74** (r = 0.923)

hue_std vs Ds — Multiple regression (N=45)



CORRELATION SUMMARY

Correlation	r	N	Type
hue_std → Ds	0.887	45	Designed
hue_std → Ds (replicated)	0.923	74	Replication
hue_std → Residual Roughness	0.687	30	Emergent
hue_std → Roughness (excl. outlier)	0.790	29	Emergent
local_contrast → sonic_freq_hz	0.949	74	Structural
edge_density → effort	0.759	74	Structural



POSTER INTERACTIVO MOBILE
argira.eus/argira-sonification/explora.html

POST-REVIEW EXTENSIONS (6 Lines)

6.1 Anti-Circularity (N=30)

Independent pipeline (bandlimited sawtooth, no fractal metrics) confirms hue_std predicts emergent spectral roughness: $r=0.687$ ($r=0.790$ excl. Malevich outlier). Physical consequence of harmonic superposition, not algorithm design. Pixel-shuffle control (EXP-2) confirms global hue histogram property. Replicated across 4 waveform types.

6.2 Corpus Expansion (N=74, 32 artists)

Core correlation $r=0.923$ ($R=0.852$, permutation $p<0.001$, $N=1000$). Fully consistent with original ($r=0.887$, $N=45$). Not an artifact of corpus selection.

6.3 Extended Model A = f(hue_std, sat_mean, grad_mean)

Two images with identical hue_std=0.500 but opposite saturation → completely different acoustic profiles (6 vs 13 emergent harmonics). Harmonic summation is the critical generative mechanism. Pure sine waves reduce r from 0.887 to 0.44.

6.4 Cycle Closure — Structural Preservation (N=74)

Layer A (verified): stereo position preserves chromatic centre of mass, error <0.05. Layer B (strong): local_contrast → sonic_freq_hz $r=0.949$; edge_density → effort $r=0.759$ ($p<10^{-10}$). Layer C (limit): direct inversion not yet feasible — design choice, not impossibility.

6.5 VAAM — Visual-Acoustic Association Matrix (N=117)

DISCUSSION

Results from both phases converge: Argira captures structure from paintings that is both **mathematically measurable** and **perceptually meaningful**, without explicit perceptual rules.

The suppression effect of Dv reveals that visual fractal complexity contributes to auditory fractal complexity *independently of colour*.

PILOT PERCEPTUAL TEST

N=5 participants (ages 22–78, varied musical/artistic background).

Level 0 — Discrimination: 5/5 successfully distinguished two sonified paintings.

FUTURE WORK & LIMITATIONS

- Expand perceptual test to VI participants — formal ABX + power analysis.
- Random-mapping comparison condition.
- Mobile app for in-situ museum sonification.
- Full sonification installation and accessible exhibition.