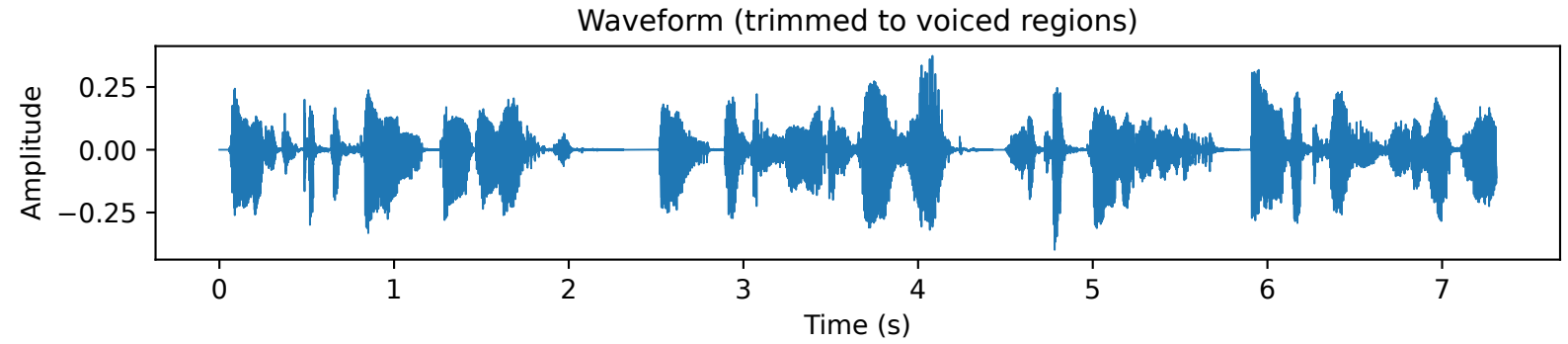
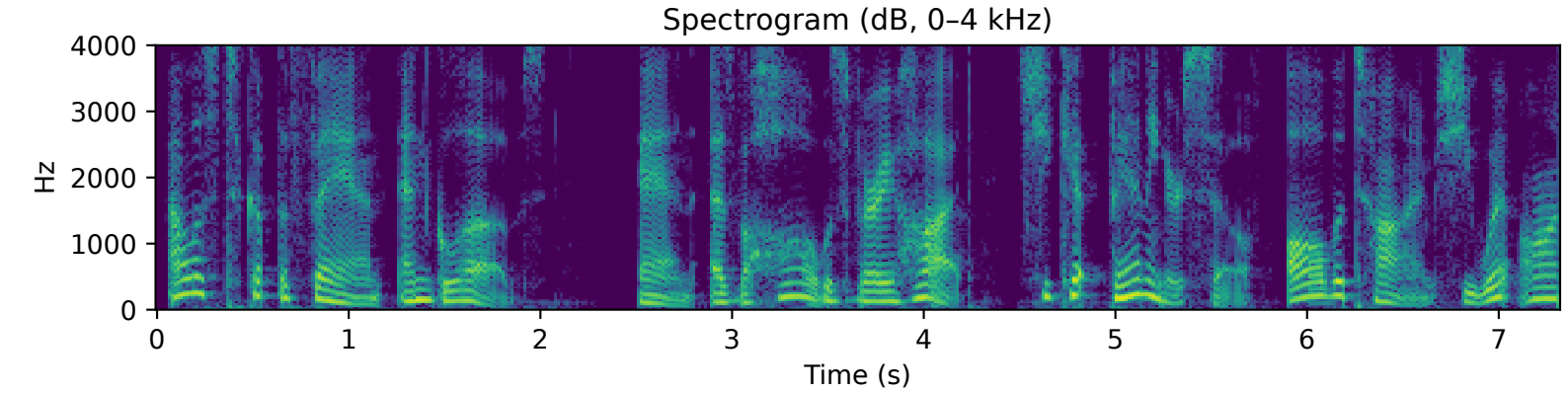
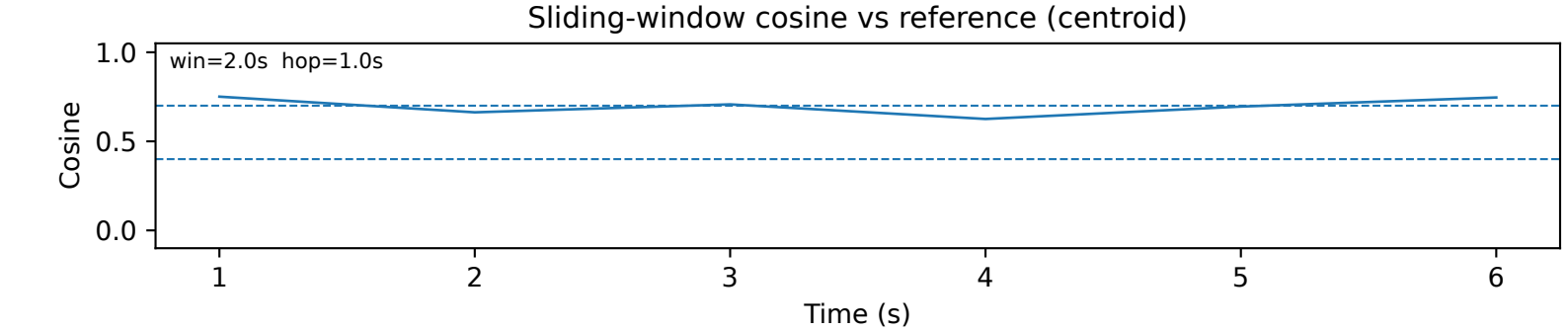
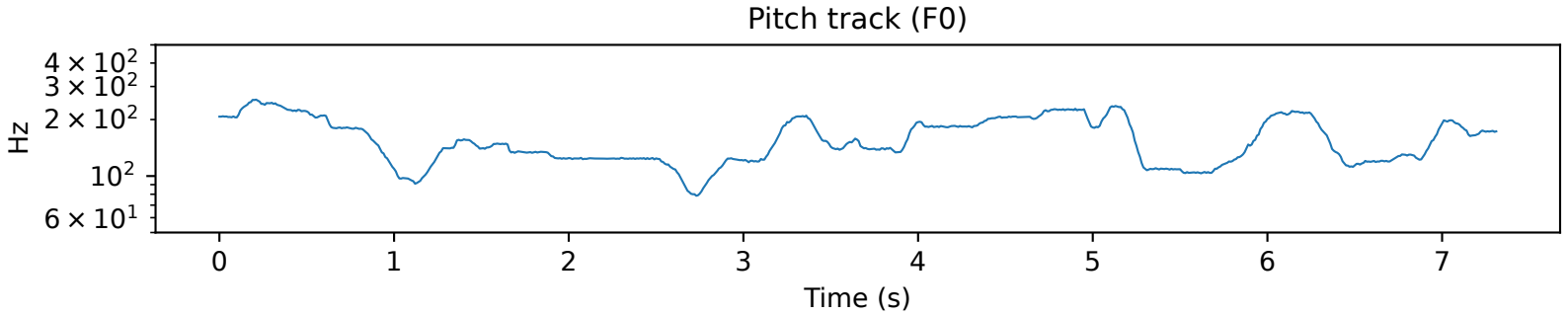
Forensic Voice Report — label: libris_260_clean

Report ID: 20251027T131347Z_dc3df386 | File: 260-123440-0004.wav | SHA256: c70586705576…
Voiced: 7.31s (VR=0.94) | SNR: 8.5 dB | Scores — Best: 0.94 | Top3 mean: 0.52 | Centroid: 0.90
Decision: high_match (thr: med=0.40, high=0.70) | CM: 0.34 | F0 median: 147.27 Hz (IQR 72.10 Hz)









Per-segment (voiced ≥ 0.2 s)

Start (s)	End (s)	Dur (s)	Cosine	Flag
0.00	7.00	7.00	0.70	HIGH

See next page for Top Enrollments and Full Provenance / Methods.

Top Enrollments and Provenance

Top enrollments by cosine

Enroll ID	Created	Cosine
a5f19f64	2025-10-25 23:15:04	0.94
d485c61c	2025-10-25 23:15:12	0.32
1b32d906	2025-10-25 23:06:20	0.29
43c245c5	2025-10-25 23:14:51	0.29
bc4bac39	2025-10-25 23:32:58	0.24
f8645e67	2025-10-25 23:34:09	0.20

Provenance & Methods

File: 260-123440-0004.wav Bytes: 250638 | SHA256: c70586705576bd458accabf9addfb4bf6cac497816a85e3e2fc3f54f984d16b6 Orig sample rate: 16000 Hz | Channels: 1 | Bit depth: 16 Duration: 7.83 s System: Windows 11 | Host: DESKTOP-32VU5AV | User: asus Python: 3.13.7 | numpy: 2.3.3 | torch: 2.8.0+cpu | torchaudio: 2.8.0+cpu speechbrain: 1.0.3 | torchcrepe: None Config snapshot: thresholds (med=0.40, high=0.70); VAD mode=2, pad_ms=50; scoring: best & mean(top-3) vs enrollments; centroid cosine.

Speech isolation (VAD): WebRTC VAD, 20 ms frames, mode 2, silence trimmed with ±50 ms padding. Voiced ratio = fraction of voiced frames. Embedding extraction: ECAPA (speechbrain/spkrec-ecapa-voxceleb) at 16 kHz mono; embeddings L2-normalized. Similarity scoring: primary = cosine to each enrollment for the label; report best and mean(top-3). Secondary = cosine to label centroid (mean of enrollments). Decision policy: thresholds medium=0.40, high=0.70 → label as low/medium/high match. Anti-spoofing (optional): CM score is reported if available.