

Formule aritmetiche – LoopbackSampler.cs

Generato: 2025-10-25 17:41:52

Notazione

- N = numero di frame nel blocco; SR = sample rate (Hz); dt = tempo tra due callback (s).
- L_i, R_i = campioni sinistro/destro in $[-1, +1]$; $|x|$ = valore assoluto; \log_{10} = logaritmo in base 10.
- $EPS = 1e-12$ (evita log di 0). Tutti i dB sono dBFS salvo diversa indicazione.

Livelli base (RMS, picco, hold, dB)

- $RMS_L = \sqrt{(1/N) * \sum_i (L_i^2)}$; $RMS_R = \sqrt{(1/N) * \sum_i (R_i^2)}$
- $Peak_L = \max_i |L_i|$; $Peak_R = \max_i |R_i|$
- $dB(x) = 20 * \log_{10}(\max(x, EPS))$
- $PeakHold_L(t) = \max(Peak_L, PeakHold_L(t-1) * 10^{(-decay_dBps * dt / 20)})$
- $PeakHold_R(t) = \max(Peak_R, PeakHold_R(t-1) * 10^{(-decay_dBps * dt / 20)})$
- Se silenzio: $PeakHold_* \leftarrow PeakHold_* * 0.5^{dt}$ (dimezza ogni 1 s)
- Silence (boolean): $20 * \log_{10}(\max(RMS_L, RMS_R) + EPS) < -65$

DC offset (EMA)

- $\alpha = 1 - \exp(-1 / (\tau * SR))$, con $\tau \approx 0.5$ s
- $DC_L(t) = DC_L(t-1) + \alpha * (L_t - DC_L(t-1))$; idem per $DC_R(t)$

Correlazione e Balance

- $\mu_L = (1/N) \sum_i L_i$; $\mu_R = (1/N) \sum_i R_i$
- $S_{LL} = \sum_i (L_i^2) - N * \mu_L^2$; $S_{RR} = \sum_i (R_i^2) - N * \mu_R^2$
- $S_{LR} = \sum_i (L_i * R_i) - N * \mu_L * \mu_R$
- $corr = S_{LR} / \sqrt{\max(1e-18, S_{LL} * S_{RR})}$, troncato in $[-1, +1]$
- $balance = (RMS_R - RMS_L) / (RMS_R + RMS_L)$, troncato in $[-1, +1]$

Crest factor (dB)

- $Crest_L_dB = 20 * \log_{10}(Peak_L / \max(RMS_L, 1e-9))$
- $Crest_R_dB = 20 * \log_{10}(Peak_R / \max(RMS_R, 1e-9))$

Mid/Side & Width

- $m_i = 0.5 * (L_i + R_i)$; $s_i = 0.5 * (L_i - R_i)$
- $\Sigma M2 = \sum_i (m_i^2)$; $\Sigma S2 = \sum_i (s_i^2)$
- $Width_dB = 10 * \log_{10}(\max(\Sigma S2, 1e-18) / \max(\Sigma M2, 1e-18))$
- MonoCompat (proxy) = corr

True-peak (parabolico + 4x Catmull–Rom)

- Parabolico su $(a, b, c) = (prev2, prev1, curr)$:
 - $x_v = 0.5 * (a - c) / (a - 2b + c)$, limitato a $[-0.5, 0.5]$
 - $TP_{par} = |b - 0.25 * (a - c) * x_v|$
- Catmull–Rom 4x ($p0, p1, p2, p3$):
 - $f(t) = a0 t^3 + a1 t^2 + a2 t + a3$, con
 - $a0 = -0.5*p0 + 1.5*p1 - 1.5*p2 + 0.5*p3$
 - $a1 = p0 - 2.5*p1 + 2.0*p2 - 0.5*p3$
 - $a2 = -0.5*p0 + 0.5*p2$
 - $a3 = p1$
 - $TP_{over} = \max(|f(0.25)|, |f(0.50)|, |f(0.75)|)$
- True-peak per canale: $TP = \max(TP_{par}, TP_{over})$
- $dBTP = 20 * \log_{10}(\max(TP, EPS))$; clipping se $TP \geq 1.0$

Finestra di Hann e normalizzazione FFT (mid = (L+R)/2)

- $w[n] = 0.5 * (1 - \cos(2\pi n / (N-1)))$, $n = 0..N-1$
- Somme: $S = \sum_n w[n]$; $D_{main} = 0.5 * S$; $D_{edge} = S$
- Per ogni bin k di FFT reale ($0..N/2$):
 - $mag_k = \sqrt{Re_k^2 + Im_k^2}$ (nessuna divisione per N)
 - $D_k = D_{edge}$ se $k \in \{0, N/2\}$ altrimenti D_{main}
 - $amp_k = mag_k / D_k$ (ampiezza di picco)
 - $amp_rms_k = amp_k / \sqrt{2}$ (tranne $k=0, N/2$)

- $\text{dBFS}_k = 20 * \log_{10}(\max(\text{amp_rms}_k, 1e-12))$, limitato a [-120, 0] dB

Feature spettrali

- Dominant bin $k_{\max} = \operatorname{argmax}_{\{k \geq 1\}} \text{dBFS}_k$
- Interpolazione parabolica (in ampiezza lineare P):
 - $\delta = 0.5 * (P_{\{k-1\}} - P_{\{k+1\}}) / (P_{\{k-1\}} - 2P_k + P_{\{k+1\}})$, limitato a [-0.5, 0.5]
 - $f_{\text{dom}} = (k_{\max} + \delta) * \text{SR} / N$
- Energia per bin: $E_k = (\text{amp_rms}_k)^2$; $f_k = k * \text{SR} / N$
- Centroid: $f_c = (\sum_k f_k * E_k) / (\sum_k E_k)$
- Roll-off 95%: più piccolo K con $\sum_{\{k=0..K\}} E_k \geq 0.95 * \sum_{\{k=0..N/2\}} E_k$; $f_{\text{roll}} = K * \text{SR} / N$
- Noise floor (dB): 10° percentile di $\{\text{dBFS}_k\}$ (esclusi DC, Nyquist e $k_{\max} \pm 1$)
- SNR (dB): $\text{Psig} = \sum_{\{k=k_{\max}-1..k_{\max}+1\}} E_k$; $\text{Pnoise} = \sum_{\{\text{altri } k\}} E_k$; $10 * \log_{10}(\text{Psig}/\text{Pnoise})$
- ENOB (bit): $(\text{SNR} - 1.76) / 6.02$

Loudness ITU-R BS.1770 / EBU R128 (approssimazione)

- K-weighting per campione: applica HP (~38 Hz, 2° ordine) + high-shelf (+4 dB @ ~1682 Hz) a L e R.
- Energia istantanea: $\text{ms}_t = L_{\text{kw}}^2 + R_{\text{kw}}^2$ (somma canali)
- Finestra scorrevole (winSec): $\text{MeanSquare} = \sum(\text{ms}_t * dt) / \sum(dt)$
- $\text{LUFS}_M = 10 * \log_{10}(\text{MeanSquare}_M) + K_{\text{offset}}$, con win M = 0.4 s
- $\text{LUFS}_S = 10 * \log_{10}(\text{MeanSquare}_S) + K_{\text{offset}}$, con win S = 3.0 s
- Integrata (blocchi da 0.4 s):
 - 1) Scarta blocchi con $\text{LUFS} < -70$ (absolute gate)
 - 2) Media preliminare: $\text{ms}_\text{pre} = \text{media}(\text{ms}_\text{gated})$; $\text{LUFS}_\text{pre} = 10 * \log_{10}(\text{ms}_\text{pre}) + K_{\text{offset}}$
 - 3) Soglia relativa: $\text{gate}_\text{rel} = \text{LUFS}_\text{pre} - 10 \rightarrow \text{ms}_\text{gate}_\text{rel} = 10^{((\text{gate}_\text{rel} - K_{\text{offset}})/10)}$
 - 4) $\text{LUFS}_I = 10 * \log_{10}(\text{media dei } \text{ms}_j \geq \text{ms}_\text{gate}_\text{rel}) + K_{\text{offset}}$
- LRA (rolling su 60 s): $\text{LRA} = \text{P95}(\text{LUFS}_S) - \text{P10}(\text{LUFS}_S)$, dopo filtro: $\text{LUFS}_S \geq \max(-70, \text{LUFS}_I - 20)$
- PLR (dB): $\max(\text{dBTP}_L, \text{dBTP}_R) - \text{LUFS}_S$; PSR (dB): $\max(\text{dBFS}_\text{peak}_L, \text{dBFS}_\text{peak}_R) - \text{LUFS}_S$

Definizione variabili principali

- L_i, R_i : campioni audio normalizzati [-1, +1]; SR: sample rate in Hz; N: numero di frame nel blocco.
- dt: durata reale del blocco processato (frames/SR); dt (peak-hold): tempo reale trascorso tra due callback.
- $K_{\text{offset}} = -0.691$ (secondo ITU-R BS.1770).
- $w[n]$: finestra di Hann; D_main, D_edge: normalizzatori per la stima di ampiezza della FFT.

— Fine —