

## 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<sub>i</sub>, R<sub>i</sub> = campioni sinistro/destro in [-1, +1]; |x| = valore assoluto; log<sub>10</sub> = 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:  $PeakHold\_L \leftarrow PeakHold\_L * 0.5^{dt}$  (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)

- $w[n] = 0.5 * (1 - \cos(2\pi n / (N-1)))$  ,  $n = 0..N-1$
- Somma:  $S = \sum_n w[n]$  ;  $D\_main = 0.5 * S$  ;  $D\_edge = S$
- 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  $k_{\max} = \arg\max_{\{k \geq 1\}} \text{dBFS}_k$
- 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):  $\max(\text{dBFS}_{\text{peak}_L, R}) - \text{LUFS}_S$

## 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;  $D_{\text{main}}, D_{\text{edge}}$ : normalizzatori per la stima di ampiezza della FFT.

— Fine —