

Università degli Studi di Padova – Dipartimento di Ingegneria
dell'Informazione

Corso di Laurea in Ingegneria Informatica

Relazione per la prova finale

«Stima del tempo di riverbero da registrazioni musicali: un
approccio blind basato su Maximum Likelihood e Spectral
Flatness»

Relatore: Prof. Sergio Canazza

Laureanda: Falezza Francesca
Matricola n.: 2101819

Padova, 21/07/2026

INTRODUZIONE

Il lavoro presentato fa parte delle attività svolte durante il tirocinio presso il **Centro di Sonologia Computazionale**.

Si inserisce all'interno dell'area di ricerca **Cultural Heritage**.

Durante il tirocinio presso il CSC è stato svolto anche un lavoro di **divulgazione**, contribuendo alla creazione della pagina Wikipedia del laboratorio.

OBIETTIVO

Nell'articolo di riferimento "**An improved algorithm for blind reverberation time estimation**" viene presentato un nuovo metodo di stima cieca del tempo di riverbero pensato per segnali di parlato.

L'obiettivo di questa tesi è **adattare tale algoritmo a registrazioni musicali** mantenendo una accuratezza accettabile.

CASO STUDIO: TEATRO LA FENICE

Il Teatro La Fenice è un caso studio ottimale poiché le misurazioni effettuate da Angelo Farina pochi mesi prima dell'incendio del 1996 offrono un riferimento scientificamente solido con cui validare l'efficacia di nuovi algoritmi applicati a registrazioni storiche.

PARAMETRI PRINCIPALI

$$T30 = 1.8s$$

Intervallo di tempo nel quale
l'energia di campo sonoro
diminuisce di 30 dB
sotto il suo livello iniziale.

$$EDT = 1.25 s$$

Tempo di decadimento, ottenuto
estrapolando a 60 dB la pendenza
calcolata sui primi 10 dB di
decadimento del suono dopo
l'interruzione della sorgente.

ALGORITMO DI LÖLLMANN

$$d_m(k) = A r^* v(k)^* e^{-\rho k T_s^*}(k)$$

MODELLO DEL DECADIMENTO

come processo casuale discreto

$$T_{60} \cong \frac{6.908}{\rho}$$

RELAZIONE

tasso di decadimento e T60

Down-sampling e preselezione
Stima massima verosimiglianza (MLE)
Istogrammi adattivi per stima finale

METODI

per ridurre carico computazionale e
migliorare le stime

LIMITI DEL METODO LÖLLMANN SULLA MUSICA

□ NATURA DEL SEGNALE

L'algoritmo assume che il decadimento sia un processo stocastico e gaussiano, ipotesi valida per il parlato non per la musica.

□ STRUTTURA TEMPORALE

L'algoritmo non isola correttamente la coda sonora nella musica, poiché è progettato per funzionare sulle pause naturali tra fonemi del parlato.

□ CONTENUTO SPETTRALE

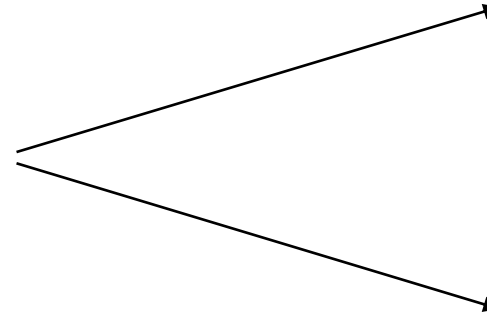
Il metodo di Löllmann fallisce perché scambia la struttura armonica e il sustain degli strumenti per riverbero.

CONTRIBUTO ORIGINALE

Selezione basata su **SPECTRAL FLATNESS**
Permette di distinguere quanto un suono sia simile a
un **tono** rispetto a quanto sia simile a un **rumore**

$SFM = \text{media geometrica} / \text{media aritmetica}$ dello
spettro di potenza del segnale.

CODA RIVERBERANTE viene modellata come un
RUMORE DIFFUSO con un decadimento energetico
esponenziale, caratterizzato da una **SFM elevata**



1 per rumore bianco (segnale
"piatto"),

0 per segnali tonali (strumenti
musicali).

SUPERAMENTO DEI LIMITI DEL METODO LÖLLMANN

✓ NATURA DEL SEGNALE

Filtrare per SF crescente significa escludere la porzione non Gaussiana dello strumento e garantisce che il segmento analizzato sia *noise-like*.

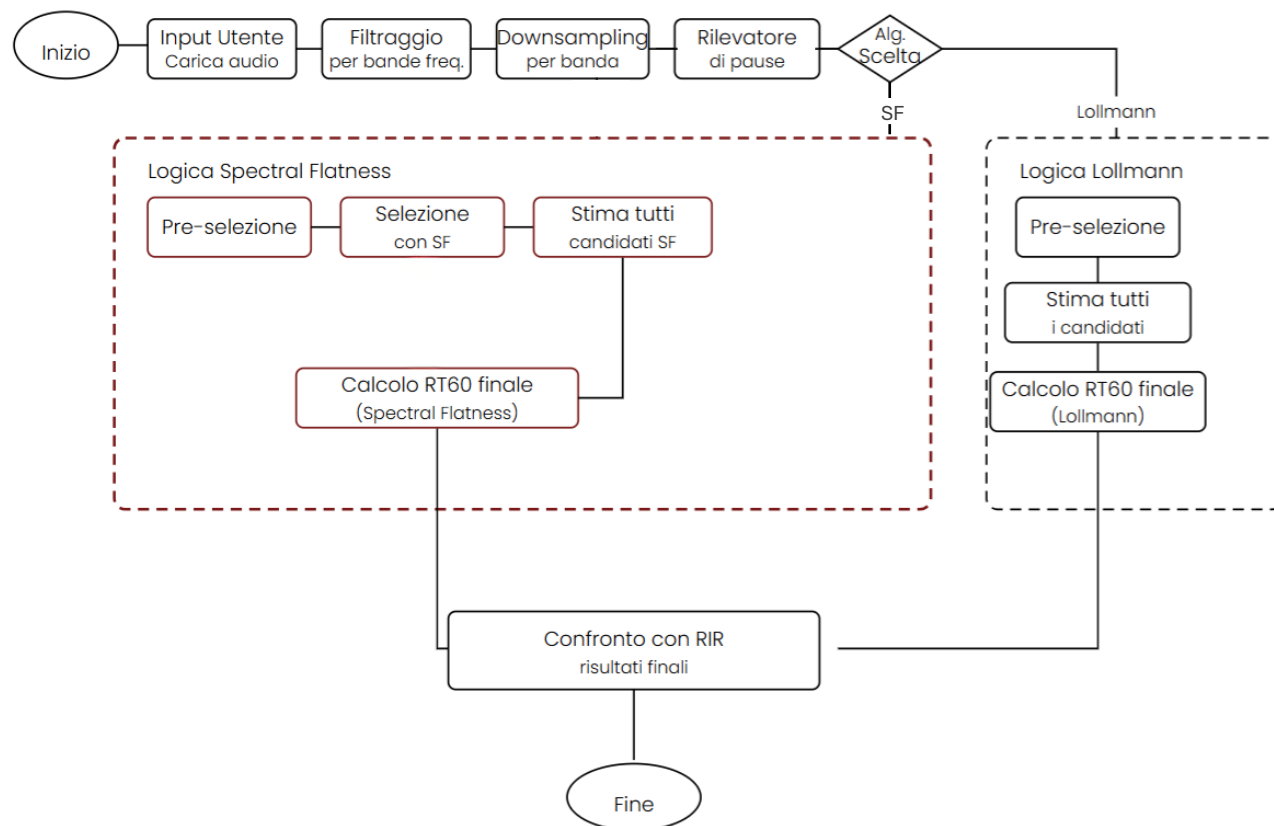
✓ STRUTTURA TEMPORALE

Il filtro funge da rilevatore di transizione. Passaggio SF bassa → SF alta indica che la sorgente originaria ha smesso di emettere.

✓ CONTENUTO SPETTRALE

Applicando SF si selezionano segmenti con spettro piatto, si isola il momento in cui la coda diventa un processo stocastico puro.

FLUSSO DI ELABORAZIONE



FUNZIONI PRINCIPALI:

Pre-selezione

Implementa la pre-selezione basata su 3 condizioni che i frame di possibile decadimento devono soddisfare.

Selezione con spectral flatness

Per ogni candidato già pre-selezionato, divide il segnale in sotto-frame e applica una ulteriore selezione su spectral flatness

Stima di un candidato

Prende un singolo segmento di decadimento e stima il parametro α (coefficiente di decadimento esponenziale) tramite massima verosimiglianza, e quindi l'RT60

Aggiornamento istogrammi

Risponde alla domanda "qual è il valore di RT60 più frequente nelle ultime N stime?".

DATASET DELL'ANALISI

SORGENTE

Registrazione del Teatro La Fenice del 1976 del secondo atto dell'opera *Madama Butterfly*.

9 SEGMENTI estratti.



N segmenti	Durata
3	120 s
3	240 s
3	480 s

RIFERIMENTO

Misurazioni condotte da Angelo Farina per il T30

Media ISO del T30 $\cong$ 1.5 s

Fattore correttivo del 15% per considerare l'assorbimento del pubblico.



Media ISO T60 corretto = 1.275s

LÖLLMANN vs SF

Segm.	Durata (s)	ISO Reale	Media ISO senza SF	Err. No SF	Media ISO con SF	Err. SF
S1	120	1.275	0.985	22.75%	1.2	5.88%
S2	120	1.275	0.915	28.24%	1.32	3.53%
S3	120	1.275	0.89	30.20%	1.17	8.24%
S4	240	1.275	1.125	11.76%	1.16	9.02%
S5	240	1.275	0.915	28.24%	1.015	20.39%
S6	240	1.275	0.9	29.41%	1.02	20.00%
S7	480	1.275	1.085	14.90%	1.29	1.18%
S8	480	1.275	0.875	31.37%	1.045	18.04%
S9	480	1.275	1.18	7.45%	1.23	3.53%

MIGLIORAMENTO maggiore sui segmenti da 2 minuti perché SF agisce come filtro di qualità e riduce il numero di falsi positivi.

S5, S6 e S8 presentano errori elevati anche con Spectral Flatness. Causati probabilmente da caratteristiche del segnale che rendono **difficile isolare la coda riverberante**.

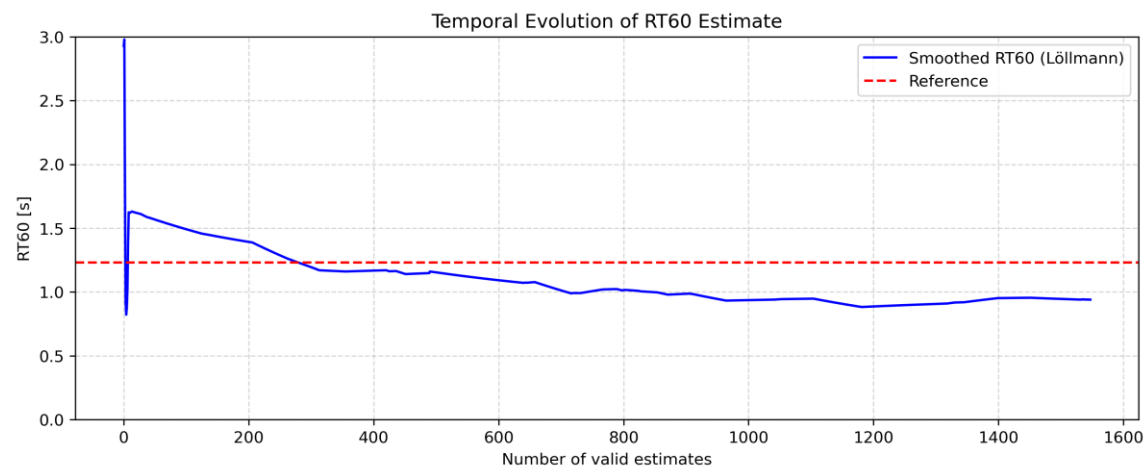
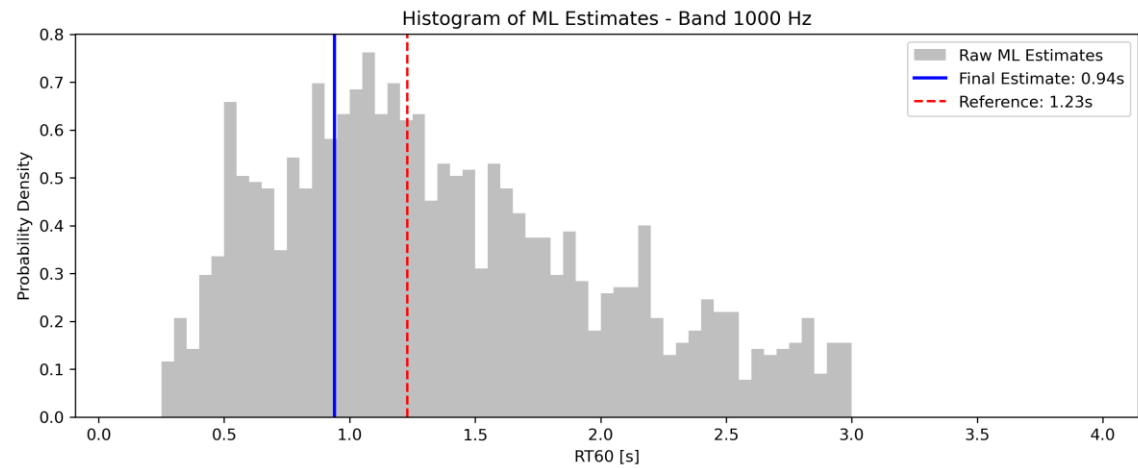
CONFRONTO PER BANDE DI FREQUENZA

SOTTOSTIMA sistematica per le basse frequenze, causata principalmente dalla finestra fissa di analisi.

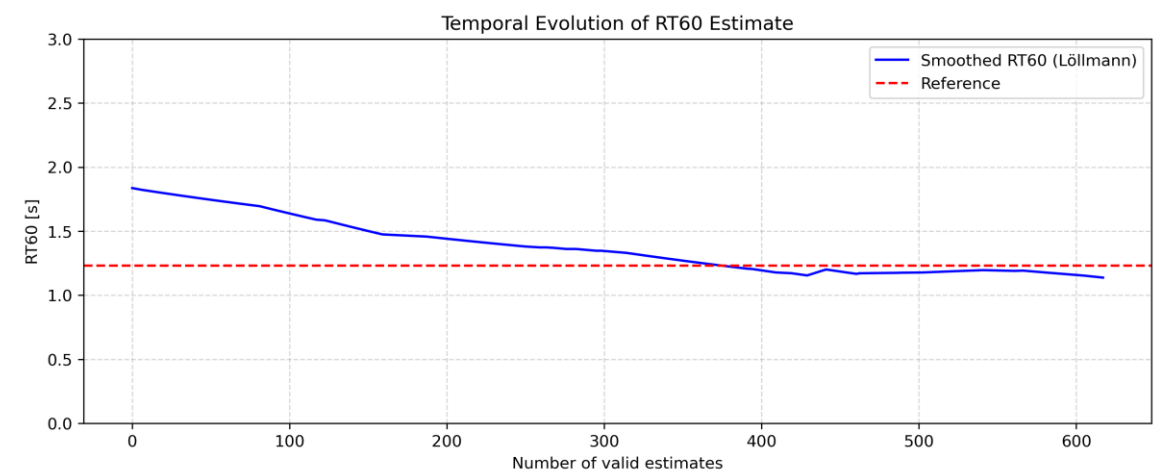
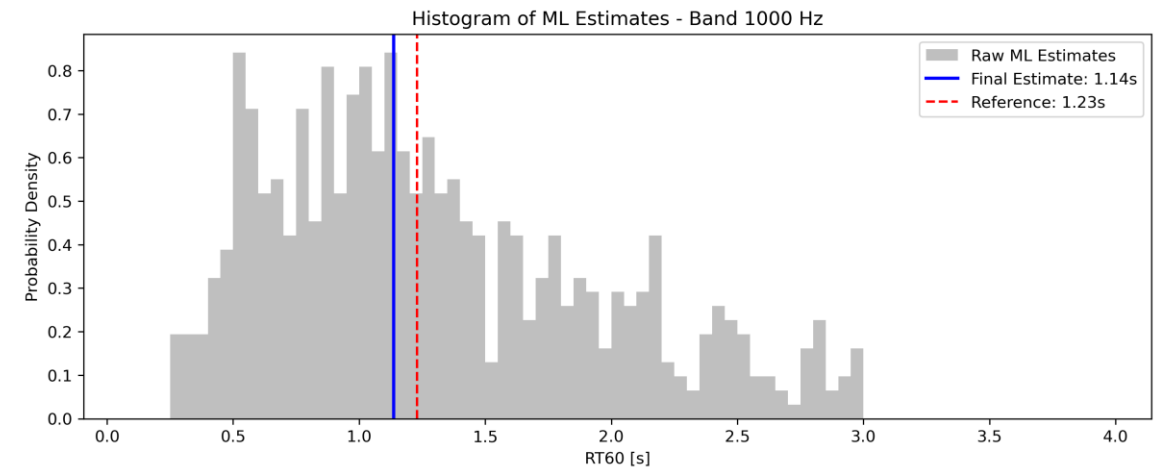
ERRORE MEDIO 24.6% → 12.1%

Parametro	125	250	500	1000	2000	4000	Media ISO	Media Tot.*
RT60	1.70	1.57	1.32	1.23	1.15	0.94	1.275	1.32
Stima no SF	1.56	1.11	0.84	0.94	0.97	0.55	0.825	0.995
Errore	8.2%	29.3%	36.4%	23.6%	15.7%	41.5%	35.3%	24.6%
Stima SF	1.72	1.08	1.20	1.14	1.20	0.62	1.17	1.16
Errore	1.2%	31.2%	9.1%	7.3%	4.3%	34.0%	8.2%	12.1%

Algoritmo Löllmann



Algoritmo SF



CONCLUSIONI E SVILUPPI

L'introduzione della selezione per Spectral Flatness migliora le stime per 9 segmenti su 9

LIMITI

Dipendenza dalla durata del segnale.

Instabilità della stima per singola banda alle frequenze più basse.

Lunghezza fissa del frame $M=1024$ campioni (116ms). La stima è meno precisa quando il frame copre solo una piccola frazione del decadimento reale.

SVILUPPI FUTURI

Adattamento della lunghezza del frame alla banda frequenziale.

Soglia SF dipendente dalla banda .

BIBLIOGRAFIA

- [1] H. W. Löllmann et al, "An improved algorithm for blind reverberation time estimation," in *Proc. IWAENC*, 2010.
- [2] M. R. Schroeder, "New method of measuring reverberation time," *J. Acoust. Soc. Am.*, vol. 37, pp. 409–412, 1965.
- [3] R. Ratnam et al, "Blind estimation of reverberation time," *J. Acoust. Soc. Am.*, vol. 114, no. 5, pp. 2877–2892, 2003.
- [4] J. Bradley, "Using ISO 3382 measures, and their extensions, to evaluate acoustical conditions in concert halls," *Acoust. Sci. Technol.*, vol. 26, pp. 170–178, 2005.
- [5] S. Dubnov, "Generalization of spectral flatness measure for non-Gaussian linear processes," *IEEE Signal Process. Lett.*, vol. 11, no. 8, pp. 698–701, 2004.
- [6] L. Tronchin and A. Farina, "Acoustics of the former Teatro La Fenice in Venice," *J. Audio Eng. Soc.*, vol. 45, no. 12, pp. 1051–1062, 1997.
- [7] N. Prodi et al, "Acoustics of Italian historical opera houses," *J. Acoust. Soc. Am.*, vol. 138, no. 2, pp. 769–781, 2015.
- [8] H. Kuttruff, *Room Acoustics*. Boca Raton, FL: CRC Press, 2016.
- [9] L. L. Beranek, *Concert Halls and Opera Houses: Music, Acoustics, and Architecture*, vol. 2. New York, NY: Springer, 2004.
- [10] J. D. Johnston, "Transform coding of audio signals using perceptual noise criteria," *IEEE J. Sel. Areas Commun.*, vol. 6, no. 2, pp. 314–323, 1988.
- [11] G. Peeters, "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," IRCAM, Paris, France, Tech. Rep., 2004.
- [12] J. F. Santos, N. Peters, and T. H. Falk, "Towards blind reverberation time estimation for non-speech signals," *Proc. Mtgs. Acoust.*, vol. 19, no. 1, p. 055059, 2013.
- [13] Nuniz, "blind_rt60," GitHub repository, 2024. [Online]. Available: https://github.com/nuniz/blind_rt60
- [14] H. Neuschmied, H. Mayer, and E. Batlle, "Content-based identification of audio titles on the Internet," in *Proc. 1st Int. Conf. WEB Delivering of Music (WEDELMUSIC)*, Florence, Italy, 2001, pp. 96–100.
- [15] Y. Ruoran and L. Tronchin, "Acoustic characteristics and objective evaluation of acoustic quality in historical Italian theaters," *Appl. Acoust.*, vol. 233, p. 110483, 2025.
- [16] L. Tronchin, "Advanced techniques for measuring and reproducing spatial sound properties of auditoria," in *Proc. RADS*, 2004.