



**Politecnico  
di Torino**

# Politecnico di Torino

Corso di Laurea Magistrale in  
Ingegneria del Cinema e dei Mezzi di Comunicazione

A.a. 2023/2024

Sessione di Laurea dicembre 2024

## **Restauro del sonoro cinematografico**

Confronto tra tecnologie di scansione e panoramica dei  
software di nuova generazione basati sulle reti neurali.

Relatori:

Prof. Vito MARTINELLI  
Prof. Antonio SERVETTI  
Dott. Gabriele Angelo PERRONE

Candidato:

Ismaele RODELLA



# Indice

<b>Abstract</b>	<b>7</b>
<b>1. INTRODUZIONE AL RESTAURO CINEMATOGRAFICO</b>	<b>9</b>
<b>1.1 Definizione e contesto</b>	<b>9</b>
<b>1.2 Formati e supporti analogici per il sonoro cinematografico</b>	<b>12</b>
1.2.1 Supporto ottico	12
1.2.2 Supporto magnetico	14
<b>1.3 Problematiche di degradazione dei supporti analogici</b>	<b>15</b>
1.3.1 Degradazione dei supporti ottici	15
1.3.2 Degradazione dei supporti magnetici	17
<b>1.4 Caso studio: “Miti e riti dell’India”</b>	<b>20</b>
<b>2. TECNOLOGIE E METODI DI SCANSIONE</b>	<b>22</b>
<b>2.1 Tipologie di scanner</b>	<b>22</b>
2.1.1 Telecine	23
2.1.2 Film scanner	24
2.1.3 Datacine	27
<b>2.2 Lasergraphics Scanstation</b>	<b>27</b>
<b>2.2 Sondor OMA/E</b>	<b>28</b>
<b>3. ANALISI DEI SEGNALI DIGITALIZZATI</b>	<b>31</b>
<b>3.1 Disturbi locali</b>	<b>31</b>
<b>3.2 Disturbi globali</b>	<b>33</b>
<b>3.3 Degradazioni presenti in “Miti e riti dell’India”</b>	<b>34</b>
3.3.1 Click, pop, crackle e clipping	34
3.3.2 Hum	36
3.3.3 Disturbi isolati	36
<b>3.4 Confronto delle due scansioni</b>	<b>38</b>
3.4.1 Confronto macroscopico	39
3.4.2 Confronto di segmenti specifici	42
3.4.2.1 Rumore di fondo	43

3.4.2.2	Segmenti generici	48
3.4.2.3	Ronzio sui segmenti parlati	51
3.4.2.4	Disturbi locali impulsivi	53
3.4.2.5	Andamento del centroide spettrale	57
3.4.2.6	Andamento del parametro di roughness	57
<b>4.</b>	<b>OPERAZIONI DIGITALI DI RESTAURO</b>	<b>59</b>
<b>4.1</b>	<b>Rimozione dei click</b>	<b>60</b>
<b>4.2</b>	<b>Rimozione del rumore di fondo a larga banda</b>	<b>62</b>
<b>4.3</b>	<b>Rimozione della riverberazione</b>	<b>65</b>
<b>4.4</b>	<b>Separazione delle sorgenti</b>	<b>67</b>
<b>5.</b>	<b>CONTRIBUTO DELLE RETI NEURALI ALLE TECNICHE DI RESTAURO</b>	<b>69</b>
<b>5.1</b>	<b>Reti neurali per la rilevazione automatica dei difetti</b>	<b>70</b>
<b>5.2</b>	<b>Reti neurali per la rimozione del rumore</b>	<b>71</b>
<b>5.3</b>	<b>Reti neurali per algoritmi di de-reverberation</b>	<b>72</b>
<b>5.4</b>	<b>Reti neurali per la separazione delle sorgenti</b>	<b>73</b>
<b>6.</b>	<b>PANORAMICA DEI SOFTWARE</b>	<b>75</b>
<b>6.1</b>	<b>Software commerciali</b>	<b>75</b>
6.1.1	Izotope RX	75
6.1.2	Steinberg SpectraLayers	78
6.1.3	Waves Clarity	79
6.1.4	Supertone CLEAR	82
6.1.5	Acon Digital Restoration Suite	83
<b>6.2</b>	<b>Implementazioni open-source</b>	<b>86</b>
6.2.1	Ultimate Vocal Remover	86
6.2.2	Werman Noise Suppression	87
6.2.3	MPAI CAE-ARP	88
<b>7.</b>	<b>CONCLUSIONI</b>	<b>89</b>
<b>7.1</b>	<b>Esempi di intervento su “Miti e riti dell’India”</b>	<b>89</b>
<b>7.2</b>	<b>Osservazioni conclusive</b>	<b>91</b>
<b>7.3</b>	<b>Futuri sviluppi</b>	<b>93</b>

<b>Bibliografia e sitografia</b>	<b>95</b>
<b>Elenco delle figure</b>	<b>100</b>
<b>Ringraziamenti</b>	<b>105</b>



# Abstract

L'argomento principale di questa tesi di laurea è il restauro del sonoro cinematografico, le sue prassi e le sue implicazioni etiche.

La tesi analizza le diverse problematiche di degrado che possono colpire i supporti analogici, come film ottici e nastri magnetici, e intende fornire una panoramica dettagliata delle tecnologie di scansione e delle tecniche digitali di restauro.

Particolare attenzione è rivolta al confronto tra due dispositivi di scansione: il Lasergraphics Scanstation, in dotazione alla Cineteca del Museo Nazionale del Cinema di Torino e il Sondor OMA/E, in dotazione al laboratorio "L'immagine Ritrovata" della Cineteca di Bologna. Sono stati effettuati confronti spettrali, analizzando svariati segmenti specifici delle tracce per trarne conclusioni caratterizzanti.

Infine, la tesi presenta una serie di software commerciali e open-source utilizzati per il restauro del suono, offrendo una overview completa delle principali soluzioni disponibili per questa tipologia di intervento, con una particolare attenzione ai recenti contributi apportati dall'implementazione delle reti neurali.



# 1. INTRODUZIONE AL RESTAURO CINEMATOGRAFICO

## 1.1 Definizione e contesto

Restaurare un'opera d'arte significa preservare un oggetto di interesse artistico, garantire la conservazione di un bene culturale, ma anche permettere il mantenimento di una testimonianza storica, in quanto l'opera filmica è influenzata in maniera sostanziale dal contesto sociale e dal periodo in cui è stata realizzata. Esistono diversi approcci finalizzati alla preservazione di queste opere: la duplicazione, il restauro e la riedizione.<sup>1</sup>

La duplicazione consiste nella realizzazione di una copia del film su un nuovo supporto. È un processo fondamentale per la conservazione del patrimonio cinematografico e vuole rispondere "sia alla necessità di recuperare la funzionalità del film (poterlo proiettare e stampare), sia a quelle di assicurare la sopravvivenza dell'opera (grazie a supporti più stabili)".<sup>2</sup>

I supporti analogici sono spesso soggetti ad una fragilità che aumenta con il passare del tempo e al peggiorare delle condizioni di conservazione, perciò la duplicazione è finalizzata a congelare lo stato di un'opera, fermando gli effetti che il degrado chimico-fisico del supporto sta avendo su di essa. La duplicazione è un processo tutt'altro che trasparente: venendo attuato nel dominio analogico, comporta intrinsecamente l'introduzione di modifiche quali la perdita di qualità dell'immagine e del suono. Inoltre, nella maggior parte dei casi, le tecnologie utilizzate per tale operazione non sono quelle originali con cui il film è stato realizzato, bensì pensate per la produzione moderna. Per questi motivi, la scelta delle tecnologie e dei metodi di duplicazione è cruciale per garantire la migliore qualità possibile del duplicato.<sup>3</sup>

---

<sup>1</sup> Plutino, 2020, 19-22

<sup>2</sup> Venturini, 2006, 91

<sup>3</sup> ibidem

Il restauro, invece, si pone l'obiettivo di affrontare il deterioramento del film, cercando di riportarlo, quanto più possibile, al suo stato originale, con un opportuno trasferimento su un supporto digitale (*riversamento*<sup>4</sup>), con la conseguente possibilità di effettuare, su tale copia, modifiche prima di registrarlo nuovamente in un formato adatto alla visualizzazione.<sup>5</sup>

Il restauro comporta un workflow più completo rispetto alla duplicazione: produce sia un duplicato del film allo stato originale, sia delle copie restaurate, (e quindi migliorate) grazie alla rimozione dei danni accumulatisi nel tempo sul supporto, consentendone la riproduzione moderna.<sup>6</sup> L'obiettivo è quello di compensare la degradazione del film, ripristinando elementi che sono stati compromessi quali colore, contrasto, rumori di fondo e disturbi, al fine di migliorare la leggibilità del film in tutte le sue dimensioni. Per questo motivo, il restauro cinematografico è un processo filologico che comporta spesso operazioni editoriali, di montaggio e di ricostruzione.<sup>7</sup>

È bene chiarire che il restauro non è un processo totalmente digitale, in quanto comincia già nel dominio analogico. Molte procedure attuate per la compensazione delle degradazioni, ad esempio, la rimozione di graffi, la riparazione di strappi, la pulizia da polvere e sporco, il consolidamento di emulsioni instabili e la sostituzione di giunte danneggiate, avvengono fisicamente sul supporto stesso e prima della transizione al dominio digitale<sup>8</sup>: un passaggio fondamentale, che rappresenta comunque, e sempre, un rischio. La digitalizzazione, infatti, introduce la possibilità dell'appiattimento dell'estetica a cui l'opera appartiene e della perdita della materialità del supporto che ha in sé una dimensione temporale e una presenza fisica.<sup>9</sup> Inoltre, le operazioni di restauro digitale, se fatte senza la dovuta

---

<sup>4</sup> Con *riversamento* si intende il trasferimento di un segnale da un supporto ad un altro. Deve avvenire esclusivamente a partire dal contenuto del supporto originale, se disponibile, e non da una sua copia, che potrebbe portare con sé manipolazioni di cui potremmo non essere a conoscenza e su cui non potremmo avere controllo. (Canazza, 2013, 3-9)

<sup>5</sup> Venturini, 2006, 38

<sup>6</sup> Plutino, 2020, 19

<sup>7</sup> Venturini, 2006, 95

<sup>8</sup> *ivi*, 80

<sup>9</sup> *ivi*, 37

consapevolezza ed esperienza, sono pericolose per il contenuto. I software utilizzati sono molto efficaci e la linea di separazione tra miglioramento e degradazione irreparabile del contenuto, a causa di un loro inadeguato utilizzo, è molto sottile.<sup>10</sup>

Infatti, poiché la procedura di restauro non è un ammodernamento dell'opera, è sempre importante cercare di mantenere su di essa una opportuna *patina*, ovvero quella serie di caratteristiche qualitative dovute al passaggio del tempo e tutte le peculiarità che sono specifiche del supporto analogico su cui è memorizzata. L'eliminazione di tali particolarità distintive può portare a snaturare la valenza artistica e soprattutto storica del prodotto originale.<sup>11</sup>

I tecnici specializzati nel restauro mirano a portare il film a uno stato più vicino possibile a quello originario utilizzato per la proiezione di allora. Per questo motivo, il restauro diviene un atto di interpretazione poiché individuare univocamente una "versione originale" del film, nella maggior parte dei casi, diventa pressoché impossibile.<sup>12</sup> Diversi approcci adottabili possono dare la priorità all'integrità storica del film, oppure alla preservazione del fascino delle sue imperfezioni o al miglioramento delle sue qualità estetiche.

Questi elementi contribuiscono a rendere il processo di restauro un lavoro di grande complessità, che non può prescindere da una conoscenza profonda della storia e della tecnica del cinema, nonché delle teorie e dei metodi critici che consentono di analizzare e interpretare un'opera filmica.<sup>13</sup>

Infine, la riedizione, come terzo approccio, comporta modifiche e sostituzioni ingenti, anche agendo su alcune caratteristiche non pensate originariamente da chi ha realizzato il film. Tale processo implica spesso alterazioni di formato, la ricolorazione, l'integrazione di materiale non previsto nella versione originale del film.

Si tratta di una serie di procedimenti che non ha il fine prioritario di garantire lo stato dell'opera originale, quanto più di adattarlo alla fruizione moderna. La

---

<sup>10</sup> Corelli, 2019, ix

<sup>11</sup> Venturini, 2006, 118

<sup>12</sup> *ivi*, 124

<sup>13</sup> *ivi*, 39

sua legittimità è oggetto di discussioni, in quanto altera in maniera consistente lo stato dell'opera originale.<sup>14</sup>

Nei prossimi capitoli verrà affrontato nello specifico l'aspetto sonoro del restauro e ne verranno dettagliate le tecnologie e le procedure.

## 1.2 Formati e supporti analogici per il sonoro cinematografico

La storia del sonoro per l'audiovisivo si può dividere sostanzialmente in tre periodi. Il periodo muto, che dura fino ai tardi anni '20 del Novecento, il periodo del suono analogico registrato su pellicola (*sound-on-film period*), che va dai tardi anni '20 ai primi anni '90 e, infine, il periodo del sonoro digitale, che arriva fino ai giorni nostri.<sup>15</sup>

Il sonoro analogico è passato attraverso una grande varietà di formati proprietari, che si possono però suddividere in due famiglie di tecnologie principali: il supporto ottico e quello magnetico.

### 1.2.1 Supporto ottico

Dai tardi anni '30 fino alla metà degli anni '50 il supporto più diffuso per la produzione e la post-produzione del sonoro per il cinema è stato quello ottico, che si presentava in due varianti: a densità variabile e ad area variabile.

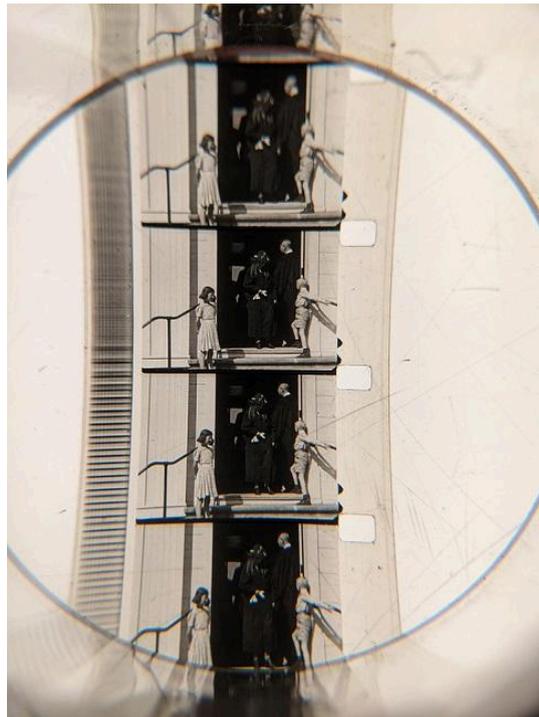
La registrazione sul supporto ottico funziona grazie alla conversione del segnale elettrico generato dal microfono in un segnale luminoso. La pista fotosensibile, situata a bordo della pellicola filmica stessa, viene esposta a questo segnale ottico e su di essa si può, così, imprimere una registrazione permanente del segnale sonoro. In fase di riproduzione, una fonte luminosa attraversa la pellicola processata illuminando una superficie fotoelettrica, in grado di convertire il segnale luminoso in segnale elettrico.

---

<sup>14</sup> Plutino, 2020, 19-21

<sup>15</sup> Enticknap, 2013, 116-117

Nel caso del supporto ottico ad area variabile (*RCA Photophone*) la registrazione avviene grazie a un fascio di luce a intensità costante che illumina una fessura la cui ampiezza varia in base al segnale di input. Questa tecnica di registrazione produce sul supporto una traccia visibile del segnale sonoro corrispondente alla sua forma d'onda, ovvero una rappresentazione in ampiezza del segnale stesso.



**Figura 1.1:** pista ottica monofonica a densità variabile.

Nel caso dell'ottico a densità variabile (*Movietone*), invece, l'apertura della fenditura è costante, mentre è l'intensità della luce a variare proporzionalmente al segnale in input.<sup>16 17</sup>

---

<sup>16</sup> Enticknap, 2013, 118

<sup>17</sup> Uccello, 1982, 329



**Figura 1.2:** pista ottica monofonica ad area variabile.

Il supporto ottico permette la realizzazione di missaggi multitraccia, grazie al posizionamento di più piste di fianco alla foratura della pellicola.

## 1.2.2 Supporto magnetico

Il nastro magnetico è uno dei supporti sonori analogici più usati e più longevi nella storia della registrazione e riproduzione audio. Dai tardi anni '50 ha iniziato a ricoprire un ruolo fondamentale anche nell'industria cinematografica. Questo supporto funziona grazie a una variazione locale di campo magnetico ottenuta intorno a un nastro, che può essere di vari materiali (triacetato, cloruro di polivinile, poliestere)<sup>18</sup>, anche se è più diffusamente costituito in Mylar (nome commerciale del polietilentereftalato).<sup>19</sup> Su questo materiale è presente uno strato magneto-sensibile, spesso in ossido di ferro, depositato sul supporto come una vernice.

I primi nastri, diffusamente larghi un quarto di pollice (6.3 mm), avevano una velocità di scorrimento di 76 cm/s e permettevano una registrazione esclusivamente monofonica. Successivamente si diffusero versioni di nastri con differenti velocità (38 cm/s, 19 cm/s) e larghezze (fino a un massimo di

---

<sup>18</sup> Uccello, 1982, 336

<sup>19</sup> Canazza, 2013, 11

due pollici), adatte anche a registrazioni multitraccia.<sup>20</sup>

La scrittura del segnale avviene grazie all'esposizione del nastro a un campo magnetico generato da piccoli elettromagneti, detti testine magnetiche. Il campo magnetico in questione ha un'intensità variabile e proporzionale alla corrente elettrica derivante dalla trasduzione elettroacustica del segnale audio originale. Tale campo, emesso dalle testine, influenza lo stato magneto-sensibile presente sul film, lasciando su di esso una traccia corrispondente al segnale sonoro (fenomeno della *magnetizzazione residua*).<sup>21</sup> La registrazione può essere effettuata con diverse curve di equalizzazione e codificata con l'utilizzo della tecnica Dolby o dBx per ridurre il rumore di fondo a larga banda. Minore è la larghezza del nastro, maggiore è il rumore introdotto nel segnale. Ciò rende i sistemi di soppressione del rumore fondamentali per supporti da ¼ di pollice e ⅛ di pollice (open-reel semi professionali e cassette). Il nastro magnetico non presenta segni visibili della presenza del segnale, al contrario, ovviamente, del supporto ottico.

## 1.3 Problematiche di degradazione dei supporti analogici

### 1.3.1 Degradazione dei supporti ottici

Poiché la registrazione del suono su supporto ottico consiste in un'immagine fotografica su pellicola del segnale, su di esso si hanno gli stessi problemi di degradazione delle immagini su pellicola.<sup>22</sup>

La principale causa di degradazione dei supporti ottici è da associare a una cattiva conservazione degli stessi. La temperatura ideale è intorno ai 4° C, mentre il valore di umidità relativa è tra il 30 e il 50%. La degradabilità del supporto dipende dal materiale della base (nitrato, acetato o poliestere).

Il problema principale che questi supporti subiscono è quello del

---

<sup>20</sup> Canazza, 2013, 10

<sup>21</sup> ibidem

<sup>22</sup> Enticknap, 2013, 118

*restringimento della pellicola*. Pellicole in avanzato stato di decadimento sono comunemente accompagnate da restringimento fino al 10%, accompagnato, nel caso di acetati, spesso da *delaminazione*, ovvero dal distacco del legante dalla base.



**Figura 1.3:** pellicola in avanzato stato di decadimento.

In altri casi, quando il restringimento è localizzato in alcuni punti, si parla di *distorsione*, che, in mancanza di ripristino delle condizioni di conservazione corrette, può portare alla perdita di elasticità del film e al conseguente abbassamento della resistenza a torsioni e piegamenti (*infragilimento*). Inoltre, le piste ottiche possono essere soggette al cosiddetto *sbiadimento delle tinte*, causato dall'instabilità dei coloranti organici. Il loro invecchiamento produce un effetto di sbiadimento dovuto soprattutto all'esposizione alla luce, ma anche a temperature e umidità non ideali.<sup>23</sup>

---

<sup>23</sup> Plutino, 2020, 48



**Figura 1.4:** un frame soggetto a sbiadimento delle tinte.

### 1.3.2 Degradazione dei supporti magnetici

Il nastro magnetico si può conservare in maniera ottimale in quanto, su di esso, le incisioni rimangono inalterate per lungo tempo rendendo possibili molti passaggi del nastro in lettura senza che questo venga rovinato.<sup>24</sup>

Tuttavia, specialmente quando le condizioni di conservazione o archiviazione non sono ottimali, il nastro magnetico non è esente da degrado. Questo può essere causato da diversi tipi di problematiche.

La prima di queste è detta *degenerazione del legante*, ovvero di quello strato che tiene le particelle magnetiche unite tra di loro e incollate al supporto. Può essere causata dalla conservazione del supporto in luoghi a elevata umidità. La presenza di acqua favorisce il fenomeno chimico dell'idrolisi, che modifica la struttura molecolare del legante, rendendolo più viscoso, creando attrito indesiderato con le testine e provocando il distacco del legante dal supporto.

Il nastro magnetico può essere, inoltre, soggetto alla *perdita del lubrificante*. Questo permette lo scorrimento del nastro a contatto della testina, senza che un attrito eccessivo ne rovini la struttura. Con il passare del tempo, il

---

<sup>24</sup> Uccello, 1982, 336

lubrificante tende a evaporare e a essere rilasciato a contatto della testina magnetica, a seguito di un elevato numero di cicli di lettura e scrittura.

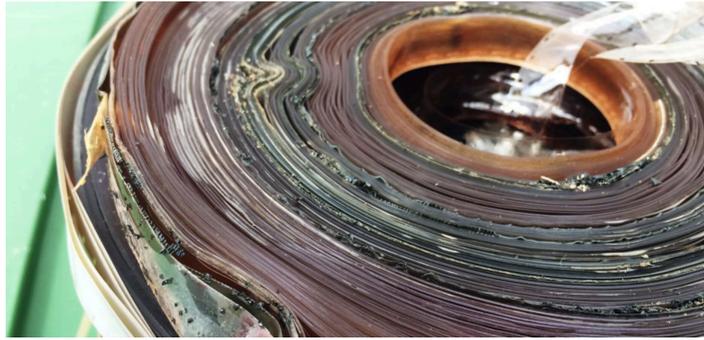
Le particelle magnetiche responsabili del fenomeno della magnetizzazione residua soffrono qualsiasi alterazione delle loro proprietà magnetiche. L'applicazione di campi magnetici esterni può agire smagnetizzando il nastro, causando quindi la perdita del segnale su esso memorizzato.<sup>25</sup> La stabilità della memorizzazione del nastro dipende anche dal tipo di sostanza magneto-sensibile applicata su di esso. L'ossido di ferro è, tra i dipoli utilizzati, quello che gode di maggiore stabilità, mentre altre sostanze come il biossido di cromo o l'MP (*Metal Particulate*) sono meno stabili nel tempo, pur godendo di più ampia banda di frequenze registrabili.

I nastri possono essere anche afflitti da deformazione del substrato, ovvero del film stesso che supporta lo strato magnetico. I nastri con film in acetati (acetato di cellulosa o triacetato di cellulosa) venivano trattati con additivi plastificanti per renderli più flessibili. Con il passare del tempo questi possono cristallizzare o evaporare portando a una graduale riduzione di flessibilità del nastro. Errate condizioni di umidità e di calore possono portare i nastri in acetato ad alterarsi nella loro struttura, portando al loro arricciamento.

Una nota problematica di questi substrati è quella della "sindrome dell'aceto" (*Vinegar Syndrome*), che può essere facilmente rilevata proprio grazie all'odore tipico dell'acido acetico. Il processo di idrolisi, di cui l'umidità è ovviamente un catalizzatore positivo, porta allo sviluppo, appunto, di acido acetico. Non sono note modalità per arrestare questo processo e, una volta che è in stato avanzato, il nastro diventa sempre più fragile e incapace di resistere alle sollecitazioni, portando alla rottura se piegato o teso.

---

<sup>25</sup> "La *coercivity* (o forza coercitiva massima) rappresenta la capacità dei dipoli di resistere alla smagnetizzazione. Indica, in altre parole, la forza del campo magnetico che deve essere applicato ad una particella magnetica affinché questa sia indotta a cambiare la direzione del suo campo magnetico. La smagnetizzazione di un nastro può essere provocata dall'applicazione di un campo magnetico esterno, come quello prodotto dai metal-detector utilizzati ai checkpoint degli aeroporti. Un nastro magnetico i cui dipoli sono caratterizzati da una bassa *coercivity*, risulta maggiormente soggetto alla smagnetizzazione e alla conseguente perdita del segnale in esso contenuto." (Canazza, 2013, 20)



**Figura 1.5:** nastro magnetico gravemente colpito da sindrome acetica.

Esiste infine il cosiddetto *print-through* (effetto spira) che consta nel trasferimento del contenuto magnetico tra spire di nastro adiacenti, quando questo è avvolto nella bobina. Il risultato è che il contenuto nominale del nastro viene sovrapposto con un segnale a livello molto più basso, corrispondente a un segmento successivo di segnale sonoro. Bisogna notare che questo effetto è fortemente influenzato dalle condizioni di conservazione (temperatura, tempo, spessore del nastro, numero di spire, interferenze magnetiche, caratteristiche in frequenza del segnale registrato).<sup>26</sup>

---

<sup>26</sup> "L'effetto spira è principalmente correlato alle seguenti cause: 1. La temperatura: all'aumentare della temperatura aumentano gli effetti del print-through. 2. Il tempo: gli effetti dovuti al print-through si intensificano con il passare del tempo. 3. Lo spessore del nastro: più il nastro è sottile, maggiore sarà l'intensità del print-through. 4. Il numero di riavvolgimenti: ad ogni riavvolgimento del nastro si riduce il print-through. 5. La presenza di campi magnetici: avvicinando il nastro a campi magnetici la cui intensità sia dell'ordine di 10-20 kAs/m (10-20 Gauss), o maggiore, si accelera il processo di print-through. 6. La lunghezza d'onda dei segnali registrati: il print-through si manifesta con maggiore intensità alle basse frequenze." (Canazza, 2013, 22-23)

## 1.4 Caso studio: “Miti e riti dell’India”

In questa tesi di laurea, realizzata con il supporto tecnico e logistico del Museo Nazionale del Cinema di Torino, si tratteranno le tecnologie e le metodologie del restauro sonoro operando su un nastro magnetico contenuto nell’archivio della Cineteca del museo.

Il materiale analogico in questione è parte del Fondo Petruccelli, composto da film di viaggio amatoriali in 16 mm, girati tra gli anni '60 e '70 del Novecento da Ettore Angioletti Sardi e Luciana Petruccelli. Il fondo include, oltre alle pellicole e alle relative piste magnetiche, anche testi che talvolta venivano utilizzati per accompagnare la proiezione, oppure impiegati in seguito per registrare la colonna sonora.

Angioletti Sardi era un ingegnere civile che per motivi di lavoro girò il mondo, assieme a sua moglie, la signora Petruccelli, trovando occasione per realizzare questi film di viaggio.

Per gli scopi della presente tesi, la Cineteca ha messo a disposizione la pista magnetica del film “Miti e riti dell’India”. Le condizioni di conservazione sono discrete: il nastro magnetico è stato trovato in condizioni di lieve sindrome acetica.



**Figura 1.6:** un frame di “Miti e riti dell’India”.

Il nastro è stato inviato alla cineteca di Bologna per essere pulito da impurità, polvere, sporco e altri residui e poi scansionato con Sondor OMA/E. E' stato poi scansionato nella sede della cineteca di Torino tramite Lasergraphics Scanstation.

I due file digitali hanno fornito la base per effettuare un confronto dei due scanner e per la disamina tecnica del presente lavoro di tesi.

## 2. TECNOLOGIE E METODI DI SCANSIONE

Per quanto riguarda il sonoro, il restauro si compone di alcune fasi fondamentali. Prima di tutto è necessario effettuare una ricerca storica, per l'individuazione del miglior supporto fisico reperibile e per fronteggiare i difetti del supporto di provenienza.

Successivamente va analizzato l'oggetto filmico, effettuando una diagnosi delle sue caratteristiche e del suo stato, a cui segue l'acquisizione del segnale tramite scansione del supporto analogico originale. Solo allora, si potrà procedere con la scansione, che deve avvenire alla qualità massima possibile: 24 o 32 bit di risoluzione e almeno 96 kHz di frequenza di campionamento.<sup>27</sup> La scansione del segnale deve avvenire tramite apparecchiature avanzate, di cui possiamo essere in grado di controllare tutti i parametri operativi.

Poiché le apparecchiature moderne offrono parametri meccanici ed elettronici migliori delle macchine con cui i supporti sono probabilmente stati registrati in origine, è necessario in fase di restauro operare una serie di compensazioni. Tali procedure devono neutralizzare alterazioni volontarie di equalizzazione, trasformazioni dovute a tarature non corrette dell'apparecchiatura (soprattutto in caso di registrazioni amatoriali) e anche introduzione di rumori di fondo, distorsioni e transitori indesiderati.

### 2.1 Tipologie di scanner

Per attuare qualsiasi operazione di restauro oggi è necessario effettuare un passaggio del contenuto filmico dal dominio analogico, ovvero quello della pellicola, a quello digitale. Per fare ciò viene utilizzato lo scanner, che permette il "processo di cattura delle immagini sia questo tramite metodi ottici, elettronici o digitali".<sup>28</sup> Si distinguono tre tecnologie di conversione.

---

<sup>27</sup> Corelli et al., 2021, 243

<sup>28</sup> Carlos, 2013, 13

## 2.1.1 Telecine

La prima è il *telecine* in grado di convertire le immagini in segnali video. Questo dispositivo nasce negli anni '50 del Novecento, grazie alla popolarità della televisione e alla conseguente necessità di rendere disponibili le pellicole cinematografiche per il broadcasting. Il telecine non ha capacità di produrre dati digitali, ma solo di effettuare una conversione real-time di tipo elettronico, compatibile con gli standard della televisione interlacciata<sup>29</sup>. Questo tipo di dispositivo produce una conversione qualitativamente insufficiente ai fini archivistici o di ristampa, di conseguenza il suo output non è considerabile come *Digital Intermediate*.<sup>30 31</sup>



**Figura 2.1:** telecine modello Marconi B3410.

---

<sup>29</sup> Enticknap, 2013, 178

<sup>30</sup> "Il Digital Intermediate o DI, in generale, è un processo che descrive la digitalizzazione del supporto cinematografico dalla scansione alla registrazione nuovamente su pellicola. Il processo include anche la manipolazione digitale del colore e di altre caratteristiche dell'immagine per la correzione e il miglioramento e, in alcuni casi, con l'aggiunta di effetti visivi generati al computer. La fase finale è quella in cui i dati manipolati, puliti e composti vengono masterizzati digitalmente per vari formati di distribuzione." (Carlos, 2013, 8)

<sup>31</sup> Carlos, 2013, 14

## 2.1.2 Film scanner

Il secondo tipo di convertitore è il film scanner. Al contrario del telecine, lo scanner di pellicola realizza una rappresentazione digitale delle immagini di un film sotto forma di frame individuali e discreti (ciascun fotogramma contiene metadati e informazioni RGB) e non ha capacità real-time. Permette di realizzare un Digital Intermediate, che corrisponde al primo, fondamentale, passo del restauro cinematografico.<sup>32</sup>

Uno scanner professionale può scansionare qualsiasi tipo di formato ottico (audio e video, da 8 al 70 mm) a risoluzioni variabili (fino a 16K). Non tutti gli scanner di pellicola, però, hanno un sistema di acquisizione del sonoro e bisogna quindi, in quei casi, ricorrere ad altri dispositivi.<sup>33</sup> Esistono quattro tecnologie di film scanner.

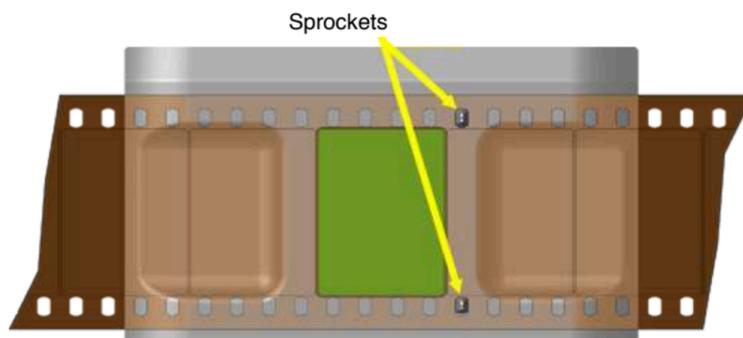
La prima è il tipo *Intermittent pull-down* ed è equipaggiato con un meccanismo intermittente. La pellicola è agganciata tramite le sue perforazioni ad un *film gate*, con dentature regolate in base al formato di pellicola scansionata. Questo tipo di scanner ha una capacità di circa 1 frame per secondo. Si tratta di una tecnologia potenzialmente pericolosa in applicazioni archivistiche o restaurative perché, in caso di pellicole soggette a *shrinking* o ad altre problematiche coinvolgenti le perforazioni, l'aggancio della pellicola può non essere stabile e il meccanismo intermittente può danneggiare irrimediabilmente il supporto.<sup>34</sup>

---

<sup>32</sup> Enticknap, 2013, 178

<sup>33</sup> Carlos, 2013, 22-23

<sup>34</sup> *ivi*, 19-20



**Figura 2.2:** schema di film gate con aggancio regolabile della pellicola.

La seconda tipologia è detta a scansione continua (*continuous-scan*) e consiste nel trasporto della pellicola a velocità costante su un sistema di rotazione privo di dentature, non vengono, quindi, coinvolte le perforazioni. Il supporto viaggia su una serie di elementi meccanici e *capstan*<sup>35</sup>, in grado di assicurare una tensione e frizione sufficiente a permettere lo scorrimento della pellicola, evitando, però, l'usura del film. La stabilizzazione dei fotogrammi viene ottenuta tramite sistema ottico, che utilizza come riferimento le perforazioni.<sup>36</sup> Per ovviare alla problematica di sporco e graffi sul frame scansionato, è stata sviluppata la tecnologia dello scanner *wet-gate*. Questo dispositivo sparge sul supporto una soluzione liquida (percloroetilene) che riempie le microscopiche fenditure sull'emulsione e ha proprietà ottiche simili a quelle del supporto, correggendo le problematiche dovute a queste piccole imperfezioni quando la luce illumina la pellicola. Questo procedimento di correzione delle imperfezioni non ha necessità di interventi digitali. Vengono prodotte immagini più nitide, ma spesso il liquido può provocare residui e bolle sul film. In alcuni scanner la modalità *wet-gate* è opzionale e utilizzabile secondo le necessità.<sup>37</sup>

<sup>35</sup> Il *capstan* è il perno rotante che permette di muovere il nastro attraverso il meccanismo di trasporto. Il nastro rimane premuto tra il capstan e due ruote gommate dette *pinch roller*, che forniscono l'attrito necessario al movimento del supporto.

<sup>36</sup> Carlos, 2013, 20

<sup>37</sup> *ivi*, 20-21



**Figura 2.3:** wet-gate ARRISCAN, i piccoli fori servono a spargere il liquido.

La maggior parte degli scanner è, invece, di tipo *standard gate* (o *dry gate*): non sono coinvolti solventi liquidi o allineatori meccanici e le operazioni di restauro vengono eseguite digitalmente. Il supporto subisce meno sollecitazioni e i macchinari di scansione sono anche più piccoli, seppur meccanicamente sofisticati da operare. <sup>38</sup>



**Fig. 2.4:** scansione dry-gate (a sinistra) e wet-gate (a destra).

Per quanto riguarda l'acquisizione del sonoro ottico, il discorso è lo stesso delle immagini, ma l'elaborazione via software sarà molto differente.

L'audio su nastro magnetico invece ha bisogno di appositi moduli di lettura e

---

<sup>38</sup> Carlos, 2013, 21-22

scrittura (testine magnetiche), di cui molti scanner sono dotati. Spesso gli scanner hanno delle componenti modulari, che permettono di utilizzare sempre lo stesso sistema di trasporto del nastro o della pellicola, modificando però i moduli di acquisizione in base al formato (testine magnetiche o moduli ottici).

### 2.1.3 Datacine

L'ultimo dispositivo è detto *datacine*: si tratta di un ibrido tra un telecine e un normale scanner di pellicola. Infatti, è in grado di produrre un output video adatto al broadcasting, ma anche file dati ad alta risoluzione (come il DPX<sup>39</sup>). Sono estremamente sofisticati e costosi, poiché integrano la massima varietà di tecnologie disponibili e una certa modularità per permettere di adattarsi a perforazioni, dimensioni e supporti differenti. Sono molto utilizzati anche nelle produzioni moderne che scelgono di utilizzare la pellicola come supporto di ripresa e devono procedere, poi, alla postproduzione digitale.

## 2.2 Lasergraphics Scanstation

La Cineteca del Museo Nazionale del Cinema di Torino ha in dotazione uno scanner LaserGraphics Scanstation.

Il suo sistema di *imaging* è in grado di supportare tutti i tipi di formati, dall'8 mm fino a 70 mm IMAX, con risoluzioni di scansione fino a 6.5 K. È basato su un sensore CFA (Color Filter Array) con schema di cattura colore di tipo Bayer.

Ha notevoli capacità di scansione dell'immagine, ma supporta anche diversi formati audio. Per quanto riguarda il sonoro ottico ha un sistema di lettura simultaneo audio-immagine. Può acquisire il suono a diverse velocità di scansione (fino a 60 fps), supportando sia il formato a densità variabile, che quello ad area variabile. A differenza di molti altri scanner, supporta la digitalizzazione dell'audio a partire dai negativi della pellicola.

---

<sup>39</sup> *Digital Moving-Picture Exchange*, un formato in grado di immagazzinare digitalmente le informazioni colore del negativo di ogni singolo fotogramma in una immagine logaritmica non compressa. (Carlos, 2013, 12)



**Figura 2.5:** Lasergraphics Scanstation caricato con pellicola.

Per quanto concerne il sonoro magnetico, supporta formati *commag* (con acquisizione monofonica) e *sepmag* tramite moduli con testine dedicate ai vari formati (8 mm, 16 mm, 17.5 mm, 35 mm a 2 o 3 canali).<sup>40</sup>

Il software proprietario dello Scanstation ha un sistema per l'estrazione automatica del file audio digitale direttamente dalle immagini della pista ottica scansionata, e un sistema di riduzione rumore basato su una sorgente luminosa diffusa, che promette di ridurre rumori e difetti (*click* e *pop*).<sup>41</sup>

## 2.2 Sondor OMA/E

Il laboratorio "L'immagine ritrovata" della Cineteca di Bologna ha in dotazione un sistema di acquisizione riservato al suono basato sul Sondor OMA/E.

Si tratta di un sistema modulare composto da un sistema di trasporto del film e una serie di componenti di lettura per diversi tipi di formati, magnetici e ottici. Il sistema OMA/E base è predisposto per l'utilizzo del nastro magnetico. Può operare a 24 o 25 fps ed è ideale per applicazioni archivistiche. Supporta un

---

<sup>40</sup> Con il termine "*sepmag*" si intende la pista audio magnetica separata rispetto alla pellicola cinematografica, solitamente riprodotta su sistemi di proiezione double-sided. Con "*commag*" si intendono quei supporti filmici per immagine con colonna audio magnetica integrata. (International Federation of Film Archives, [fiafnet.org](http://fiafnet.org))

<sup>41</sup> LaserGraphics, Inc., n.d, [lasergraphics.com/scanstation](http://lasergraphics.com/scanstation)

sistema basato su due bobine e diversi pignoni dinamici in grado di gestire pellicole con problemi di restringimento fino al 4% senza creare danni.<sup>42</sup>



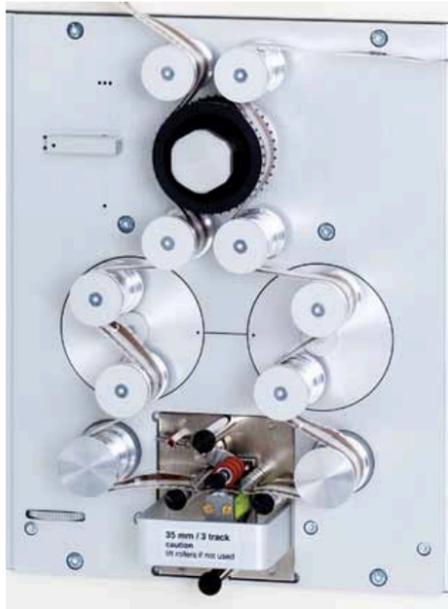
**Figura 2.6:** Sondor OMA/E.

Per quanto riguarda la scansione ottica, OMA-E è compatibile con il sistema Sondor Resonances, composto da un sensore CCD. Resonances si comporta come un datacine, ma permette, tramite il software, di individuare un'area di interesse (ROI), per poter confinare la scansione soltanto intorno alla pista della colonna sonora ottica. Il sistema è dotato di un processore digitale di immagine che, tramite il software proprietario, permette l'individuazione e

---

<sup>42</sup> Sondor & Digital Film Technology GmbH, n.d.

rimozione dei difetti ottici prima che avvenga la conversione a segnale audio.<sup>43 44</sup>



**Figura 2.7:** sistema di trasporto dell'OMA/E. **Figura 2.8:** acquisizione ottica Resonances.

---

<sup>43</sup> Sondor & Digital Film Technology GmbH, n.d.

<sup>44</sup> Sondor Versa & Resonances, n.d., dft-film.com

## 3. ANALISI DEI SEGNALI DIGITALIZZATI

Si riconoscono diverse tipologie di degradazioni del segnale audio, riconducibili, però, a due categorie: disturbi locali e disturbi globali.

### 3.1 Disturbi locali

I disturbi di tipo locale possono essere definiti come discontinuità nella forma d'onda che coinvolgono solo alcuni campioni del segnale. Tra questi si possono distinguere vari tipi di difetti, come *clicks*, *crackles*, *scratches* e *clipping*.<sup>45</sup>

Nello specifico, "si definisce *click* un generico tipo di degradazione localizzata di durata limitata e presente in posizione casuale lungo la forma d'onda".<sup>46</sup>

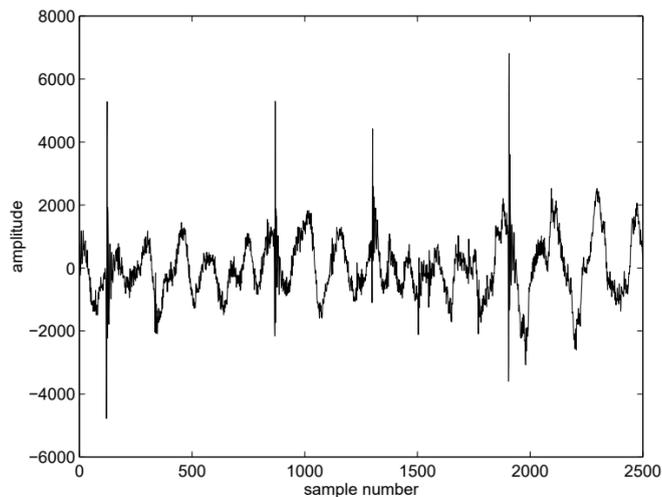
Si tratta di disturbi impulsivi a largo spettro, molto diffusi in molti segnali registrati su supporti analogici. Sono solitamente meno frequenti sui nastri magnetici che nei supporti ottici.

Possono essere causati da molteplici problematiche: sporco depositato, irregolarità superficiali o piccoli graffi dovuti all'usura. Talvolta possono essere dovuti a disturbi di natura elettrica o elettromagnetica, interferenze e falsi contatti delle apparecchiature. Ma possono avere un'origine acustica, per esempio sono tipici degli schiocchi di saliva del parlatore.

---

<sup>45</sup> Godsill & Rayner, 1998, 6-7

<sup>46</sup> Canazza, 2013, 62



**Figura 3.1:** forma d'onda di un segnale audio degradato da alcuni click.

Tra i vari tipi di click, troviamo i *pop*, impulsivi e caratterizzati da componenti in bassa e medio-bassa frequenza (quindi timbricamente più scuri).

Gli *scratches* sono disturbi temporalmente più estesi e in alta e medio-alta frequenza.

I *crackles* sono gruppi di click e pop, densamente accumulati uno vicino all'altro sull'asse temporale. Sono tipici di sporcizia e polvere depositate sul nastro magnetico o sulle testine.<sup>47</sup>

È importante notare che questo tipo di degradazione non è tipico soltanto dei supporti analogici, ma si presenta di frequente anche nel dominio digitale, spesso a causa di errori di scrittura o di temporizzazione.<sup>48</sup>

Nei segnali audio analogici troviamo spesso disturbi di *clipping*, caratterizzati dall'introduzione di distorsione armonica all'interno del segnale audio. Questo tipo di problematica è causata dal superamento del range dinamico dell'amplificatore, con conseguente saturazione del segnale. Il risultato è un suono "*grattato*", soprattutto nelle zone delle forma d'onda ad elevata ampiezza.<sup>49</sup>

---

<sup>47</sup> Corelli, 2019, 24

<sup>48</sup> Canazza, 2013, 62

<sup>49</sup> Corelli, 2019, 25

## 3.2 Disturbi globali

I disturbi globali colpiscono tutti i campioni di un dato segmento del segnale. Nello specifico, tra questi, troviamo il rumore a banda larga, che rappresenta il più diffuso tipo di degradazione, presente nei segnali audio, che il restauratore si ritrova ad affrontare. Questo tipo di rumore viene percepito dall'ascoltatore come un "hiss", termine onomatopeico, come i nomi di molti altri disturbi, che indica un rumore bianco. Viene spesso prodotto dal sistema di registrazione, dal degrado del supporto o dal rumore ambientale. Avendo componenti in frequenza in ogni parte dello spettro, non è in alcun modo eliminabile con filtri equalizzatori. Si può rimuovere con diversi tipi di sistemi, di cui parleremo in seguito.<sup>50</sup>

Il rumore bianco è da distinguere dal "hum" (ovvero un ronzio, anche detto "ground loop"), un disturbo globale a banda stretta, in bassa frequenza, generato dalle interferenze dell'impianto elettrico di rete con cui le apparecchiature elettroacustiche vengono alimentate.

Esistono altri tipi di ronzii, come il "buzz" che ha la stessa natura dell'hum, ma è localizzato nello spettro sulle alte frequenze. Questi disturbi possono essere affrontati anche tramite semplici operazioni di equalizzazione.<sup>51</sup>

Per quanto riguarda soprattutto i dischi analogici, spesso si parla di *low frequency noise pulse*. Si tratta di disturbi globali a bassissima frequenza, tipicamente causati da graffi e rotture sulla superficie dei dischi. Percettivamente è simile ad un pop, ma al transiente si sussegue una lunga pulsazione a bassa frequenza molto difficile da rimuovere, perchè si amalgama al segnale originale.<sup>52 53</sup>

---

<sup>50</sup> Canazza, 2013, 72

<sup>51</sup> Corelli, 2019, 26-28

<sup>52</sup> Godsill & Rayner, 1998, 154

<sup>53</sup> Canazza, 2013, 70-71

### 3.3 Degradazioni presenti in “Miti e riti dell’India”

La pista sonora di “Miti e riti dell’India” è costituita da un montaggio di registrazioni vocali in voice-over, unite e alternate a tracce musicali con atmosfere tipiche della cultura indiana.

Come già accennato nel paragrafo 1.4, il nastro è stato pulito dalle impurità, dalla polvere accumulata e dai residui depositatisi nel tempo. Nelle fasi preliminari di questo progetto di tesi, a fini di test, è stata effettuata anche una scansione prima della pulizia del supporto, in cui il risultato è estremamente più disturbato. Questo sottolinea l’importanza delle fasi analogiche del restauro.<sup>54</sup>

A un’analisi preliminare la traccia magnetica sembra in uno stato di discreta conservazione. Al primo ascolto si notano diversi tipi di problematiche, ma non tutti sono dovuti al degrado del supporto, bensì alle caratteristiche dello stesso e alle apparecchiature di registrazione.<sup>55</sup>

È doveroso premettere che tutte le degradazioni discusse in questo paragrafo sono state rilevate in entrambe le scansioni effettuate. Le differenze osservate tra gli scanner saranno oggetto dei prossimi paragrafi.

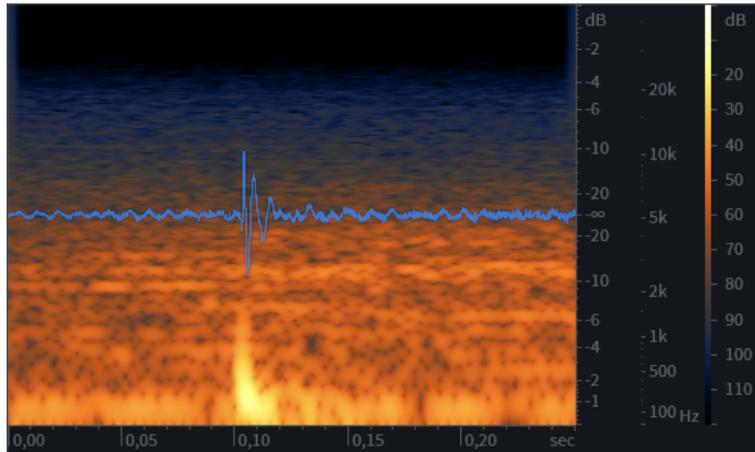
#### 3.3.1 Click, pop, crackle e clipping

La pista è popolata di diversi click e pop che sono udibili, ma non rovinano, fortunatamente, l’intelligibilità della traccia, sia durante il parlato, sia durante i segmenti musicali. Esistono brevi segmenti con crackle appena udibili, dove i click sono, sì, addensati temporalmente, ma a ridotta intensità rispetto ad altri disturbi più impulsivi, presenti altrove.

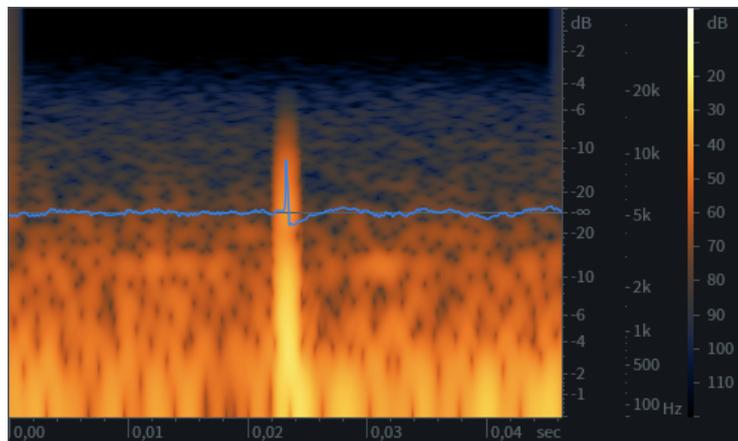
---

<sup>54</sup> Il nastro è stato lavato nel laboratorio della Cineteca di Bologna.

<sup>55</sup> Si ricorda che il film in questione è una produzione amatoriale, realizzata con mezzi limitati.



**Figura 3.2:** spettrogramma e forma d'onda di un pop. Scansione: MNC.

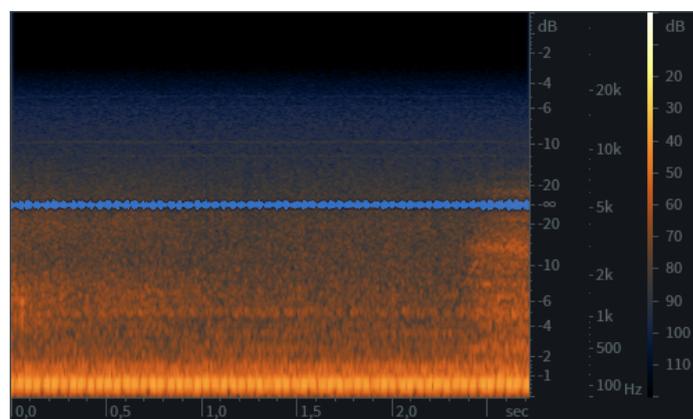


**Figura 3.3:** spettrogramma e forma d'onda di un click. Scansione: MNC.

Inoltre, la traccia è in molti punti deteriorata da clipping. La voce molto spesso raggiunge il livello di saturazione, ciò è udibile, ma non tanto da rovinare il segnale in maniera preoccupante. Questo disturbo pare associabile a un volume di registrazione troppo alto ed è, in molti tratti, assimilabile a un'estetica sonora tipica dei nastri magnetici.

### 3.3.2 Hum

I passaggi con il parlato sono accompagnati da un *hum* chiaramente distinguibile, che cessa solamente nei segmenti della traccia che sono esclusivamente musicali. Questo è probabilmente dovuto all'apparecchiatura di registrazione, di cui non è nota alcuna caratteristica, né per quanto riguarda il trasduttore microfonico, né per quanto concerne l'apparato di acquisizione. Così come non conosciamo nel dettaglio il metodo di realizzazione di queste tracce.

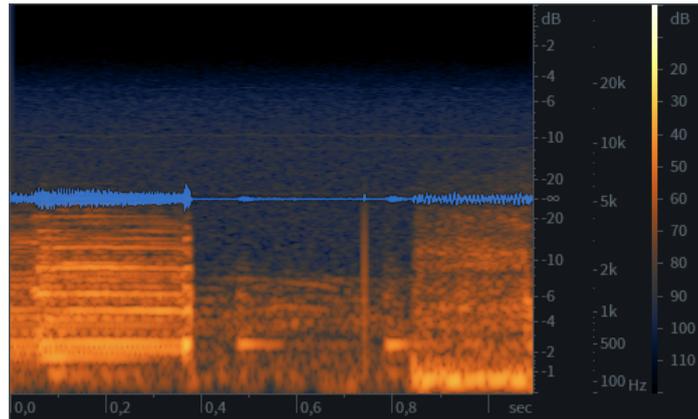


**Figura 3.4:** spettrogramma e forma d'onda di un *hum* isolato. Scansione: MNC.

### 3.3.3 Disturbi isolati

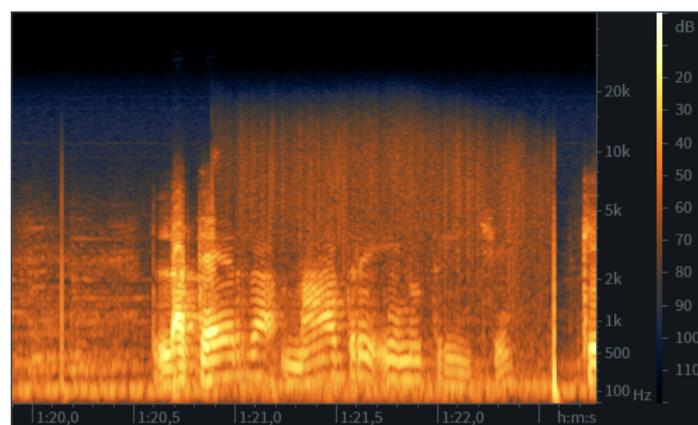
In alcuni tratti della pista audio sono presenti dei disturbi isolati. Nello specifico uno scratch, probabilmente dovuto ad un graffio fisicamente presente sul nastro magnetico, produce un effetto sonoro simile a quello della puntina di un giradischi che esce dal solco.

Un altro disturbo particolare è caratterizzato da una ovattatura transitoria della musica. È difficile dire con certezza quale sia la causa, osservando lo spettrogramma si nota come ci sia un effetto passa-basso transitorio, della durata di circa 40 ms, con una rispettiva riduzione di volume.



**Figura 3.5:** spettrogramma relativo a ovattatura transitoria. Scansione: MNC.

La più macroscopica delle problematiche osservate, però, riguarda un segmento della traccia di circa 2 secondi. Il disturbo è delimitato da due click chiaramente udibili. Dopo il primo click, si sente una intensificazione del rumore di fondo, con un relativo effetto pitch-down progressivo sul contenuto vocale della traccia. Un secondo click delimita la fine della degradazione e il ripristino del rumore di fondo ai livelli osservati nel resto della traccia. La problematica è chiaramente descritta dalla figura 3.6. Si notano i due click, all'inizio e alla fine, con la loro caratteristica impulsiva a largo spettro. Si osserva chiaramente l'aumento di intensità del rumore a larga banda e l'andamento discendente in frequenza delle componenti a maggiore intensità.



**Figura 3.6:** spettrogramma dell'effetto pitch-down transitorio. Scansione: MNC.

### 3.4 Confronto delle due scansioni

Sulla carta, il Sondor è uno dei sistemi di acquisizione del suono analogico più efficaci e performanti in circolazione, ciò è testimoniato dal fatto che si tratta di una macchina esclusivamente pensata e ottimizzata per applicazioni audio, oltre che dalla sua diffusione nelle maggiori cineteche mondiali.<sup>56</sup>

Paragonando le schede tecniche, però, il Lasergraphics non ha nulla da invidiare, almeno per quanto riguarda le specifiche dichiarate. Come già detto, il Lasergraphics Scanstation è pensato, prima di tutto, per la scansione dell'immagine e rappresenta una soluzione all-in-one.

Il senso del confronto fra queste due macchine nasce dal tentativo di comprendere le differenze tra i due scanner e proporre un'analisi dettagliata delle loro possibilità, per fornire indicazioni oggettive finalizzate a comprendere come arginare il deficit che può nascere in una digitalizzazione effettuata con un macchinario piuttosto che l'altro.

I due macchinari, dal punto di vista di questa tesi, appaiono approssimabili a due black-box: abbiamo a disposizione sia i manuali utente, sia le schede tecniche, ma questo materiale offre più che altro una guida d'utilizzo. Abbiamo informazioni complete e dettagliate a proposito dell'input e dell'output, ma è piuttosto complicato conoscere con precisione cosa avviene all'interno dello scanner durante il processo di digitalizzazione, perché le componenti non sono note.

Per studiare e comprendere il divario tra gli output delle due macchine, a partire dallo stesso materiale analogico di partenza, si cercherà di tracciare un profilo delle differenze tra i due file digitalizzati e offrire un punto di partenza per il tecnico restauratore che, guidato dalla propria sensibilità personale ed esperienza, deciderà come muoversi per ottenere il risultato migliore possibile dai file a disposizione.

Conoscere nello specifico queste differenze potrà infatti servire in fase di

---

<sup>56</sup> Hanno in dotazione uno scanner Sondor: il NARA di Washington D.C., il BFI di Londra, l'INA di Parigi, il Bundesarchiv di Berlino e l'Immagine Ritrovata di Bologna. (Sondor & Digital Film Technology GmbH, n.d.)

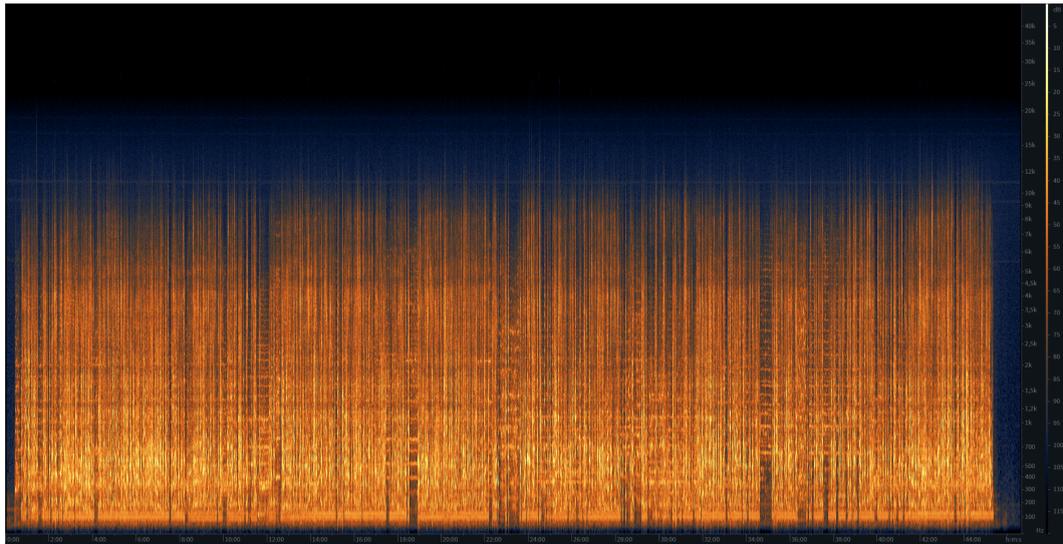
postproduzione per sapere, partendo dalle caratteristiche della scansione a propria disposizione, come compensare le imperfezioni del macchinario tramite gli strumenti software, allo scopo di avere un controllo totale e mirato in ogni fase del restauro.

Per trarre delle conclusioni utili a livello operativo, serve confrontare le due scansioni, prima su un livello macroscopico e poi dettagliare le differenze esistenti negli specifici disturbi individuati.

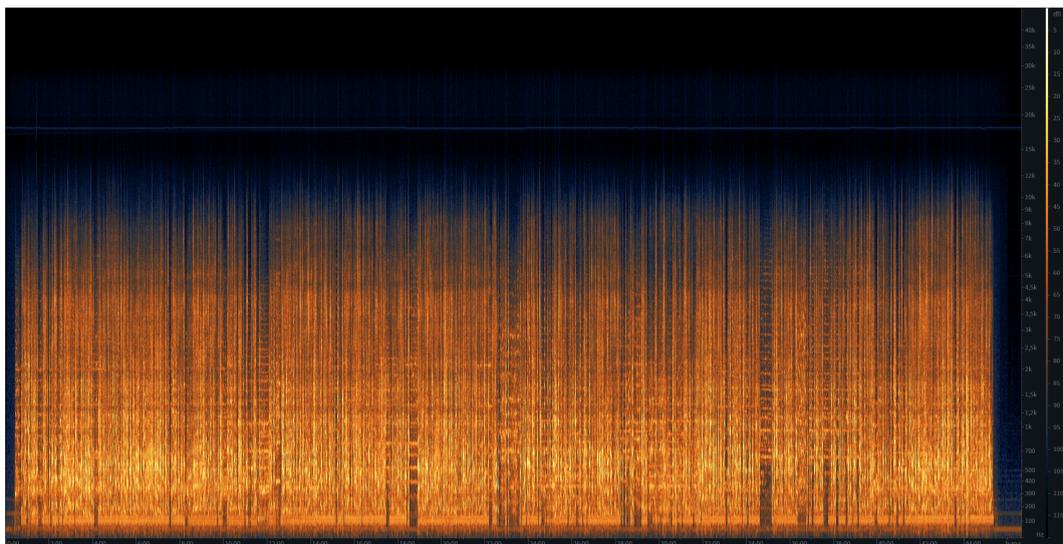
Affinché i confronti siano validi e leggibili, i due segnali devono essere normalizzati. Si possono operare vari tipi di normalizzazione, in base alle necessità, basandosi sul valore di picco (true peak), ma anche sul valore di intensità RMS e su valori tarati percettivamente, come i LUFS (Loudness Unit relative to Full Scale). Una volta normalizzati i segnali si possono fare delle osservazioni di tipo spettrale, per comprendere le macro-differenze.

### **3.4.1 Confronto macroscopico**

Prima di tutto, l'elemento più evidente che emerge dallo spettrogramma è la differenza in rumore di fondo: la scansione fatta con il Sondor appare molto meno popolata da disturbi a larga banda rispetto a quella del Lasergraphics, ciò è molto agilmente verificabile anche con un semplice ascolto. È probabile che la scansione effettuata a Bologna con il Sondor sia stata fatta con un sistema di riduzione del rumore di fondo applicato di default. Bisogna notare che questo tipo di sistema è presente anche sul Lasergraphics, ma è stato scelto di tenerla disattivata per evitare che il segnale venisse modificato da sistemi su cui non abbiamo il perfetto controllo. Il parametro di denoising della scansione effettuata con Sondor, sempre che sia stato applicato, non è noto.



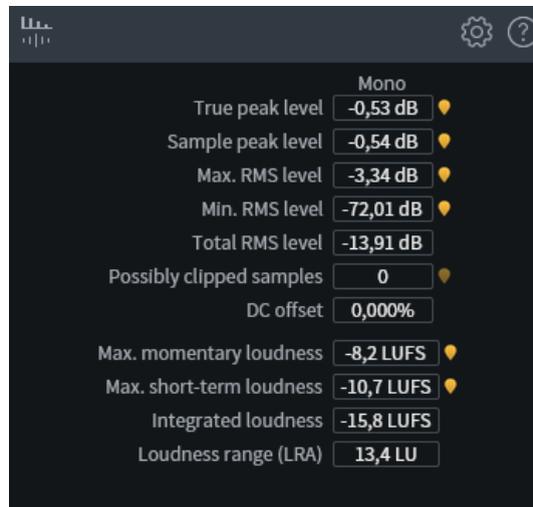
**Figura 3.7:** spettrogramma della scansione effettuata con Lasergraphics (MNC).



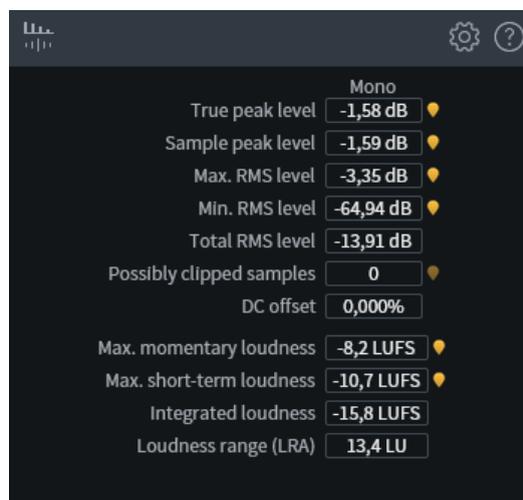
**Figura 3.8:** spettrogramma scansione effettuata con Sondor (Cineteca di Bologna).

Si riportano alcune statistiche di base dei due segnali normalizzati, compensando i valori RMS tra le due tracce con un boost sulla scansione Sondor di 7.3 dB (figure 3.9 e 3.10). Questa procedura ha come scopo un

pareggiamento del valore di loudness (misurato in LUFS): pur essendoci differenze nei valori di picco, il volume percepito è esattamente lo stesso.<sup>57</sup>



**Figura 3.9:** statistiche scansione Lasergraphics.



**Figura 3.10:** statistiche scansione Sondor.

---

<sup>57</sup> La loudness è una misura psicoacustica del volume, basata sulle caratteristiche timbriche del segnale e sul modo con cui il nostro cervello le interpreta. Particolarmente interessanti sono i valori di LUFS-I, che sono mediati lungo tutta la durata del segnale. (Stewart & iZotope, Inc., n.d.)

### 3.4.2 Confronto di segmenti specifici

Per tentare di caratterizzare i due scanner a partire dalla scansione della stessa pista magnetica, si procederà ad analizzare svariate porzioni del segnale, a partire prima da alcuni segmenti generici e poi da specifiche sezioni, contenenti i più significativi disturbi e degradazioni.

Per osservare queste differenze sono stati tracciati i grafici dello spettro di ciascuna porzione, nelle sue due versioni. Sono, inoltre, stati scelti alcuni parametri per quantificare le differenze tra una scansione e l'altra.

Per primo il *centroide spettrale*, ovvero il centro di massa dello spettro. Viene espresso con un valore frequenziale, che rappresenta il baricentro. Più alto il valore del baricentro, più il segnale è, mediamente, brillante, mentre un valore basso indica un segnale timbricamente più scuro.

Altro parametro che è stato considerato è la *spectral flatness*, indicatore di rumorosità del segnale. Il valore è compreso tra 0 e 1, dove 0 corrisponde a un tono puro, mentre un valore che si sposta verso 1 corrisponde a un rumore gaussiano bianco.

Inoltre, si considera il parametro di distanza spettrale logaritmica LSD (*Log-Spectral Distance*), che corrisponde, appunto, ad una quantificazione della differenza tra i due spettri.<sup>58</sup>

Per avere una visualizzazione ottimale degli spettri sovrapposti, in ciascun grafico il valore della FFT è stato normalizzato rispetto all'energia media di ogni segnale preso in considerazione, in modo da rendere i comportamenti spettrali perfettamente sovrapposti, permettendo un confronto più agile.

Sull'asse verticale, il riferimento per la scala in dB è stato preso rispetto a una sinusoide pura a 440 Hz, esportata da ProTools a 0 dBFS. Il valore in dB, quindi, non corrisponde al volume di quel dato segmento nella traccia originale, ma è ottimizzato per agevolare il confronto spettrale.

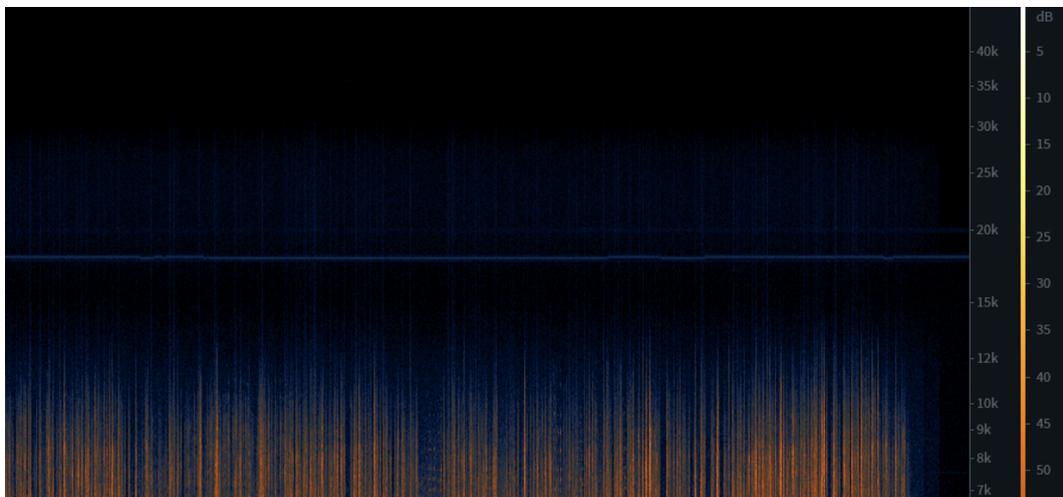
---

<sup>58</sup> Rabiner, 1993, 158-162

I segmenti selezionati per i confronti sono stati estrapolati dalle due tracce. La traccia di Bologna ha ricevuto una compensazione di +7.3 dB<sup>59</sup>, sufficiente a pareggiare i valori RMS e ottenere uguali livelli di loudness tra le due tracce.

### 3.4.2.1 Rumore di fondo

Sebbene il rumore di fondo della traccia scansionata con il Lasergraphics sia generalmente superiore, bisogna notare che la scansione del Sondor presenta un picco nel rumore di fondo a 18 kHz, che è ben visibile sia nello spettrogramma<sup>60</sup>, sia in tutti i grafici che saranno proposti in seguito.<sup>61</sup>



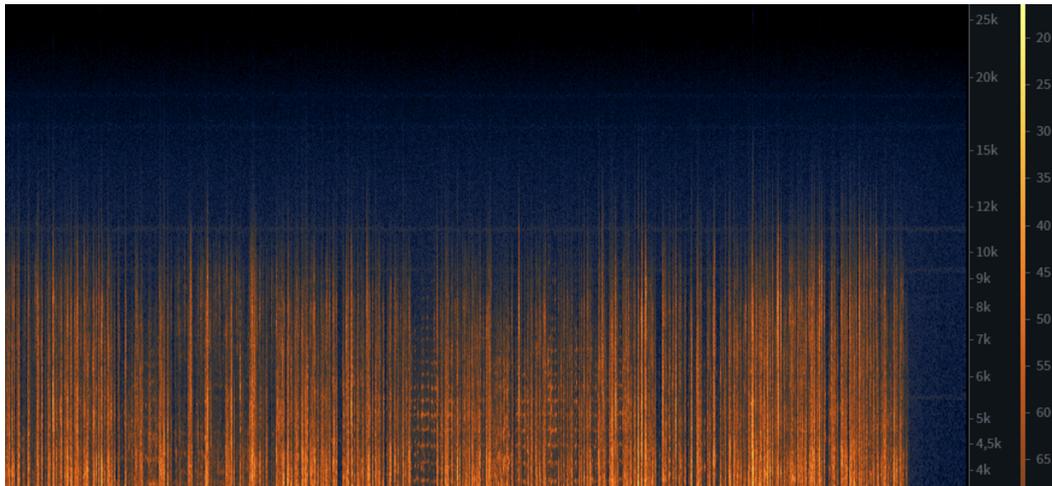
**Figura 3.11:** dettaglio dello spettrogramma con picco intorno ai 18 kHz. (Sondor)

In maniera molto simile, la scansione del Lasergraphics presenta dei picchi nel rumore di fondo, presenti a 5.5 kHz, 9.5 kHz, 11 kHz, 17 kHz. Questi sono chiaramente visibili dai grafici, ma spiccano meno, perché sono mascherati dal rumore a larga banda presente e perché meno intensi.

<sup>59</sup> Vedere paragrafo 3.4.1

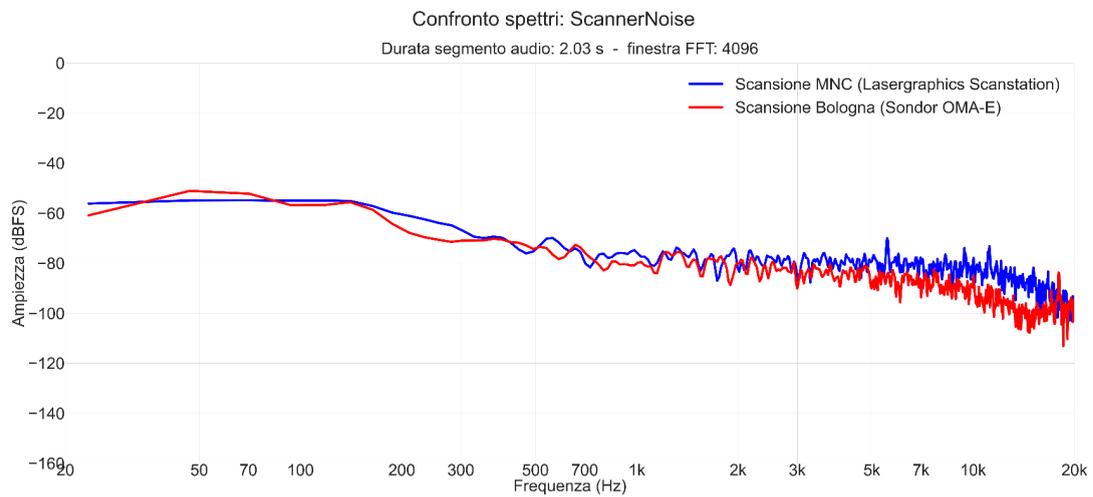
<sup>60</sup> Per vederlo, bisogna osservare la linea orizzontale biancastra che si nota in alto, sopra a tutti i picchi. Si trova appunto a 18 kHz ed ha una intensità rilevante rispetto al rumore di fondo della traccia scansionata con Lasergraphics.

<sup>61</sup> In questa zona dello spettro siamo vicini al limite di udibilità dell'orecchio umano e per molti udire un tono a 18 kHz è pressoché impossibile.



**Figura 3.12:** dettagli di picchi nel rumore di fondo della scansione Lasergraphics.

Si riporta, anche il grafico del confronto tra gli spettri dei rumori di fondo dei due scanner, dove si possono osservare nitidamente i picchi menzionati sopra, su entrambe le tracce.



**Figura 3.13:** confronto spettrale di un segmento di rumore di fondo.

Il rumore di fondo della scansione Sondor ha un comportamento maggiormente tonale, riportando un valore di spectral flatness più alto rispetto a quello del Lasergraphics.

Scanner Noise	Lasergraphics	Sondor
Spectral Centroid [Hz]	5903.37	6275.25
Spectral Flatness	0.000464765	0.002607362

**Figura 3.14:** descrittori utili del primo segmento di rumore di fondo.

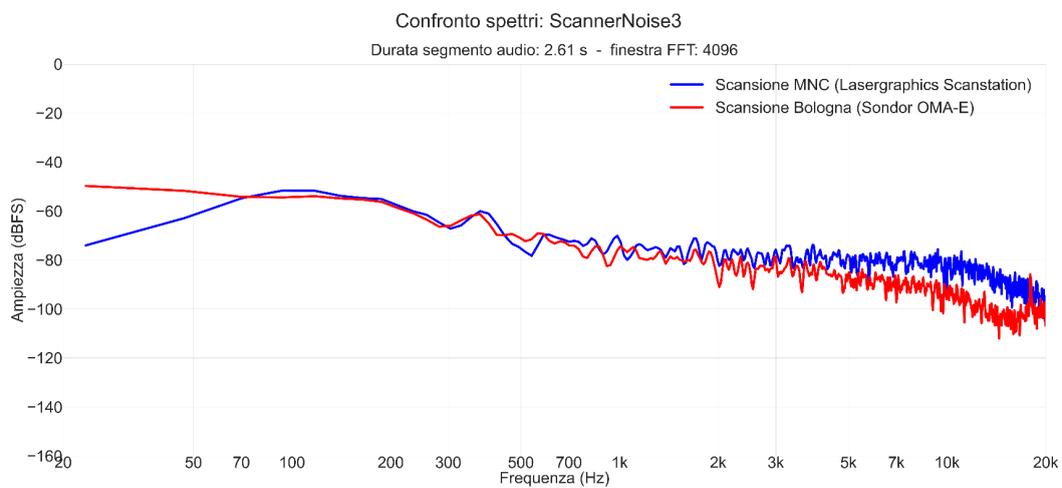
Questo viene anche sottolineato dall'osservazione della LSD, calcolata per bande d'ottava, che tra i due segmenti di rumore di fondo è maggiore in bassissima e in alta frequenza.

<b>LSD Globale [dB]</b>	1.34
LSD 20-40 Hz	0.10
LSD 40-80 Hz	1.57
LSD 80-160 Hz	0.92
LSD 160-315 Hz	1.05
LSD 315-630 Hz	0.63
LSD 630-1250 Hz	0.75
LSD 1250-2500 Hz	0.71
LSD 2500-5000 Hz	0.82
LSD 5000-10000 Hz	1.03
LSD 10000-20000 Hz	1.16

**Figura 3.15:** distanza spettrale globale e per bande d'ottava del rumore di fondo.

Prendendo in considerazione un altro segmento di rumore di fondo, situato in un altro punto della traccia e di durata simile (intorno ai 2 secondi), osserviamo uno schema paragonabile.

Troviamo la stessa tendenza sia nella spectral flatness che nella LSD (figura 3.17). In bassissima frequenza si osservano 4.5 dB di differenza, nonostante il valore globale sia minore.<sup>62</sup>



**Figura 3.16:** confronto spettrale di un secondo segmento di rumore di fondo.

---

<sup>62</sup> Si noti che nella zona 20-40 Hz il nastro magnetico non ha capacità di registrare informazioni utili del segnale, per sua caratteristica intrinseca. In fase di restauro si sceglierà probabilmente di usare un filtro passa-banda per eliminare informazioni inutili in bassissima e altissima frequenza.

<b>LSD Globale [dB]</b>	0.90
LSD 20-40 Hz	4.50
LSD 40-80 Hz	1.01
LSD 80-160 Hz	0.43
LSD 160-315 Hz	0.48
LSD 315-630 Hz	0.79
LSD 630-1250 Hz	0.67
LSD 1250-2500 Hz	0.70
LSD 2500-5000 Hz	0.90
LSD 5000-10000 Hz	1.17
LSD 10000-20000 Hz	1.31

**Figura 3.17:** distanze spettrali di un secondo segmento di rumore di fondo.

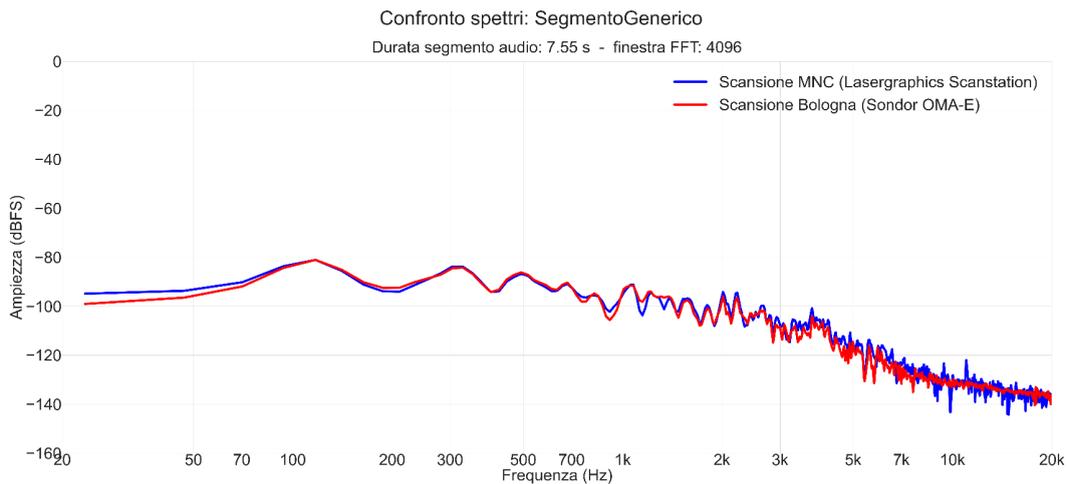
I valori di spectral flatness corrispondono a quelli dell'altro segmento, anche se in questo tratto la tendenza per quanto riguarda il centroide spettrale si inverte, avendo un valore più alto sul rumore del Lasergraphics.

<b>Scanner Noise 2</b>	<b>Lasergraphics</b>	<b>Sondor</b>
Spectral Centroid [Hz]	5675.26	5170.89
Spectral Flatness	0.000471910	0.001945591

**Figura 3.18:** descrittori utili, secondo segmento di rumore di fondo.

### 3.4.2.2 Segmenti generici

Si procede a commentare i dati estratti da segmenti generici, il cui contenuto è misto (musica e voce). Queste porzioni di segnale non riportano nessuna caratteristica specifica o degradazione, ma riassumono bene il contenuto medio della traccia.



**Figura 3.19:** confronto spettrale di una porzione mista, voce e musica.

Il segmento dura poco meno di 8 secondi e le differenze sono minime, con spettri praticamente perfettamente allineati. Come riscontrato in molti altri segmenti specifici, la scansione del Lasergraphics ha un centroide spettrale più alto, che ne caratterizza una maggiore brillantezza. Si può notare che il picco a 18 kHz nella scansione di Bologna qui è invisibile, perché la sua intensità è trascurabile rispetto al volume del contenuto del segmento audio.

<b>SegmentoGenerico</b>	<b>Torino</b>	<b>Bologna</b>
Spectral Centroid	1553.02	1401.86
Spectral Flatness	0.000000634	0.000009216

**Figura 3.20:** descrittori utili, porzione mista.

La distanza spettrale globale è minima, con un valore intorno a 0.24 dB, che va contestualizzato prendendo in considerazione le LSD per bande d'ottava.

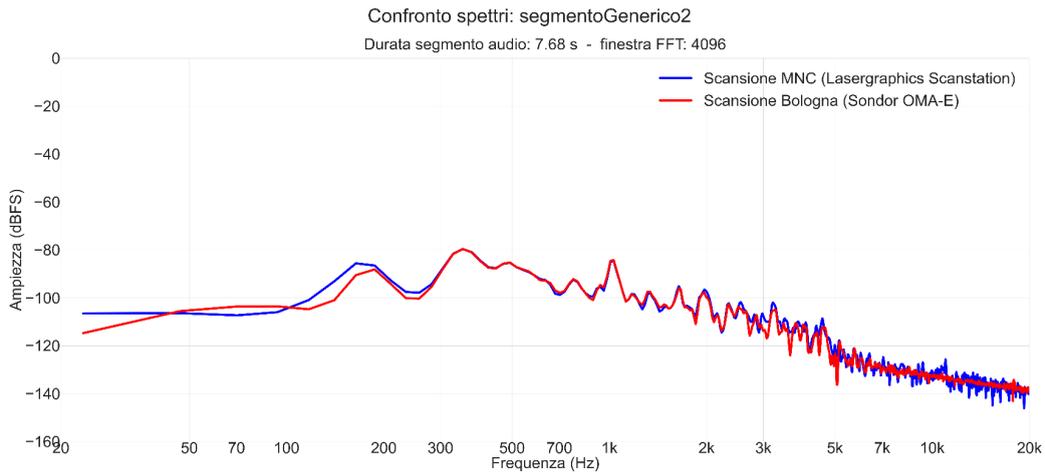
<b>LSD Globale [dB]</b>	0.24
LSD 20-40 Hz	0.57
LSD 40-80 Hz	0.23
LSD 80-160 Hz	0.05
LSD 160-315 Hz	0.14
LSD 315-630 Hz	0.09
LSD 630-1250 Hz	0.29
LSD 1250-2500 Hz	0.30
LSD 2500-5000 Hz	0.41
LSD 5000-10000 Hz	0.47
LSD 10000-20000 Hz	0.29

**Figura 3.21:** distanze spettrali, segmento misto.

Nella maggior parte delle bande analizzate c'è un allineamento notevole fra i due spettri, che si differenziano di più nel range 20-40 Hz (in maniera nettamente visibile anche nel grafico dello spettro) e tra i 2.5 e i 10 kHz, con valori comunque molto bassi.

Si può, dunque, procedere a osservare altri due segmenti generici, anche in questo caso porzioni miste di segnale composte da musica e voce, di durata di circa 8 secondi.

Nel primo dei due si nota un comportamento pressoché identico, con un allineamento considerevole tra gli spettri delle due scansioni.



**Figura 3.22:** confronto spettrale di un altro segmento misto.

Considerando le distanze spettrali si nota che, anche in questo caso, i valori sono comparabili, ma le differenze nelle singole bande d'ottava sono lievemente diverse. Infatti in questo caso la distanza maggiore si ha nella banda 80-160 Hz, quindi in una zona dello spettro più alta rispetto al segmento audio precedente. La distanza media è allineata e comunque minima.

<b>LSD Globale [dB]</b>	0.22
LSD 20-40 Hz	0.07
LSD 40-80 Hz	0.33
LSD 80-160 Hz	0.97
LSD 160-315 Hz	0.18

LSD 315-630 Hz	0.05
LSD 630-1250 Hz	0.11
LSD 1250-2500 Hz	0.22
LSD 2500-5000 Hz	0.46
LSD 5000-10000 Hz	0.32
LSD 10000-20000 Hz	0.27

**Figura 3.23:** distanze spettrali, secondo segmento misto.

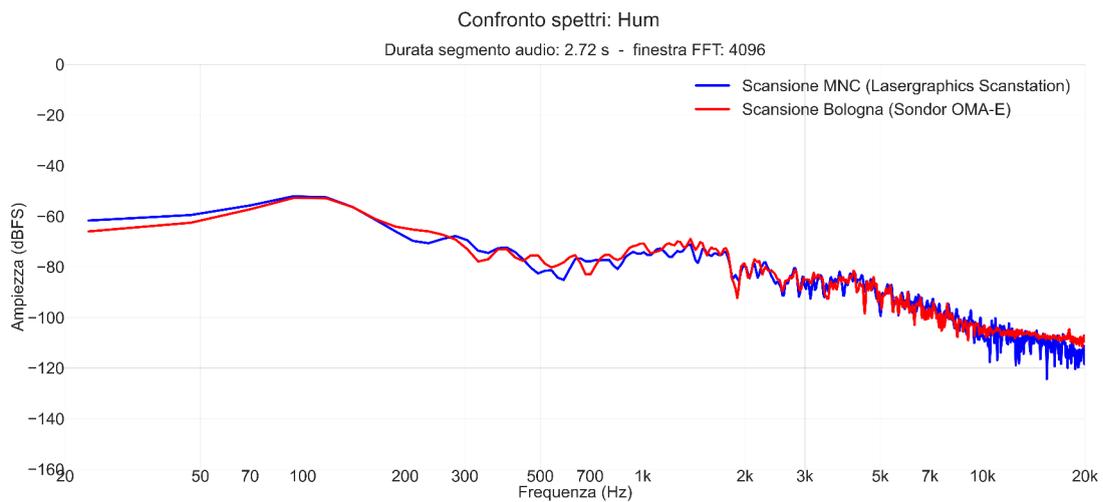
Per quanto riguarda i valori di spectral flatness il valore, come nel caso del segmento precedente, è minore nella scansione del Lasergraphics. Il Sondor, in perfetta tendenza con il resto dei valori, riporta un centroide spettrale minore.

<b>Segmento generico 2</b>	<b>Lasergraphics</b>	<b>Sondor</b>
Spectral Centroid	1385.76	1271.83
Spectral Flatness	0.000004710	0.000009720

**Figura 3.24:** descrittori utili, secondo segmento a contenuto misto.

### 3.4.2.3 Ronzio sui segmenti parlati

Come già discusso nel paragrafo 3.3.2 i voiceover di "Miti e riti dell'India" sono caratterizzati da un consistente hum. Per effettuare il confronto è stato isolato un frammento di poco meno di tre secondi, in cui non c'è nè contenuto vocale nè musicale, ma l'hum è chiaramente udibile.



**Figura 3.25:** confronto spettrale del ronzio isolato dal contenuto vocale.

Su questo tipo di elemento le differenze tra le due scansioni sono davvero minime, tanto che la LSD globale è di 0.46 dB. Osservando la LSD per bande d'ottava si notano valori massimi di differenza intorno ai 0.7 dB.

<b>LSD Globale [dB]</b>	0.46
LSD 20-40 Hz	0.70
LSD 40-80 Hz	0.34
LSD 80-160 Hz	0.15
LSD 160-315 Hz	0.52
LSD 315-630 Hz	0.76
LSD 630-1250 Hz	0.51
LSD 1250-2500 Hz	0.69
LSD 2500-5000 Hz	0.55
LSD 5000-10000 Hz	0.58
LSD 10000-20000 Hz	0.54

**Figura 3.26:** tabella delle distanze spettrali tra i ronzii scansionati.

Si osserva un comportamento comparabile anche nella spectral flatness, mentre il centroide spettrale sottolinea che la scansione del Lasergraphics è timbricamente più brillante, con uno spostamento del centro di massa dello spettro verso destra.

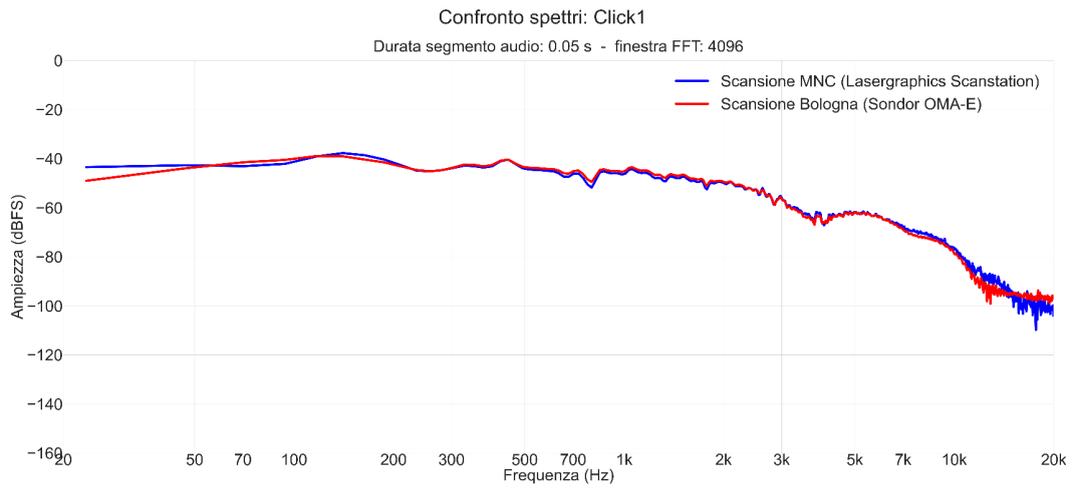
<b>Hum</b>	<b>Lasergraphics</b>	<b>Sondor</b>
Spectral Centroid	1938.44	1744.92
Spectral Flatness	0.000023755	0.000053142

**Figura 3.27:** tabella di descrittori utili per il ronzio.

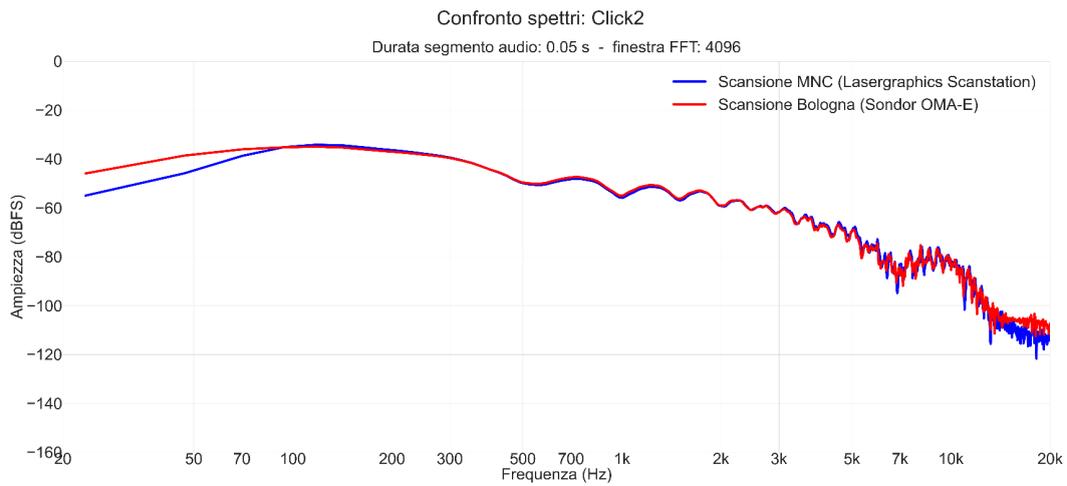
#### 3.4.2.4 Disturbi locali impulsivi

Per quanto riguarda i disturbi impulsivi come click e pop sono stati presi in considerazione diversi segmenti. Si riportano le porzioni audio che conducono alle osservazioni ritenute più rilevanti.

Prima di tutto, i click sono caratterizzati, come atteso, da una componente energetica ad ampio spettro. Gli spettri si discostano nella banda delle bassissime frequenze, alternando i risultati: in un caso raccoglie maggiore energia il Lasergraphics (figura 3.28), nell'altro il Sondor (figura 3.29). Nella parte alta dello spettro, sopra i 13 kHz, la scansione Sondor ha intensità maggiore.

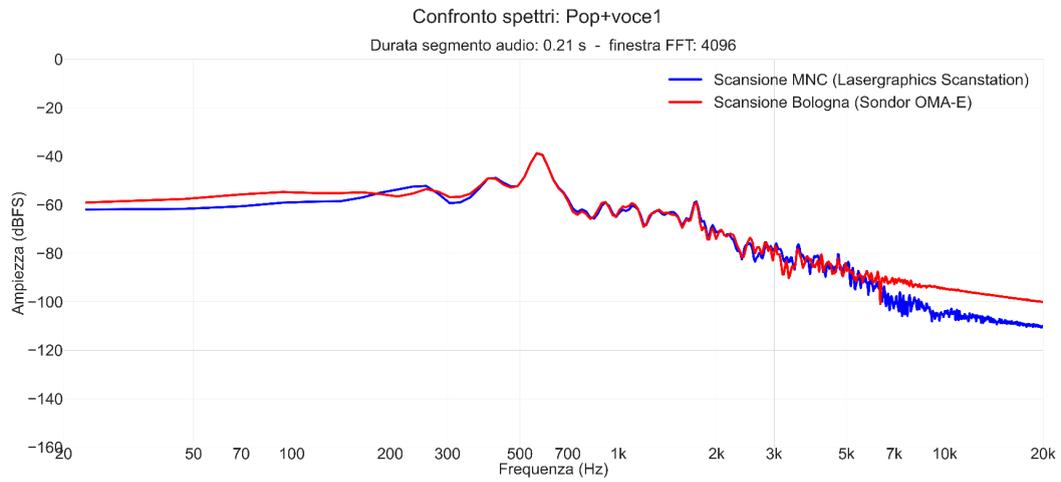


**Figura 3.28:** confronto spettrale di un disturbo impulsivo di tipo click.

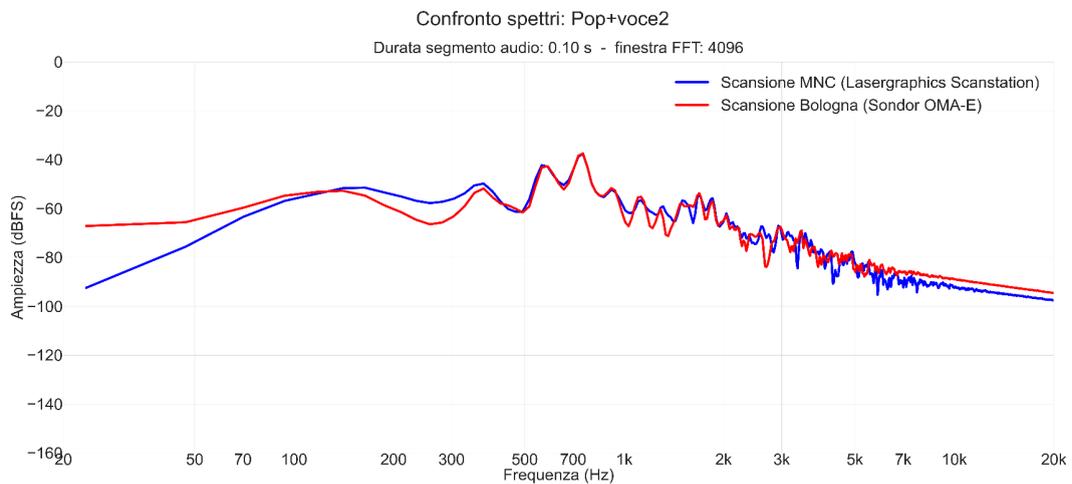


**Figura 3.29:** confronto spettrale di un altro click.

Per quanto riguarda i pop, sono spesso sovrapposti a segmenti di parlato (figure 3.30 e 3.31).



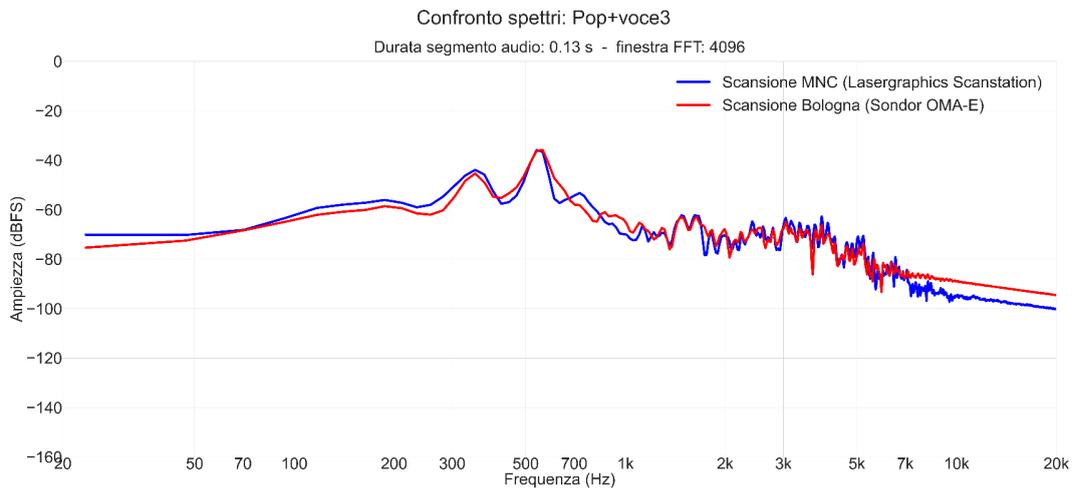
**Figura 3.30:** confronto spettrale di un pop.



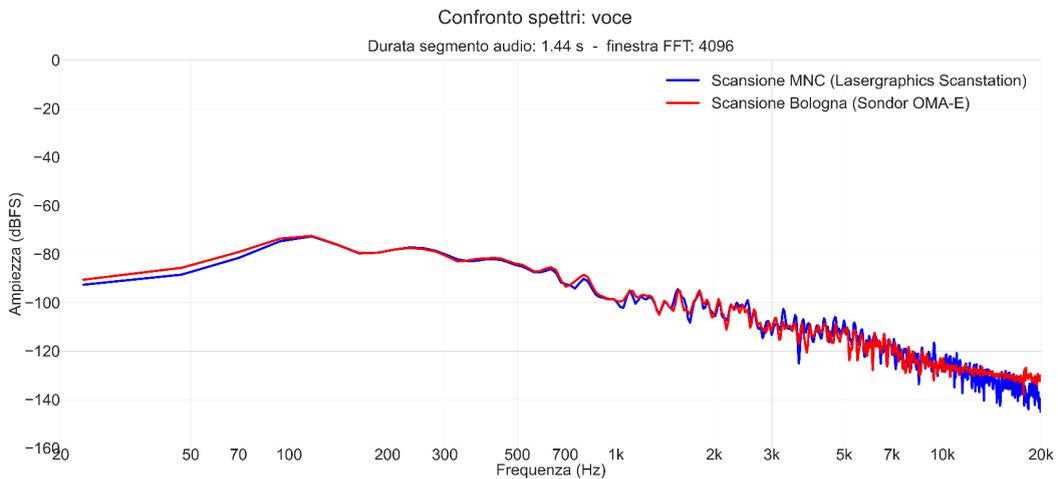
**Figura 3.31:** confronto ulteriore di un altro segmento contenente pop.

Nei grafici qui sotto (figure 3.32 e 3.33) si riporta il confronto tra un segmento di parlato degradato da pop e uno privo di disturbi particolari.

Tra le scansioni si osserva una certa differenza sotto i 200 Hz e sopra i 7 kHz.



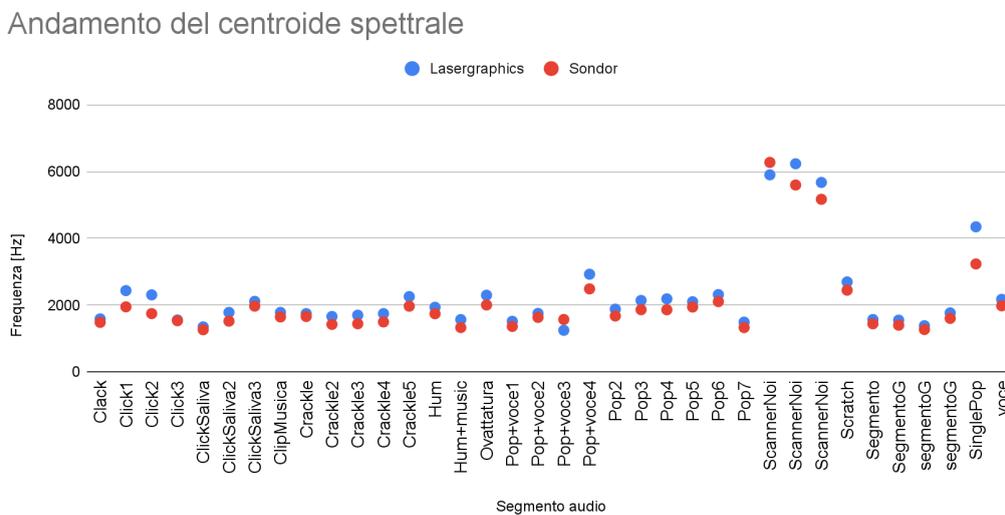
**Figura 3.32:** confronto spettrale di una porzione di parlato danneggiata da un pop.



**Figura 3.33:** confronto spettrale di una porzione di parlato priva di degradazioni.

### 3.4.2.5 Andamento del centroide spettrale

Come ultimo passaggio di questo confronto, si vuole osservare le differenze tra le due scansioni, tracciando l'andamento del centroide spettrale dei vari segmenti presi in considerazione. I valori sono tendenzialmente più elevati nella scansione effettuata con il Lasergraphics, dato che il centro di massa dello spettro è maggiore. La causa di questa tendenza risiede nel maggior contenuto in altissima frequenza nella scansione effettuata con il Lasergraphics, dovuto alla presenza di rumore a larga banda, che nel Sondor è invece stato eliminato (vedere paragrafo 3.4.2.1).



**Figura 3.34:** andamento del centroide spettrale nelle due scansioni.

### 3.4.2.6 Andamento del parametro di roughness

Per aggiungere un'ulteriore osservazione che possa ancora confermare i risultati del confronto, si è scelto di osservare il parametro psicoacustico della roughness.

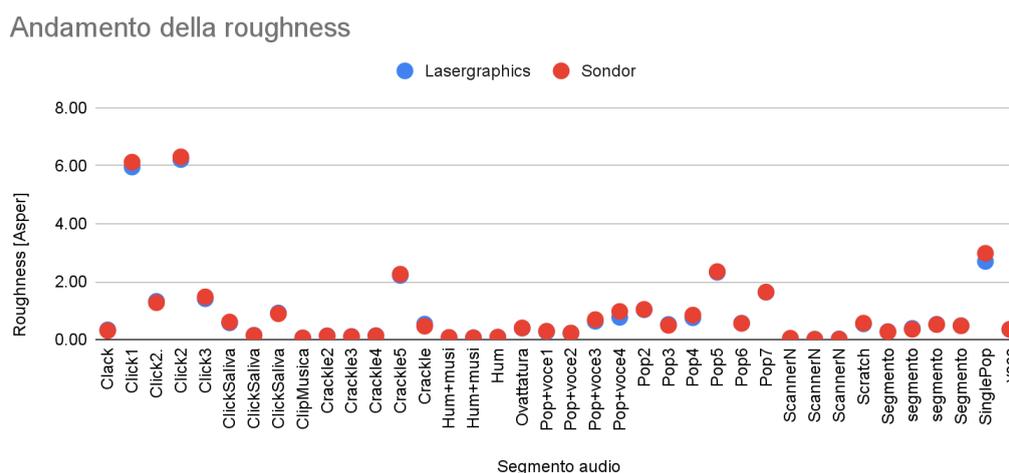
La roughness ("ruvidità") è definita come "sensazione psicoacustica di base associata a rapide variazioni dell'ampiezza che riducono la piacevolezza

sensoriale e la qualità dei rumori”.<sup>63</sup>

All'aumentare del valore di ruvidità, si riduce la piacevolezza di ascolto di un suono o di un rumore e funge da classificatore importante in molti contesti di valutazione dell'impatto dei rumori sui soggetti ad essi esposti.

Si è quindi deciso di estrarre i valori di roughness di tutti i segmenti analizzati, concludendo che la differenza riscontrata è sotto alla *jnd* di riferimento.<sup>64</sup>

Soltanto nel caso del rumore di fondo isolato si ottengono valori di differenza sopra alla soglia minima udibile, ma questi, su tre diversi segmenti di rumore di fondo, sono incoerenti, ovvero non identificano con congruenza che uno dei due scanner abbia un rumore di fondo più “ruvido”, ma concludono soltanto che su tre diverse misurazioni, i segmenti hanno ruvidità percettibilmente differenti, seppur così basse da essere approssimabili a 0.



**Figura 3.35:** andamento della roughness nelle due scansioni.

<sup>63</sup> Daniel & Weber, 1997, 113

<sup>64</sup> La soglia *jnd* (just noticeable difference) per la roughness è fissata al 17% della differenza relativa tra i valori: al di sotto di questo valore non è percepibile alcuna differenza in termine di roughness tra due suoni. (Daniel & Weber, 1997, 114)

## 4. OPERAZIONI DIGITALI DI RESTAURO

Nel passare degli anni le tecnologie ed i supporti ad alta qualità hanno aumentato in maniera consistente le aspettative dell'utente per quanto riguarda la qualità sonora. Le procedure di restauro si sono rese fondamentali non soltanto a fini filologici, ma anche a fini di riedizione per quel tipo di operazioni commerciali che intendono riportare al grande pubblico opere di interesse storico. Ciò ha creato una grande richiesta di restauri di materiale degradato.<sup>65</sup>

Come già anticipato nel capitolo 1, il restauro cinematografico inizia nel dominio analogico. Per molto tempo sono state utilizzate tecniche di restauro compiute totalmente nel dominio analogico, basate su editing manuale, sulla pulizia del supporto e poi sull'equalizzazione nel dominio della frequenza. Se è vero che tali operazioni sono meno rischiose ad un utilizzo inconsapevole e non mirato, non permettono di avere la stessa flessibilità garantita da processi digitali, che però sono rischiosi e talvolta conducono a degradazioni del segnale tali da rendere il restauro non ottimale.<sup>66</sup>

I processi digitali di restauro audio sono basati su software di vario genere. Spesso sono composti da moduli specializzati nel processamento delle singole problematiche, mentre, altre volte, sono composti da interfacce più immediate, che permettono, per esempio, di procedere a isolare una voce dal rumore di fondo con pochi semplici interventi.

Verranno trattate in seguito le metodologie di gestione di alcune delle più diffuse degradazioni e alcuni software utilizzati per trattarle.

---

<sup>65</sup> Canazza, 2013, 60

<sup>66</sup> ibidem

## 4.1 Rimozione dei click

Le caratteristiche e le origini dei click sono già state argomentate in precedenza nel paragrafo 3.3.1.

Esistono diversi approcci per il restauro dei disturbi impulsivi come i click. In generale, la procedura si costituisce di due fasi: la localizzazione dei disturbi, basata sulla stima della loro posizione all'interno della forma d'onda, e la loro successiva rimozione, ricostruendo il segnale audio nei punti danneggiati.

Il rilevamento, nella sua metodologia più semplice, è basato sul filtraggio passa-alto: dato che la maggior parte delle informazioni del segnale sono sotto gli 8 kHz, un rilevatore a soglia dopo al filtro equalizzatore può permettere di rilevare i disturbi impulsivi, che, invece, hanno componenti in frequenza in tutto lo spettro. Questo sistema ha come unico parametro il valore di soglia, ma diventa inutilizzabile se i click da rimuovere vanno a degradare un segnale che ha un contenuto importante in alta frequenza, come ad esempio nella musica elettroacustica.

Esistono altre tecniche di rilevamento, basate su modelli autoregressivi (AR). Il segnale viene rappresentato da un processo autoregressivo stazionario a breve termine, dove "il segnale all'istante  $nT$  è espresso dalla somma pesata di  $P$  (ordine del modello) campioni precedenti e di un *residuo*  $e(n)$ :"

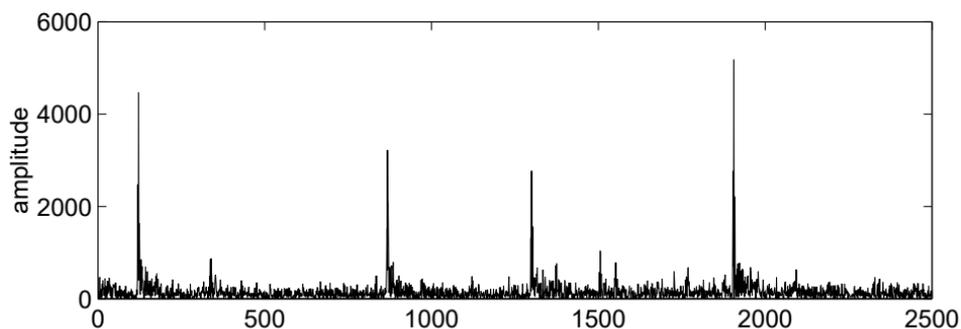
$$s(n) = \sum_{i=1}^P s(n-i) a_i + e(n)$$

Questo modello rappresenta bene tanto segnali armonici che rumorosi." <sup>67</sup>

In questo caso viene calcolata la differenza tra il segnale originale e la parte armonica del modello AR. I click, che non sono rappresentati dal modello AR, risultano amplificati e il resto del segnale originale invece ridotto, permettendo la rilevazione tramite una soglia (vedere figura 4.1).

---

<sup>67</sup> Canazza, 2013, 61



**Figura 4.1:** rilevamento click con modello AR.

La precisione di localizzazione temporale di questo sistema diminuisce all'aumentare dell'amplificazione relativa dei click, rendendolo meno affidabile, ad esempio, per la rimozione di crackle, perché i click che sono tra loro molto ravvicinati possono non venire rilevati, specialmente se hanno ampiezze molto diverse fra loro.<sup>68 69</sup>

Una volta localizzati i click è necessario procedere alla loro rimozione, ricostruendo il segnale a partire dai campioni non deteriorati adiacenti a quelli rimossi, per esempio si può procedere all'interpolazione con metodo LSAR<sup>70</sup>. I dati corrotti sono considerati mancanti, non si assume alcuna ipotesi sul valore che il segnale assume durante i click, bensì vengono stimati i parametri del modello AR utilizzando i campioni del segnale disponibili prima e dopo la sezione danneggiata. Il segnale nelle posizioni dei click viene quindi interpolato usando il modello AR stimato. Ciò avviene utilizzando il problema dei minimi quadrati per trovare i valori dei campioni mancanti che meglio si adattano al modello AR stimato. L'algoritmo è efficace e preciso, ma ha delle limitazioni riguardo click di durata elevata e entra in difficoltà importanti con segnali che non sono stazionari a breve termine, generando artefatti udibili.<sup>71</sup>

<sup>68</sup> Canazza, 2013, 63-69

<sup>69</sup> Godsill & Rayner, 1998, 131

<sup>70</sup> Least Squares AR-based, sviluppato da Vaseghi e Rayner nel 1990 e originariamente pensato per la rimozione di tick digitali nei sistemi CD. (Canazza, 2013, 66)

<sup>71</sup> Canazza, 2013, 63-69

## 4.2 Rimozione del rumore di fondo a larga banda

Per la rimozione del rumore di fondo a banda larga, che, come già detto, rappresenta una delle degradazioni più diffuse nel materiale audio d'archivio, verrà discusso un insieme di tecniche denominato *attenuazione spettrale a breve termine* (STSA).

Queste tecniche operano applicando un'attenuazione time-based allo spettro a breve termine del segnale degradato. Quest'ultimo viene suddiviso in brevi segmenti, utilizzando, ad esempio, la trasformata di Fourier a breve termine (STFT), permettendo di analizzare il contenuto in frequenza del segnale in funzione del tempo. Ad ogni finestra viene applicato un guadagno, determinato da una stima della quantità di rumore presente.

Queste stime possono essere calcolate con varie *regole di soppressione*, le più diffuse sono basate sul *filtro di Wiener* o sulla *sottrazione di potenza*.

Il filtro di Wiener mira a preservare il più possibile il segnale desiderato, cercando di ridurre al minimo la potenza del rumore di fondo. Nello specifico, per realizzare questo filtro, si procede calcolando lo spettro di potenza del segnale rumoroso e, successivamente, si stima lo spettro di potenza del rumore in sé, analizzando sezioni di rumore isolato dal contenuto di interesse del segnale (pause o porzioni di silenzio). Si procede calcolando il rapporto segnale-rumore (SNR) tra i due spettri di potenza, per ogni banda di frequenza. Viene quindi calcolato un *guadagno di Wiener* per ciascuna banda, calcolato in base all'SNR e pensato per attenuare bande di frequenza in cui l'SNR è basso e mantenere invece intatte le bande in cui l'SNR risulta maggiore (e quindi il segnale di interesse è più forte del rumore). Il guadagno di Wiener alla frequenza  $\omega$  si calcola come il rapporto tra lo spettro di potenza del segnale rumoroso  $X$  e la somma degli spettri di potenza del segnale  $X$  e del suo rumore  $N$ :

$$G(\omega) = \frac{S_X(\omega)}{S_X(\omega) + S_N(\omega)}$$

Mentre i vantaggi di questo filtro sono la sua relativa facilità di implementazione e la sua capacità di ridurre al minimo la distorsione complessiva del segnale, esso può comportare artefatti udibili in caso di SNR troppo bassi. Inoltre è da sottolineare che il metodo di Wiener richiede una stima accurata dello spettro di potenza, che potrebbe non essere sempre facile da ottenere in tutte le condizioni.

La sottrazione di potenza semplifica l'approccio, rispetto al filtro di Wiener, perché non richiede una stima esplicita dell'SNR. Per prima cosa, va calcolato, anche in questo caso, lo spettro di potenza sia del segnale rumoroso  $X$ , sia del rumore  $N$ . Lo spettro di potenza del rumore stimato viene sottratto allo spettro di potenza del segnale intero, evitando valori negativi grazie a un valore di soglia, che porta la differenza a zero se la differenza risulta negativa.

Il segnale pulito si ottiene applicando al segnale un guadagno pari alla radice quadrata del guadagno di Wiener.

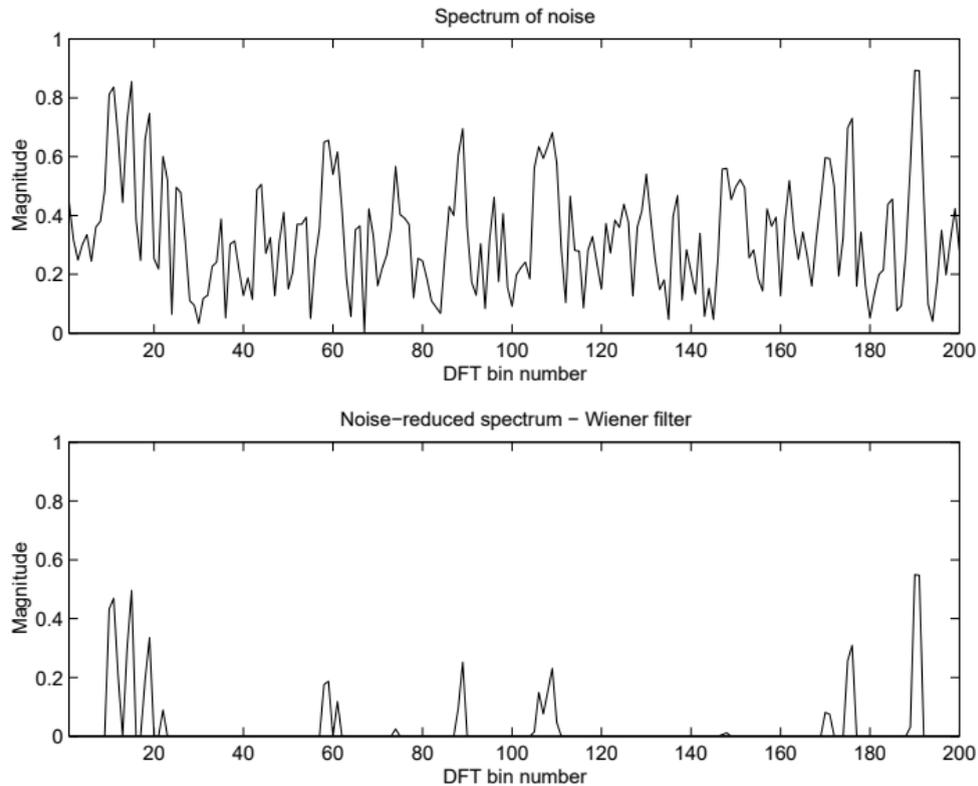
$$G(\omega) = \sqrt{\frac{S_X(\omega)}{S_X(\omega) + S_N(\omega)}}$$

La maggior parte degli algoritmi di soppressione del rumore utilizzati nel campo del restauro audio sono basati sulle tecniche STSA. Questo è probabilmente dovuto al fatto che questo approccio è non-parametrico e, quindi, può essere applicato a molti tipi di segnali eterogenei fra loro.

Esistono altre tecniche, ottimizzate nello specifico per il parlato, ma la loro applicazione non è ottimale nella maggioranza dei casi, dove i segnali da restaurare non sono esclusivamente vocali.

Il maggior rischio delle tecniche STSA riguarda l'introduzione del cosiddetto *rumore musicale*, costituito da artefatti udibili simili a brevi fischi o scampanellii, molto fastidiosi e talvolta più spiacevoli del rumore di fondo a banda larga di partenza. Questi artefatti sono dovuti alle fluttuazioni randomiche nella stima dello spettro di potenza. Le regole di soppressione non

sono lineari e sebbene riescano a attenuare efficacemente le componenti a bassa ampiezza, quelle a elevata ampiezza lasciano dei residui, isolati fra di loro e quindi percepibili come componenti tonali all'interno del segnale restaurato, con un effetto influenzato anche dalla loro rapida variazione temporale (figura 4.2).



**Figura 4.2:** segnale rumoroso (sopra), componenti residue della riduzione (sotto).

Il rumore musicale si può attenuare. Usando tecniche di soppressione standard (citate sopra), l'unico mezzo efficace per farlo è effettuare una pesante sovrastima del livello di rumore, provocando una cancellazione, però, anche di componenti del segnale utile.

Esistono, però, svariate tecniche alternative a quelle standard, che producono stime statisticamente più stabili dello spettro di potenza, come la *regola di soppressione di Ephraim e Malah*, che basa il calcolo del guadagno su stime multiple del SNR.

Gli sviluppi più importanti della riduzione del rumore a banda larga nel dominio spettrale sono quelli che ottimizzano psicoacusticamente gli algoritmi, dotandoli della capacità di considerare criteri percettivi.<sup>72 73</sup>

### 4.3 Rimozione della riverberazione

La riverberazione è una caratteristica non trascurabile del suono, specialmente nei contesti in cui un suono non viene acquisito in un contesto controllato, come quello di uno studio di registrazione o, nel caso della produzione cinematografica, in un teatro di posa.

In un ambiente riverberante il segnale si propaga nell'aria e viene riflesso dai muri, dal soffitto, dal pavimento e da qualsiasi oggetto riflettente della stanza: il segnale registrato è una somma di tutte le infinite copie del segnale, ritardate e decadute nel tempo. La riverberazione inoltre aumenta con l'aumentare della distanza tra la sorgente sonora e il microfono.<sup>74</sup>

La quantità di riverbero registrata in un segnale audio può incidere in maniera significativa sull'intelligibilità del contenuto audio. La rimozione, o più diffusamente la riduzione, di questa quantità si può escludere dalle operazioni fondamentali di restauro, ma i software commerciali hanno reso più facile l'accesso a queste tecnologie i risultati sono spesso stupefacenti.

Sempre ricordando quanto sia fondamentale evitare l'alterazione consistente dell'originalità dell'opera, nelle fasi di postproduzione digitale del restauro utilizzare in maniera delicata e non invasiva questo tipo di algoritmi può aiutare a recuperare una nitidezza alle informazioni audio, ma, ovviamente, ci si rifà sempre alla sensibilità del restauratore, che deve saper usare questi strumenti in maniera non invasiva.

Gli approcci principali adottati dai sistemi di de-riverberazione sono tre: il *beamforming*, lo *speech enhancement* e la *blind deconvolution*.

---

<sup>72</sup> Godsill & Rayner, 1998, 136-148

<sup>73</sup> Canazza, 2013, 72-78

<sup>74</sup> Wang & Wang, 2020, 941

Il primo, ovvero il beamforming, è un processo *"multiple input, single output"*, dipendente dalla disponibilità di segnali multi-microfonici e quindi, nella maggior parte dei casi, dall'uso di trasduttori di tipo *array*. Il segnale ricevuto dai diversi microfoni viene filtrato e pesato in modo da enfatizzare ciò che proviene dalla sorgente di interesse e ridurre tutti gli altri suoni. Viene effettuato, in sostanza, un filtraggio spaziale in grado di costruire un fascio di sensibilità (*"beam of sensitivity"*) nella direzione desiderata.

Questo tipo di approccio funziona molto meglio nel trattamento di sorgenti fortemente localizzate, ma perde di efficacia nel campo diffuso.<sup>75 76</sup>

Un'altra direzione adottabile per la riduzione della riverberazione nel segnale è quella dello *speech enhancement*. Si tratta di un processo *"single input, single output"* dove i segnali sono modificati per corrispondere a un modello a priori, basato sulla forma d'onda o sullo spettro di un segnale pulito.

Negli anni sono stati proposti diversi modelli, basati su filtraggio avanzati di specifiche caratteristiche del segnale. In poche parole, migliorando la pulizia del segnale, individuando specifiche caratteristiche in forma d'onda e spettro e, alterandole, si riduce anche la riverberazione, guadagnando in chiarezza e intelligibilità.

Questo approccio è adatto esclusivamente al parlato, ma i modelli sono efficienti e adattabili anche ad applicazioni real-time.<sup>77 78</sup>

Il terzo approccio fondamentale è quello della *blind deconvolution*, dove le risposte all'impulso vengono identificate "alla cieca", ovvero usando soltanto il segnale registrato per realizzare un filtro inverso della risposta all'impulso del riverbero da rimuovere, in modo da compensare l'effetto acustico della riverberazione stessa. Questo tipo di processo è tecnicamente e computazionalmente molto complesso.

---

<sup>75</sup> Naylor & Gaubitch, 2010, 4-5

<sup>76</sup> *ivi*, 8-10

<sup>77</sup> *ivi*, 4-5

<sup>78</sup> *ivi*, 10-11

## 4.4 Separazione delle sorgenti

Nel contesto del restauro di segnali audio, può essere di cruciale importanza la *source separation*, ovvero la procedura di separazione di un segnale mono o stereofonico nelle sue componenti.

Gli utilizzi di questo tipo di tecnologia sono molteplici, infatti ottenendo le tracce separate delle sorgenti che caratterizzano un segnale audio complesso si può ribilanciare il volume di ogni elemento o per esempio correggere la sua posizione nello spazio stereofonico. Con la separazione delle sorgenti, inoltre, viene reso tecnicamente possibile il passaggio di una traccia da mono a stereo, o da stereo a surround, cosa che, nel contesto di una riedizione, può avere notevoli conseguenze.

È facile dedurre la varietà di possibilità che una tecnologia del genere, se ben ottimizzata, può offrire, ma anche la quantità di implicazioni etiche che, perlomeno per quanto riguarda il restauro cinematografico, vengono chiamate in causa.<sup>79 80</sup>

È necessario sottolineare che maggiore è la complessità del segnale originale, maggiore è la difficoltà nel creare sistemi in grado di separare il segnale in componenti diverse, elaborabili separatamente.

I sistemi MSS usano la STFT per convertire il segnale dal dominio del tempo a quello della frequenza.

Successivamente avviene lo step di modellazione della sorgente, che si concentra sull'analisi dello spettrogramma per individuare sia le caratteristiche tipo di ogni sorgente, sia la sua posizione nello spazio stereofonico (nel caso di segnali multicanale).

Per esempio, i modelli KAM (*Kernel Additive Model*) sono in grado di sfruttare le caratteristiche locali dello spettrogramma per separare le sorgenti.

Invece, i modelli di fattorizzazione dello spettrogramma, come ad esempio i NMF (*Non-negative Matrix Factorization*), si basano sulla scomposizione dello

---

<sup>79</sup> Stöter et al., 2019, 1-2

<sup>80</sup> Cano et al., 2019, 31

spettrogramma in componenti più semplici, in grado di ricostruire il comportamento di ogni singola sorgente e riconoscerlo rispetto alle altre. Infine, esistono modelli basati sul principio della sintesi additiva, secondo cui ogni segnale è approssimabile come somma di sinusoidi modulate, sia nella frequenza che nell'ampiezza. Questo tipo di rappresentazione funziona bene con segnali armonici, ma gli MSS sinusoidali sono poco diffusi, a causa della difficoltà richiesta per raggiungere una rappresentazione realistica delle sorgenti.<sup>81</sup>

---

<sup>81</sup> Cano et al., 2019, 35-40

## 5. CONTRIBUTO DELLE RETI NEURALI ALLE TECNICHE DI RESTAURO

Lo sviluppo dei metodi visti nel capitolo 4 ha inizio già negli anni '90 con metodi che potremmo definire tradizionali, poiché sfruttano le tecniche di elaborazione dei segnali classiche.

La rivoluzione dell'efficacia delle applicazioni di denoising e rimozione delle imperfezioni audio arriva, però, con gli algoritmi basati su machine learning.

Negli anni sono stati sviluppati molti diversi software pensati per gestire le problematiche di degradazione del suono. Il più famoso e diffuso di questi software è Izotope RX, che nel 2024 è arrivato alla sua undicesima versione e rappresenta, almeno nella suite *Advanced*, una collezione completa di moduli potenti e sofisticati per la gestione di una varietà di disturbi impressionante. Izotope RX è un software commerciale molto utilizzato negli ambiti dell'industria musicale e cinematografica, in cui svetta come standard dell'*audio restoration* da diversi anni.

Già da diverse versioni, ben prima dell'arrivo della cosiddetta "AI Spring"<sup>82</sup>, la RX Advanced Suite di Izotope è dotata di moduli basati su tecnologie di Machine Learning<sup>83</sup>, ma queste tecnologie sono sempre state poco diffuse fuori da contesti specificatamente professionali a causa, soprattutto, del prezzo di vendita del software.<sup>84</sup>

Negli ultimi anni si sono diffusi molti strumenti, sia commerciali che open-source, che hanno reso disponibile a tutti algoritmi immediati ed efficaci per la gestione di problematiche disparate. Prima di descriverne alcuni (nel

---

<sup>82</sup> Il boom dell'intelligenza artificiale, iniziato nei tardi anni '10 del nostro secolo e concretizzatosi a partire dalla diffusione di Chat GPT, nel 2022.

<sup>83</sup> I moduli in questione sono "De-rustle" e "Dialogue Isolate" disponibili in RX 6 Advanced. (Wichern, 2017)

<sup>84</sup> La versione più avanzata, la Advanced Suite, è in vendita a circa 1300 dollari.

capitolo 6), verranno presentate alcune soluzioni tecnologiche, basate su AI, per migliorare le tecniche di restauro viste nel capitolo precedente.

## 5.1 Reti neurali per la rilevazione automatica dei difetti

Di recente le Deep Neural Networks hanno iniziato ad essere utilizzate per sviluppare sistemi di rimozione del rumore, ma anche per sistemi di *audio inpainting*, ovvero per la ricostruzione di porzioni del segnale mancanti o corrotte, con risultati molto migliori rispetto alle tecniche convenzionali basate su modelli autoregressivi.

Una delle applicazioni dove le reti neurali profonde (DNN) possono risultare utili è la rilevazione dei difetti. A questo fine è necessario disporre di grandi librerie per allenare la rete a individuare determinate specifiche problematiche in automatico. Per esempio, si può creare un rilevatore di click che non necessiti di tecniche standard di elaborazione dei segnali.<sup>85</sup>

Sono stati già proposti dei framework che fanno questo tipo di operazione sulla musica e sono stati trainati su dataset di tracce mp3 per rilevare click digitali e glitch. La sfida dell'approccio data-driven è quella di creare un sistema di rilevazione automatica che sia però in grado di distinguere i difetti da elementi che sono invece caratteristici del contenuto audio.

Come esempio di questo tipo di sistema si può fare riferimento Hook-Net, un modello presentato per la rilevazione di artefatti nei segnali audio. La rete è alimentata con segmenti di forma d'onda grezzi, mentre in output viene prodotta una classificazione binaria ("originale" o "degradato"). Questi risultati sono prodotti in una serie temporale che indica in quali campioni l'audio presenta artefatti oppure no.

Per la rilevazione dei click, sono stati introdotti click digitali simulati dentro a svariate decine di migliaia di tracce mp3 di diversi generi musicali presenti nella libreria open source FMA-Large. Sono stati confrontati i risultati di

---

<sup>85</sup> Vedere paragrafo 4.1

rilevazione con un rilevatore di click standard della libreria Essentia.<sup>86</sup> La precisione del rilevatore basato su Hook-Net è maggiore, soprattutto nella sua capacità di estendersi a più generi musicali, distinguendo artefatti da elementi impulsivi a largo spettro intenzionali, che sono molto più frequenti in generi come il pop e l'elettronica.<sup>87</sup>

Questa soluzione, agli scopi del restauro di piste analogiche, potrebbe presentare un grande limite: come detto per allenare la DNN si usano click digitali introdotti artificialmente, in questo modo si può addestrare la rete per distinguere degradazioni indesiderate da elementi impulsivi che fanno parte del contenuto intenzionale del segnale audio. Le caratteristiche dei disturbi analogici potrebbero essere diverse da quelle dei click digitali e quindi questo sistema potrebbe essere più complesso da utilizzare, a causa della difficoltà nel reperire grandi data-set adatti allo scopo.

## 5.2 Reti neurali per la rimozione del rumore

Come anticipato nel capoverso precedente, negli ultimi anni le reti neurali si sono dimostrate molto promettenti per applicazioni di rimozione del rumore nei segnali audio. I sistemi che utilizzano questa famiglia di tecnologie per ora sono più diffuse nell'ambito dello *speech enhancement* e si basano sul tentativo di previsione dei valori del rumore nel segnale in input, in maniera da poterlo poi sottrarre dal file di partenza, ottenendo una versione pulita dell'informazione sonora. Utilizzando una rete neurale questo tipo di operazione può essere fatta senza l'utilizzo di alcuna stima statistica, a differenza dei metodi tradizionali usati fino a questo momento.

I dataset con cui queste reti neurali sono alimentate sono formati da una enorme quantità di coppie di segnali (uno pulito e uno rumoroso). La rete apprende il modo di produrre la migliore stima possibile del rumore da rimuovere, in modo da rimuovere il rumore dal segnale in questione.<sup>88</sup>

---

<sup>86</sup> Universitat Pompeu Fabra, Music Technology Group, [www.essentia.upf.edu](http://www.essentia.upf.edu)

<sup>87</sup> Wolff et al., 2022, 2-8

<sup>88</sup> Mohammed & Radhika, 2022, 34

Sono stati proposti diversi modelli in grado di sfruttare le DNN per diversi obiettivi, dal miglioramento della riduzione del rumore in tempo reale per scopi legati alle telecomunicazioni, alla classificazione di segmenti parlati, rumore e suoni ambientali a partire da sorgenti complesse.<sup>89</sup>

Ad esempio, partendo da un'architettura usata per la segmentazione delle immagini (UNet) si può costruire un modello addestrato con gli spettrogrammi dei segnali rumorosi. L'80% del dataset di partenza viene usato per allenare la rete, mentre il restante 20% per testare i risultati. La qualità del sistema può essere valutata con dei parametri di quality assessment pensati per il parlato, come il PESQ<sup>90</sup> e lo STOI<sup>91</sup>.

È interessante notare che poiché UNet è pensato per lavorare con le immagini offre risultati ottimi con la previsione del risultato a partire dagli spettrogrammi del segnale. Regolando la rete UNet sui parametri specifici legati ai segnali vocali si possono ottenere risultati ancora migliori.

È chiaro che riguardo l'applicazione di questo specifico modello per applicazioni di restauro ci sono dei limiti, dato che i sistemi proposti sono adattati specificamente alla qualità del parlato e alla rimozione di tutto il resto. L'intenzione, però, è quella di mostrare la versatilità di questo tipo di reti, se addestrate su un dataset correttamente inerente all'applicazione di cui si necessita.

### **5.3 Reti neurali per algoritmi di de-reverberation**

Le reti neurali sono applicate anche nel miglioramento delle tecnologie di rimozione della riverberazione, dove possono venire addestrate, per esempio, per prevedere le componenti reali e immaginarie della componente diretta del segnale, a partire da segnali rumorosi e riverberati. I risultati ottenuti sono

---

<sup>89</sup> Mohammed & Radhika, 2022, 35-36

<sup>90</sup> Perceptual Evaluation of Speech Quality, usato nell'industria delle telecomunicazioni per la valutazione della qualità oggettiva della voce. (ivi, 39)

<sup>91</sup> Short Term Objective Intelligibility, usato per valutare l'intelligibilità del parlato e non la sua qualità generale. (ibidem)

migliori rispetto a quelli garantiti dalle tecniche tradizionali, sia per quanto riguarda la qualità del parlato che il riconoscimento vocale, e sono applicabili sia a segnali monofonici che a segnali multicanale (registrati con array microfonic).

Questo tipo di approccio sfrutta due DNN: la prima riesce a prevedere le componenti reali e immaginarie che descrivono il segnale diretto; la seconda, invece, riesce a migliorare la stima del segnale diretto effettuata dalla prima rete, sfruttando le informazioni spaziali ottenute dall'elaborazione del segnale tramite *beamforming*<sup>92</sup>, permettendo un migliore isolamento del segnale target dalle componenti riverberate.

Le reti neurali sono addestrate su dataset composti da una grande quantità di risposte all'impulso simulate. È necessario sottolineare che questo sistema è ottimizzato per il parlato, sfruttando, dunque, la premessa che la sorgente sia fortemente localizzata.<sup>93</sup>

Ai fini delle applicazioni al restauro, è già stato discusso che l'utilizzo delle tecniche di rimozione del riverbero va valutato con estrema delicatezza. L'intenzione di questo paragrafo è di introdurre gli sviluppi e i vantaggi, in termini di risultati, che le reti neurali profonde possono portare, giustificando anche la diffusione di strumenti commerciali dedicati, che verranno esposti successivamente.

## 5.4 Reti neurali per la separazione delle sorgenti

Sono stati presentati anche modelli che cercano di applicare le DNN alla separazione delle sorgenti. Sia per quanto riguarda le applicazioni di tale tecnologia alla musica, sia alla separazione del parlato da un segnale misto, un sistema basato su reti neurali profonde può essere usato per modellare lo spettro delle diverse sorgenti sonore, con ottimi risultati su segnali monofonici, ma con ottimi presupposti di estensione a sistemi multicanale.

---

<sup>92</sup> Ovvero le tecniche di *beamforming* (vedere paragrafo 4.3).

<sup>93</sup> Wang & Wang, 2020, 941-949

Questo tipo di approccio può, ad esempio, essere portato a termine sfruttando un algoritmo iterativo in grado di stimare i parametri spaziali e spettrali che descrivono ciascuna sorgente nel segnale audio, utilizzandoli, poi, per produrre dei filtri di Wiener multicanale in grado di migliorare la qualità della separazione.<sup>94</sup>

---

<sup>94</sup> Nugraha et al., 2016, 1652-1664

# 6. PANORAMICA DEI SOFTWARE

## 6.1 Software commerciali

### 6.1.1 Izotope RX

Come è già stato chiarito in diversi momenti di questa trattazione, Izotope RX è la suite di software per il restauro audio più completa e nota, rappresentando l'*industry standard* per le applicazioni nel campo dell'editing, del restauro, della rimozione dei disturbi, sia nelle produzioni cinematografiche che musicali.

Il software offre una interfaccia di editing audio complessa, che permette sia di visualizzare lo spettrogramma del suono, dando la possibilità di intervenire direttamente su di esso con strumenti di ritocco, sia di trattare il suono con appositi moduli dedicati a specifiche problematiche come clicks, crackle, hum e offre diversi denoiser di alto livello per rimuovere il rumore di fondo.



**Figura 6.1:** modulo Spectral De-noise di Izotope RX 11.

Nella versione *Advanced* esistono moduli in grado di effettuare operazioni come la regolazione dell'intonazione del parlato, la rimozione del riverbero, un sistema di *ambience matching*<sup>95</sup> e la possibilità di creare profili di equalizzazione per compensare, per esempio, le differenze timbriche tra tracce diverse e ottenere lo stesso suono da più microfoni differenti.

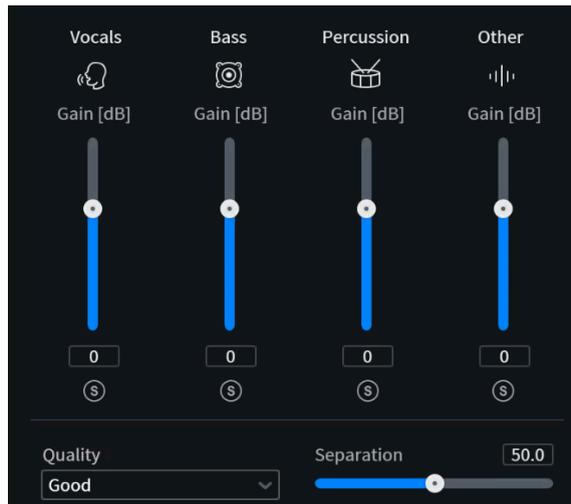


**Figura 6.2:** modulo De-reverb di Izotope RX 11.

Tra i moduli disponibili ne esiste uno di source separation, chiamato Music Rebalance.

---

<sup>95</sup> Ricrea un roomtone basandosi su un input costituito da un segmento di rumore di fondo, permettendo di coprire eventuali buchi e di uniformare lo sfondo sonoro di tracce differenti.



**Figura 6.3:** modulo Music Rebalance di Izotope RX 11.

Le ultime versioni di RX includono un modulo chiamato *Repair Assistant* che analizza il contenuto del segnale audio cercando di individuare autonomamente problematiche e proporre alcune soluzioni rapide per migliorare la qualità complessiva della traccia in maniera rapida e intuitiva, tramite interfaccia semplificata.<sup>96 97</sup>



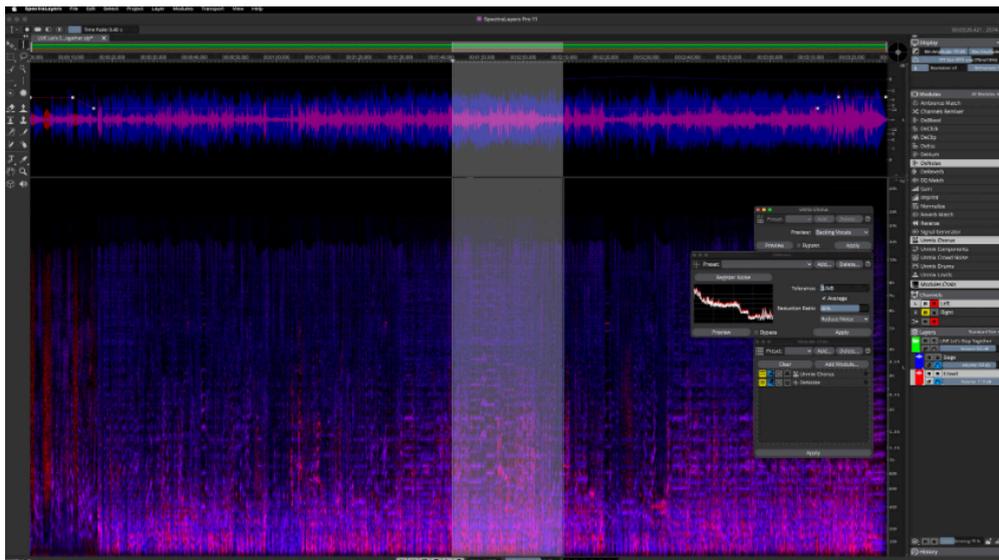
**Figura 6.4:** modulo Repair Assistant di Izotope RX 11.

<sup>96</sup> Corelli, 2019

<sup>97</sup> iZotope, Inc., [www.izotope.com](http://www.izotope.com)

## 6.1.2 Steinberg SpectraLayers

SpectraLayers è l'alternativa sviluppata da Steinberg a Izotope RX. Offre, prima di tutto, dei potenti strumenti di visualizzazione del suono. Rispetto a Izotope concentra maggiormente le sue possibilità nell'ambito dell'editing spettrale, ma offre anche una discreta quantità di moduli per affrontare e gestire disturbi specifici del suono.



**Figura 6.5:** Interfaccia dell'editor spettrale di Steinberg SpectraLayers 11.

La suite è più diffusamente utilizzata nell'ambito del sound design e per applicazioni sperimentali di ricerca sonora, piuttosto che nel restauro del suono, ma è uno strumento versatile anche per la risoluzione di diversi tipi di degradazioni.

Il software è potenziato da algoritmi basati su reti neurali in grado di analizzare lo spettro e riconoscere pattern, oltre che aumentare la velocità di calcolo.<sup>98</sup>

---

<sup>98</sup> Steinberg Media Technologies GmbH, [www.steinberg.net](http://www.steinberg.net)

### 6.1.3 Waves Clarity

Un esempio notevole di tool commerciale in grado di semplificare molto il processo di denoising e di isolamento del parlato dai disturbi è la serie Clarity della Waves.

La serie consiste di due plugin, a loro volta disponibili in due versioni, quella base e quella pro.

Clarity Vx è un tool di rimozione del rumore, in grado di isolare il segnale parlato dai disturbi e ridurre il rumore di fondo, i rumori ambientali e le imperfezioni. Utilizza algoritmi basati su reti neurali, ne sono disponibili tre, due chiamate "board", da alternare in base al tipo di rumore da gestire, e una di tipo "eco", in grado di ottimizzare la richiesta hardware e ridurre il carico computazionale sulla CPU.

La comodità di questo tipo di plugin risiede nella sua semplicità di utilizzo: con un solo intervento sull'interfaccia si può decidere come intervenire sul segnale, senza rovinare le caratteristiche originali e raggiungendo il desiderato livello di isolamento della voce.



**Figura 6.6:** Waves Clarity Vx.

Clarity Vx Pro rende l'approccio meno immediato, ma offre una quantità elevata di controlli pensati per garantire una versatilità superiore. Si può effettuare fine-tuning per un controllo multibanda sul denoiser.



**Figura 6.7:** Waves Clarity Vx PRO.

Il secondo plugin della serie si chiama Clarity Vx DeReverb ed è pensato specificamente per la rimozione della riverberazione. Anche questo tool è basato sulla tecnologia *Waves Neural Networks* e offre diverse possibilità in base al tipo di voce da pulire e alle necessità del sistema hardware.



**Figura 6.8:** Waves Clarity Vx DeReverb.

La versione Pro permette una regolazione di fino su tutto lo spettro, dando la possibilità di effettuare un processo di rimozione del riverbero con controllo multibanda. Offre parametri aggiuntivi come l'attenuazione della coda della riverberazione, oltre che una maggiore quantità di modelli di machine learning a disposizione.<sup>99</sup>



**Figura 6.9:** Waves Clarity Vx DeReverb PRO.

<sup>99</sup> Waves Audio Ltd., [www.waves](http://www.waves)

## 6.1.4 Supertone CLEAR

Preceduto da una popolarissima versione beta gratuita chiamata Goyo, il Supertone CLEAR è un plugin di denoising molto efficace e intuitivo. È in grado di scomporre una traccia in componente ambientale (e relativo rumore di fondo), contenuto vocale e riverberazione e permette di bilanciare con tre comandi diretti il quantitativo di ciascuna di queste sorgenti all'interno della traccia. Utilizza una rete neurale proprietaria che promette di ridurre al minimo gli artefatti.<sup>100</sup>



**Figura 6.10:** Supertone CLEAR.

---

<sup>100</sup> Supertone Inc., [www.supertone.ai](http://www.supertone.ai)

## 6.1.5 Acon Digital Restoration Suite

Restoration Suite è un insieme di plugin sviluppato da Acon Digital composto da DeNoise, DeHum, DeClick e DeClip.

DeNoise è un denoiser disegnato per rimuovere o ridurre rumori di fondo con caratteristiche diverse fra loro. Dispone di una modalità adattiva che riconosce il rumore e modella una maschera adattiva in grado di rispondere alla variabilità in frequenza del profilo del rumore di fondo. Promette di gestire anche i rumori che variano nel tempo, come i *rustle* tipici dello sfregamento dei microfoni lavalier contro i vestiti.



**Figura 6.11:** modulo DeNoise 2 della Acon Digital Restoration Suite.

DeHum è un plugin per la rimozione di ronzii e disturbi elettrici di vario genere, ma anche rumori con caratteristiche tonali.



**Figura 6.12:** modulo DeHum 2 della Acon Digital Restoration Suite.

DeClick è progettato per la rimozione di click e crackle, ma è in grado anche di individuare e rimuovere plosive indesiderate nel parlato.



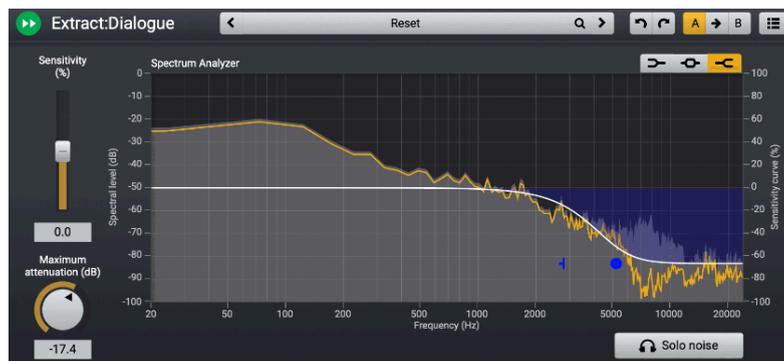
**Figura 6.13:** modulo DeClick 2 della Acon Digital Restoration Suite.

DeClip offre un'intuitiva interfaccia per la rimozione delle distorsioni indesiderate, siano esse analogiche o digitali.



**Figura 6.14:** modulo DeClip 2 della Acon Digital Restoration Suite.

La Acon Digital sviluppa anche un apprezzato plugin di isolamento dei dialoghi, chiamato Extract:Dialogue. Si tratta di una versione semplificata del DeNoise, con dei preset di rimozione del rumore, in grado di isolare efficacemente e soprattutto molto rapidamente le componenti parlate del segnale audio, tramite algoritmi migliorati basati su reti neurali.<sup>101 102</sup>



**Figura 6.15:** interfaccia di Extract:Dialogue.

<sup>101</sup> Acon Digital, [www.acondigital.com](http://www.acondigital.com)

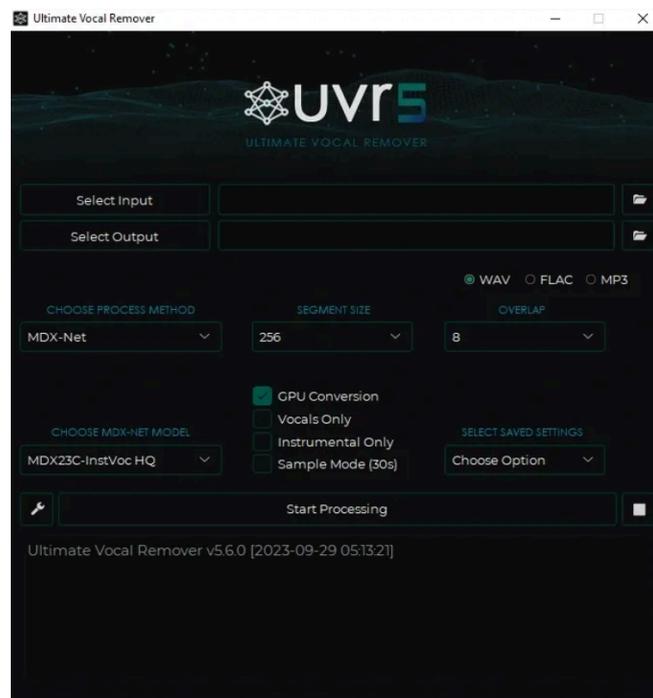
<sup>102</sup> ibidem

## 6.2 Implementazioni open-source

### 6.2.1 Ultimate Vocal Remover

Tra le soluzioni di open-source il progetto UVR (Ultimate Vocal Remover) è uno dei più interessanti.

Il sistema è progettato sulla soluzione di source separation, basata su reti neurali, presentata nel 2017 nell'articolo "Multi-Scale multi-band densenets for audio source separation".<sup>103</sup> L'approccio è basato sull'utilizzo di reti neurali convoluzionali di tipo densely connected<sup>104</sup> che permettono di realizzare modelli spettrali dei diversi strumenti con risultati molto soddisfacenti.



**Figura 6.16:** GUI di UVR5.

<sup>103</sup> Takahashi & Mitsufuji, 2017, 21-25

<sup>104</sup> Un tipo di rete neurale convoluzionale che utilizza connessioni dense tra i layer, dove tutti i layer sono connessi direttamente tra loro. Per preservare la natura feedforward, ogni layer riceve input aggiuntivi da tutti i layer precedenti e passa i propri output a tutti i layer successivi. (Papers With Code, [www.paperswithcode.com](http://www.paperswithcode.com))

UVR è arrivato alla sua quinta versione ed è disponibile in maniera gratuita su GitHub, può essere installato in standalone nel sistema ed essere utilizzato attraverso una semplice GUI.

Supporta formati WAV, FLAC e MP3 e diversi modelli di processamento dei file, tra cui Demucs e MDX/Net. Il primo è il più versatile per diversi generi musicali, il secondo offre le migliori performance di estrazione della voce senza introdurre artefatti.

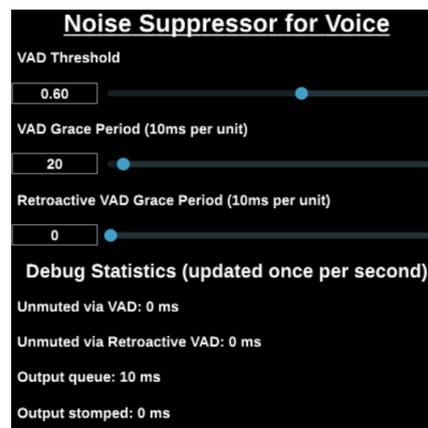
UVR offre anche una modalità *Audio Tools*, in cui sono disponibile una funzione di time stretching e pitch shifting, una di allineamento temporale, oltre che la modalità ensemble che permette di combinare i vari modelli di separazione sia in modalità automatica che manuale.<sup>105</sup>

## 6.2.2 Werman Noise Suppression

Werman Noise Suppression è un plugin open-source basato sulla libreria RNNoise, costruita per la rimozione di varie classi di rumore di fondo.

Questo plugin offre risultati promettenti ma soffre di ingenti limitazioni, sia sulla frequenza di campionamento, che sulla risoluzione in bit (rispettivamente 48 kHz, 16 bit).

Werman Noise Suppression è disponibile con una GUI minimale che permette un facile utilizzo, ma poca possibilità di regolazione.<sup>106</sup>



**Figura 6.17:** GUI di Werman Noise Suppression.

<sup>105</sup> Ultimate Vocal Remover, [www.github.com/Anjok07](https://www.github.com/Anjok07)

<sup>106</sup> Werman Noise Suppression Project, [www.github.com/werman](https://www.github.com/werman)

### 6.2.3 MPAI CAE-ARP

Nel 2023 è stato presentato lo standard ARP (Audio Recording Preservation), che costituisce un componente fondamentale dello standard CAE (Context-based Audio Enhancement) sviluppato dal MPAI (Moving Picture, Audio and Data Coding by Artificial Intelligence). Questo standard si occupa di specificare una serie di tecnologie basate su AI per disparate applicazioni audio, dalle telecomunicazioni, all'intrattenimento, passando per la post-produzione e la preservazione.

CAE-ARP è un valido strumento, utile nella digitalizzazione e nella gestione di archivi di varie dimensioni, composti da tracce sonore registrate su supporto ottico magnetico open-reel. Utilizza algoritmi di AI per l'estrazione di informazioni rilevanti, allo scopo di creare copie adatte alla preservazione e a generare metadati utili per l'archiviazione efficace di tutte le informazioni.

Lo standard è composto di diversi moduli, due per l'analisi dei file (audio e video) in grado di individuare e catalogare degradazioni sia nella traccia audio, che in una registrazione video del nastro che scorre nel registratore magnetico.

Altri moduli sono realizzati per catalogare le irregolarità del nastro magnetico e per adottare operazioni preliminari di restauro, quali compensazione dell'equalizzazione e ripristino di problemi dovuti a variazioni della velocità di scorrimento del nastro.

Pur essendo stato pubblicato come strumento open-source, lo standard non è disponibile per l'utilizzo tramite un'interfaccia e di conseguenza il suo utilizzo è poco user-friendly, anche se implementazioni di tale natura sono attualmente in fase di sviluppo.<sup>107</sup>

---

<sup>107</sup> Bosi et al., 2023, 1-8

## **7. CONCLUSIONI**

### **7.1 Esempi di intervento su “Miti e riti dell’India”**

Per avvicinarsi alla conclusione della presente trattazione, è necessario descrivere alcuni tentativi di intervento che sono stati fatti su un segmento di “Miti e riti dell’India” (della durata di circa due minuti) e che hanno funzionato da test per valutare le differenze di risultato ottenibili tramite tecniche di restauro base applicate alle due scansioni.

Il primo intervento effettuato è stato il ripristino della velocità originale e del pitch. Infatti, il film è stato originariamente girato a 20 frame per secondo, ma le scansioni sono state effettuate a una velocità di scorrimento del nastro di 24 fps, perché il Sondor non ha la possibilità di scansionare a velocità inferiori e, per rendere le scansioni perfettamente paragonabili, è stata impostata la stessa velocità anche sul Lasergraphics. La velocità è stata pareggiata con un rallentamento del 16.67% e con un conseguente abbassamento del pitch di circa 2 semitoni.

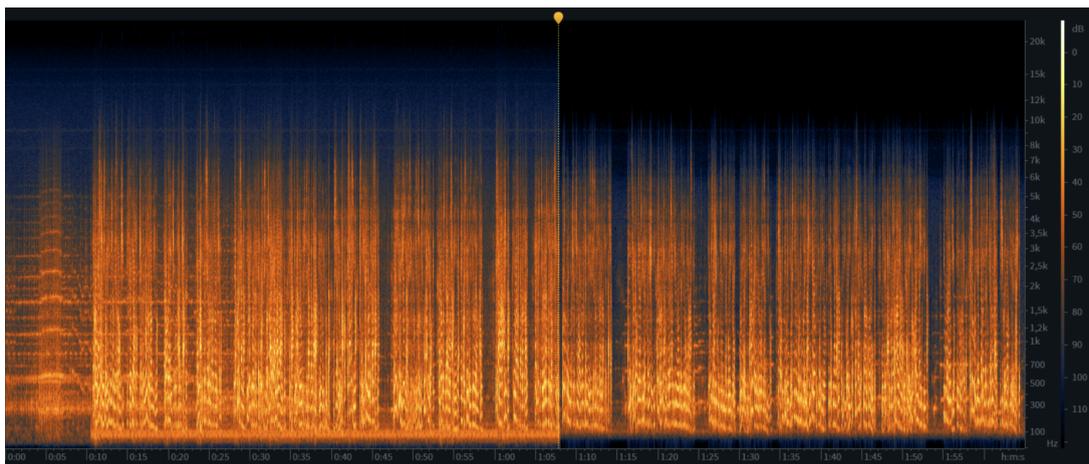
Inoltre, per testare quale tipo di risultato si possa ottenere sulla traccia magnetica messa a disposizione dalla Cineteca, sono stati realizzati alcuni interventi di rimozione dei disturbi.

Nello specifico i file sono stati processati con la seguente catena di moduli della suite Advanced di iZotope RX 10: Voice De-noise, De-click, De-hum e infine De-Clip.

Il primo modulo ha rimosso gran parte del rumore di fondo con risultati molto soddisfacenti, e, dopo aver processato i segmenti con il modulo De-click, il risultato è molto più chiaro. Si è scelto di utilizzare il modulo De-hum per rimuovere parte del ronzio in bassa frequenza che era rimasto nonostante il processamento con il Voice De-noise e, infine, si è provata ad attenuare la distorsione del segnale tramite il modulo De-clip. Quest’ultimo, nello specifico,

è progettato per rimuovere il clipping digitale e per ricostruire il segnale che ha ecceduto la soglia massima rappresentabile in conversione digitale. La distorsione caratteristica della traccia magnetica in questione, però, è una di tipo analogico (probabilmente dovuta a una registrazione non professionale), e il modulo non può fare molto per rimuoverlo.

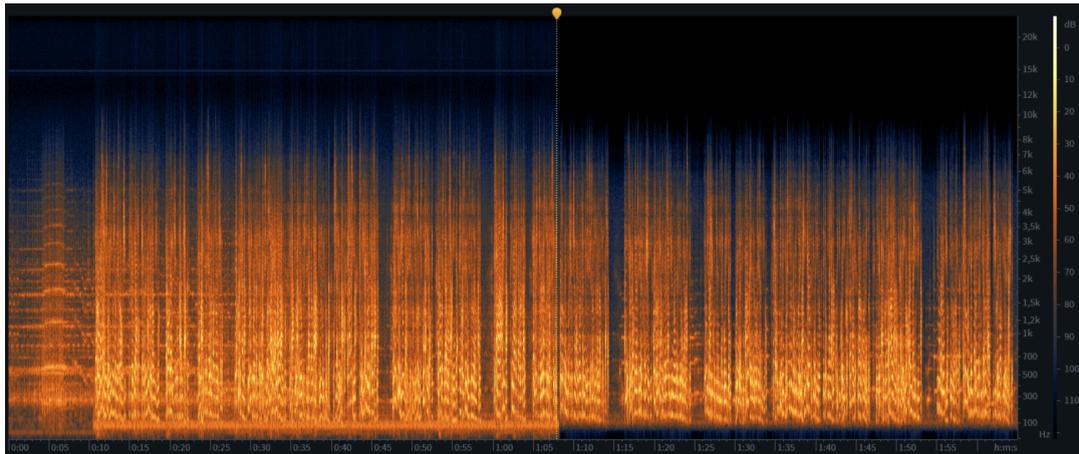
Ciò che, invece, aiuta molto nel miglioramento del suono del parlato è una buona compressione, per controllare la dinamica del segnale<sup>108</sup>, oltre che una accurata equalizzazione, per ridurre il comportamento eccessivamente medioso del nastro magnetico e rendere il parlato più chiaro e piacevole. Queste ultime sono, ovviamente, scelte estetiche opinabili, ma se portate a termine correttamente aumentano la qualità percepita del risultato finale, sempre tenendo a mente di non snaturare le caratteristiche sonore tipiche, in questo caso, del sonoro registrato su nastro magnetico.



**Figura 7.1:** spettrogramma del segmento prima e dopo gli interventi (scansione Lasergraphics).

---

<sup>108</sup> È assai probabile che nel contesto di produzione amatoriale specifico di questo caso studio, non sia stata effettuata nessuna elaborazione del segnale audio, eccetto per le operazioni di montaggio. Il segnale è, perciò, particolarmente grezzo.



**Figura 7.2:** spettrogramma del segmento prima e dopo gli interventi (scansione Sondor).

## 7.2 Osservazioni conclusive

L'analisi delle scansioni di "Miti e riti dell'India" rappresenta un caso studio difficile per testare le apparecchiature di scansione. Questo è dovuto al fatto che si tratta di un film amatoriale e, in quanto tale, non è stato realizzato nelle migliori condizioni tecniche possibili, tipiche di un contesto cinematografico industriale professionale, dove spesso gli sforzi di restauro si concentrano.

Ciò che si può effettivamente trarre, a conclusione di questa trattazione, è che le due scansioni di "Miti e riti dell'India" non presentano differenze sostanziali udibili da un punto di vista spettrale che non siano correggibili o compensabili in post-produzione. Entrambe le digitalizzazioni della colonna magnetica presa in esame possono condurre a un restauro efficace del suono. I test di restauro effettuati con la suite iZotope RX 10, portano a risultati paragonabili e qualitativamente soddisfacenti su entrambe le digitalizzazioni.

Il Lasergraphics Scanstation e il Sondor OMA/E sono macchinari sicuramente differenti fra loro, entrambi permettono risultati qualitativamente elevati e non riportano differenze consistentemente udibili che non possano essere trattate in sede di restauro e che possano pregiudicare la sua buona riuscita, almeno per quanto riguarda il caso preso in esame.

Inoltre, è fondamentale sottolineare che il risultato della scansione non può prescindere da operazioni analogiche preliminari sul supporto che siano condotte a regola d'arte. Le operazioni di pulizia e il trattamento del supporto analogico, antecedenti al restauro nel dominio digitale, sono cruciali per la sua riuscita e permettono di escludere una grande quantità di problematiche, senza ricorrere ai delicati e potenti strumenti digitali. La pulizia di film analogici può essere condotta con lavatrici ultrasoniche, che rimuovono le impurità utilizzando onde sonore ad altissima frequenza propagate attraverso un bagno chimico riscaldato in cui il supporto viene immerso. Alternativamente, esistono metodi basati su solventi chimici specifici o su sistemi meccanici chiamati PTR (Particle Transfer Roller), che permettono tramite contatto diretto di eliminare sporcizia accumulata nel tempo.<sup>109</sup>

Occorre sottolineare che, nella maggior parte dei casi, il lavaggio di tipo chimico e ultrasonico viene evitato sui supporti magnetici, dove si preferisce usare metodi di pulizia meccanici, che permettono di allontanare meglio i rischi di danneggiamento.

Nella valutazione della qualità delle scansioni e del restauro finale è necessario tenere a mente tra gli obiettivi il formato di fruizione pensato per l'opera. La prevalenza dei restauri condotti mira a riportare una data opera alla situazione di proiezione per cui era stata pensata, ovvero alla proiezione in sala. Quest'ultima ha determinate caratteristiche, a livello acustico e elettroacustico, che non possono essere ignorate. L'ascolto cinematografico medio non è un ascolto Hi-Fi, questo è un elemento non trascurabile nella fase di determinazione degli obiettivi preliminari e va sempre ricordato, anche nella valutazione delle tecnologie di scansione da utilizzare.

Infine, come è già stato discusso ampiamente, i risultati delle operazioni di restauro dipendono in maniera preponderante dalle scelte del singolo tecnico di restauro, che agisce sul segnale e sul supporto secondo regole etiche ed estetiche stabilite, ma interpretate in maniera personale, secondo la sua sensibilità e gusto. A parità di scansione, ogni singola scelta del tecnico

---

<sup>109</sup> Enticknap, 2013, 87-91

restauratore può portare a risultati qualitativamente e percettivamente molto diversi.

### **7.3 Futuri sviluppi**

Per portare avanti il discorso fatto sul confronto degli scanner, sarebbe sicuramente prioritario e interessante analizzare scansioni di tipo ottico. La mancanza di un'analisi su una scansione di sonoro su supporto diverso dal magnetico amplierebbe ulteriormente il discorso e permetterebbe ancora meglio di capire le limitazioni di una o dell'altra tecnologia di digitalizzazione.

Inoltre, un approccio statistico, ovvero condotto su molteplici opere differenti su supporti diversi, permetterebbe maggiore chiarezza sulla tecnologia di scansione migliore da fare, operando valutazioni progetto per progetto.

Un altro possibile futuro sviluppo di questi argomenti riguarda la possibilità di condurre vari tipi di test di ascolto soggettivi. Questo potrebbe permettere di testare statisticamente come cambia la percezione di un segnale restaurato, in base a chi compie le scelte tecniche e artistiche di intervento e anche alla tipologia di scansione, comprendendo come questa possa correlare con il risultato finale del restauro.

Per quanto riguarda le tecnologie basate su reti neurali, gli sviluppi possibili sono tantissimi. Come emerso dalle analisi nel capitolo 5, la maggioranza delle soluzioni presentate non è pensata a fini di restauro, ma perfettamente adattabile anche al contesto della preservazione. Con la giusta disponibilità di dataset si potrebbero studiare modelli appositamente pensati per il rilevamento dei disturbi tipici dei supporti analogici (portando avanti le possibilità esplorate dal modello MPAI CAE-ARP, paragrafo 6.2.3) e per la source separation adattata al contesto cinematografico (piuttosto che musicale, dove invece fino ad oggi si è principalmente concentrata).

Questa famiglia di tecnologie, vista come strumento a supporto, e non a sostituzione, dell'esperienza e della sensibilità filologica del tecnico restauratore, rappresenta una delle più interessanti aperture tecnologiche da osservare e in cui sperare per ottenere risultati sempre migliori.



# Bibliografia e sitografia

- Acon Digital. (n.d.). *Extract:Dialogue*. Acon Digital. Retrieved October 21, 2024, from <https://acondigital.com/products/extract-dialogue>
- Acon Digital. (n.d.). *Restoration Suite*. Acon Digital. Retrieved October 21, 2024, from <https://acondigital.com/products/restoration-suite>
- ARRI. (n.d.). *ARRISCAN XT*. ARRI. Retrieved September 16, 2024, from <https://www.arri.com/en/applications/archive-technologies/arriscan-xt>
- Bosi, M., Canazza, S., Russo, A., Pretto, N., & Chiariglione, L. (2023). An MPAI/IEEE International Standard for Audio: Overview of CAE Audio Recording Preservation (ARP) Technology. *AES Intl. Conference on Audio Archiving, Preservation & Restoration*.
- Boston, G., & International Association of Sound and Audiovisual Archives. Technical Committee. (2014). *Handling and Storage of Audio and Video Carriers: Technical Committee Standards, Recommended Practices, and Strategies* (D. Schüller & A. Häfner, Eds.; Traduzione italiana, 2016. ed.). International Association of Sound and Audiovisual Archives. <https://www.iasa-web.org/book/export/html/3930>
- Canazza, S. (2013). *Restauro materiale audio*. Università di Padova, Centro di Sonologia Computazionale. <https://www.dei.unipd.it/~musica/IM/Restauro.pdf>
- Canazza, S., Cosi, P., Russo, A., & Rodà, A. (2018). Audio Documents Restoration as a Documentary Source in the Linguistic Research: Comparison of Instruments. In *Speech Audio Archives: Preservation, Restoration, Annotation Aimed at Supporting the Linguistic Analysis : International Workshop (Rome, 18-19 May 2017)*. Accademia Naz. dei Lincei.
- Cano, E., FitzGerald, D., Liutkus, A., Plumbey, M. D., & Stöter, F. -R. (2019). Musical Source Separation: An Introduction. *IEEE Signal Processing Magazine*, 36, 31-40. 10.1109/MSP.2018.2874719

- Carlos, M. S. (2013). A comparison of scanning technologies for archival motion picture film.  
[https://www.academia.edu/5918000/A\\_COMPARISON\\_OF\\_SCANNING\\_TECHNOLOGIES\\_FOR\\_ARCHIVAL\\_MOTION\\_PICTURE\\_FILM?source=swp\\_share](https://www.academia.edu/5918000/A_COMPARISON_OF_SCANNING_TECHNOLOGIES_FOR_ARCHIVAL_MOTION_PICTURE_FILM?source=swp_share)
- Corelli, S. (2019). *Radiografare il suono con iZotope RX*. Lambda.
- Corelli, S., Felici, F., & Martinelli, G. (2006). *Elementi di cinematografia sonora*. Lambda.
- Corelli, S., Mainetti, S., & Martinelli, G. (2021). *Dialoghi, musica, effetti: il suono nell'audiovisivo*. Lambda.
- Daniel, P., & Weber, R. (1997). Psychoacoustical Roughness: Implementation of an Optimized Model. *ACUSTICA - Acta Acustica, Vol. 83*, 113-123.
- Enticknap, L. (2013). *Film Restoration: The Culture and Science of Audiovisual Heritage*. Palgrave Macmillan.
- Godsill, S. J., & Rayner, P. J.W. (1998). *Digital audio restoration: a statistical model based approach*. Springer London.
- International Federation of Film Archives. (n.d.). *Glossary of Technical Terms*. Retrieved September 2, 2024, from  
<https://www.fiafnet.org/pages/E-Resources/Technical-Terms-Full-List.html>
- iZotope, Inc. (n.d.). *RX 11 Features*. iZotope. Retrieved October 18, 2024, from  
<https://www.izotope.com/en/products/rx/features.html>
- LaserGraphics, Inc. (n.d.). *ScanStation 8mm to 70mm Motion Picture Film Scanner*. LaserGraphics. Retrieved September 2, 2024, from  
<https://lasergraphics.com/scanstation.html>
- Mohammed, S. J., & Radhika, N. (2022). Audio Denoising Using Deep Neural Networks. In D. J. Hemanth, D. Pelusi, & C. Vuppapapati (Eds.), *Intelligent Data Communication Technologies and Internet of Things: Proceedings of ICICI 2021*. Springer Nature Singapore.

- National Archives. (2023, October 25). *Motion Picture Film Guidance: Identifying Motion Picture Film Formats*. National Archives. Retrieved September 9, 2024, from <https://www.archives.gov/preservation/formats/motion-picture-film-identify-formats.html>
- National Film Preservation Foundation. (n.d.). *Why Preserve Film?* Retrieved September 16, 2024, from <https://www.filmpreservation.org/preservation-basics/>
- Naylor, P. A., & Gaubitch, N. D. (Eds.). (2010). *Speech Dereverberation*. Springer.
- Nugraha, A. A., Liutkus, A., & Vincent, E. (2016). Multichannel Audio Source Separation With Deep Neural Networks. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24. 10.1109/TASLP.2016.2580946
- Papers With Code. (n.d.). *DenseNet Explained*. Papers With Code. Retrieved October 23, 2024, from <https://paperswithcode.com/method/densenet>
- Plutino, A. (2020). *Tecniche di restauro cinematografico. Metodi e pratiche tra analogico e digitale*. Audino.
- Preferred Media. (n.d.). *Why you should digitise video, tape and film assets now*. Retrieved September 16, 2024, from <https://preferredmedia.com.au/our-resources/digitise-video-and-film/>
- Pretto, N., Russo, A., Bressan, F., Burini, V., Rodà, A., & Canazza, S. (2020). Active preservation of analogue audio documents: A summary of the last seven years of digitization at CSC. In *Proceedings of the 17th Sound and Music Computing Conference (Torino, Italy, 24th-26th June 2020)*. Axa.
- Rabiner, L. R. (1993). *Fundamentals of Speech Recognition*. PTR Prentice Hall.
- Russo, A., Canazza, S., & Rodà, A. (2016). Il caso dell'archivio storico del Teatro Regio di Parma: applicazione di una metodologia scientifica per la conservazione attiva di documenti sonori. In V. Burini (Ed.), *Extending Interactivity. Atti Del 26o CIM Colloquio Di Informatica Musicale*. Dip. Arti e Design Industriale.

- Russo, A., Pretto, N., Roda, A., & Canazza, S. (2019). *L'informatica per la gestione e la conservazione di informazioni acustiche (musica e voce)*. aisv.it.  
[https://www.aisv.it/StudiAISV/2019/vol\\_6/027\\_RussoPrettoEtAlII.pdf](https://www.aisv.it/StudiAISV/2019/vol_6/027_RussoPrettoEtAlII.pdf)
- Sondor & Digital Film Technology GmbH. (n.d.). *Sondor OMA-E Handbook* [User manual].
- Sondor & Digital Film Technology GmbH. (n.d.). *Resonances. Optical soundtrack scanning system. User Manual 3.0.0*.
- Sondor Versa & Resonances*. (n.d.). Dft-Film.com. Retrieved September 8, 2024, from <https://dft-film.com/products/versa-resonances-2/>
- Steinberg Media Technologies GmbH. (n.d.). *SpectralLayers: Advanced Spectral Audio Editor*. Steinberg. Retrieved October 30, 2024, from <https://www.steinberg.net/spectralayers>
- Stewart, I., & iZotope, Inc. (n.d.). *What Are LUFS? The Complete Guide*. izotope.com. Retrieved November 12, 2024, from <https://www.izotope.com/en/learn/what-are-lufs.html>
- Stöter, F.-R., Uhlich, S., Liutkus, A., & Mitsufuji, Y. (2019). Open-Unmix - A Reference Implementation for Music Source Separation. *Journal of Open Source Software, Open Journals*, 10.21105/joss.01667
- Supertone Inc. (n.d.). *Supertone Clear*. Supertone - Product. Retrieved October 21, 2024, from <https://product.supertone.ai/clear>
- Takahashi, N., & Mitsufuji, Y. (2017). Multi-Scale multi-band densenets for audio source separation. *2017 IEEE Workshop on Applications of Signal Processing to Audio and Acoustics (WASPAA)*. 10.1109/WASPAA.2017.8169987
- Uccello, P. (1982). *Cinema: tecnica e linguaggio*. Edizioni paoline.
- Ultimate Vocal Remover. (n.d.). *Anjok07/ultimatevocalremovergui: GUI for a Vocal Remover that uses Deep Neural Networks*. GitHub. Retrieved October 23, 2024, from <https://github.com/Anjok07/ultimatevocalremovergui>

- Universitat Pompeu Fabra, Music Technology Group. (n.d.). Homepage –  
Essentia 2.1-beta6-dev documentation. Retrieved October 13, 2024, from  
<https://essentia.upf.edu/index.html>
- Venturini, S. (Ed.). (2006). *Il restauro cinematografico: principi, teorie, metodi*.  
Campanotto.
- Walsh, D., & FIAF Technical Commission. (2021). Choosing a Film Scanner.  
*Resources of the Technical Commission (TC)*.  
[https://www.fiafnet.org/images/tinyUpload/2021/07/Choosing\\_a\\_film\\_scanner\\_2021.pdf](https://www.fiafnet.org/images/tinyUpload/2021/07/Choosing_a_film_scanner_2021.pdf)
- Wander, T., & Curteis, M. (n.d.). marconiintelevision [licensed for  
non-commercial use only] / Marconi in Television. Retrieved September  
16, 2024, from <http://marconiintelevision.pbworks.com/>
- Wang, Z.-Q., & Wang, D. (2020). Deep Learning Based Target Cancellation for  
Speech Dereverberation. *IEEE/ACM Transactions on Audio, Speech, and  
Language Processing*, vol. 28. 10.1109/TASLP.2020.2975902
- Waves Audio Ltd. (n.d.). *Clarity Vx*. waves.com.  
<https://www.waves.com/plugins/clarity-vx>
- Werman Noise Suppression Project. (n.d.).  
*werman/noise-suppression-for-voice: Noise suppression plugin based  
on Xiph's RNNoise*. GitHub. Retrieved October 30, 2024, from  
<https://github.com/werman/noise-suppression-for-voice>
- Wichern, G. (2017, aprile 24). *What the Machine Learning in RX 6 Advanced  
Means for the Future of Audio Repair Technology*. izotope.com.  
[https://www.izotope.com/en/learn/what-the-machine-learning-in-rx-6-advanced-means-for-the-future-of-audio-repair-technology.html?srsId=AfmBOoo2VsdF-t7jdhWt7GIPgHtXNLT6ng\\_OyQkmZoWg8FjMBKTYvliC](https://www.izotope.com/en/learn/what-the-machine-learning-in-rx-6-advanced-means-for-the-future-of-audio-repair-technology.html?srsId=AfmBOoo2VsdF-t7jdhWt7GIPgHtXNLT6ng_OyQkmZoWg8FjMBKTYvliC)
- Wolff, D., Mignot, R., Roebel, A., Analysis-Synthesis Team, UMR 9912 STMS, &  
IRCAM, CNRS, Sorbonne Université, Paris, France. (2022, Feb 11). Audio  
defect detection in music with Deep Networks. *Proceedings of the 22nd  
International Society for Music Information Retrieval Conference*.

# Elenco delle figure

1. Introduzione al restauro cinematografico
  - 1.1. Pista ottica monofonica a densità variabile. (National Archives, 2023)
  - 1.2. Pista ottica monofonica ad area variabile. (National Archives, 2023)
  - 1.3. Pellicola in avanzato stato di decadimento. (National Film Preservation Foundation, [www.filmpreservation.org](http://www.filmpreservation.org))
  - 1.4. Un frame soggetto a sbiadimento delle tinte. (National Film Preservation Foundation, [www.filmpreservation.org](http://www.filmpreservation.org))
  - 1.5. Nastro magnetico gravemente colpito da sindrome acetica. (Preferred Media, [www.preferredmedia.com](http://www.preferredmedia.com))
  - 1.6. Un frame di "Miti e riti dell-India". (Cineteca del Museo Nazionale del Cinema di Torino, Fondo Petruccelli)
  
2. Tecnologie e metodi di scansione
  - 2.1. Telecine modello Marconi B3410. (Wander & Curteis, [www.marconitelevision.pbworks.com](http://www.marconitelevision.pbworks.com))
  - 2.2. Schema di film gate con aggancio regolabile della pellicola. (Carlos, 2013, 20)
  - 2.3. Wet/gate ARRISCAN, i piccoli fori servono a spargere il liquido. (ARRI, [www.ari.com](http://www.ari.com))
  - 2.4. Scansione dry/gate (a sinistra) e wet/gate (a destra). (Carlos, 2013, 21)
  - 2.5. Lasergraphics Scanstation caricato con pellicola. (LaserGraphics, Inc., [www.lasergraphics.com](http://www.lasergraphics.com))
  - 2.6. Sondor OMA/E. (Sondor & Digital Film Technology GmbH, n.d.)

- 2.7. Sistema di trasporto dell'OMA/E. (Sondor & Digital Film Technology GmbH, n.d.)
  - 2.8. Acquisizione ottica Resonances. (Sondor & Digital Film Technology GmbH, n.d.)
3. Analisi dei segnali digitalizzati
- 3.1. Forma d'onda di un segnale audio degradato da clicks. (Godsill & Rayner, 1998, 129)
  - 3.2. Spettrogramma e forma d'onda di un pop. Scansione: MNC.
  - 3.3. Spettrogramma e forma d'onda di un click. Scansione: MNC.
  - 3.4. Spettrogramma e forma d'onda di un hum isolato. Scansione: MNC.
  - 3.5. Spettrogramma relativo a ovattatura transitoria. Scansione: MNC.
  - 3.6. Spettrogramma dell'effetto pitch-down transitorio. Scansione: MNC.
  - 3.7. Spettrogramma della scansione effettuata con Lasergraphics (MNC).
  - 3.8. Spettrogramma scansione effettuata con Sondor (Cineteca di Bologna).
  - 3.9. Statistiche scansione Lasergraphics.
  - 3.10. Statistiche scansione Sondor.
  - 3.11. Dettaglio dello spettrogramma con picco intorno ai 18 kHz. (Sondor)
  - 3.12. Dettagli di picchi nel rumore di fondo della scansione Lasergraphics
  - 3.13. Confronto spettrale di un segmento di rumore di fondo.
  - 3.14. Descrittori utili del primo segmento di rumore di fondo.

- 3.15. Distanza spettrale globale e per bande d'ottava del rumore di fondo.
- 3.16. Confronto spettrale di un secondo segmento di rumore di fondo.
- 3.17. Distanze spettrali di un secondo segmento di rumore di fondo.
- 3.18. Descrittori utili, secondo segmento di rumore di fondo.
- 3.19. Confronto spettrale di una porzione mista, voce e musica.
- 3.20. Descrittori utili, porzione mista.
- 3.21. Distanze spettrali, segmento misto.
- 3.22. Confronto spettrale di un altro segmento misto.
- 3.23. Distanze spettrali, secondo segmento misto.
- 3.24. Descrittori utili, secondo segmento generico.
- 3.25. Confronto spettrale del ronzio isolato dal contenuto vocale.
- 3.26. Tabella delle distanze spettrali tra i ronzii scansionati.
- 3.27. Tabella di descrittori utili per il ronzio.
- 3.28. Confronto spettrale di un disturbo impulsivo di tipo click.
- 3.29. Confronto spettrale di un altro click.
- 3.30. Confronto spettrale di un pop.
- 3.31. Confronto ulteriore di un altro segmento contenente pop.
- 3.32. Confronto spettrale di una porzione di parlato danneggiata da un pop.
- 3.33. Confronto spettrale di una porzione di parlato priva di degradazioni.
- 3.34. Andamento del centroide spettrale nelle due scansioni.
- 3.35. Andamento della roughness nelle due scansioni.

#### 4. Operazioni digitali di restauro

- 4.1. Rilevamento click con modello AR. (Godsill & Rayner, 1998, 129)
- 4.2. Segnale rumoroso (sopra), componenti residue della riduzione (sotto). (Godsill & Rayner, 1998, 146)
5. Contributo delle reti neurali alle tecniche di restauro  
-
6. Panoramica dei software
  - 6.1. Modulo Spectral De-noise di Izotope RX 11. (iZotope, Inc., [www.izotope.com](http://www.izotope.com))
  - 6.2. Modulo De-reverb di Izotope RX 11. (iZotope, Inc., [www.izotope.com](http://www.izotope.com))
  - 6.3. Modulo Music Rebalance di Izotope RX 11. (iZotope, Inc., [www.izotope.com](http://www.izotope.com))
  - 6.4. Modulo Repair Assistant di Izotope RX 11. (iZotope, Inc., [www.izotope.com](http://www.izotope.com))
  - 6.5. Interfaccia dell'editor spettrale di Steinberg SpectraLayers 11. (Steinberg Media Technologies GmbH, [www.steinberg.net](http://www.steinberg.net))
  - 6.6. Waves Clarity Vx. (Waves Audio Ltd., [www.waves.com](http://www.waves.com))
  - 6.7. Waves Clarity Vx PRO. (Waves Audio Ltd., [www.waves.com](http://www.waves.com))
  - 6.8. Waves Clarity Vx DeReverb. (Waves Audio Ltd., [www.waves.com](http://www.waves.com))
  - 6.9. Waves Clarity Vx DeReverb PRO. (Waves Audio Ltd., [www.waves.com](http://www.waves.com))
  - 6.10. Supertone CLEAR. (Supertone Inc., [www.supertone.ai](http://www.supertone.ai))
  - 6.11. Modulo DeNoise 2 della Acon Digital Restoration Suite. (Acon Digital, [www.acondigital.com](http://www.acondigital.com))
  - 6.12. Modulo DeHum 2 della Acon Digital Restoration Suite. (Acon Digital, [www.acondigital.com](http://www.acondigital.com))
  - 6.13. Modulo DeClick 2 della Acon Digital Restoration Suite. (Acon Digital, [www.acondigital.com](http://www.acondigital.com))

- 6.14. Modulo DeClip 2 della Acon Digital Restoration Suite. (Acon Digital, [www.acondigital.com](http://www.acondigital.com))
- 6.15. Interfaccia di Extract:Dialogue. (Acon Digital, [www.acondigital.com](http://www.acondigital.com))
- 6.16. GUI di UVR5. (Ultimate Vocal Remover, [www.github.com/Anjok07](http://www.github.com/Anjok07))
- 6.17. GUI di Werman Noise Suppression. (Werman Noise Suppression Project, [www.github.com/werman](http://www.github.com/werman))

## 7. Conclusioni

- 7.1. Spettrogramma del segmento prima e dopo gli interventi (scansione Lasergraphics).
- 7.2. spettrogramma del segmento prima e dopo gli interventi (scansione Sondor).

# Ringraziamenti

Questa tesi di Laurea è stata possibile grazie alla guida di Vito Martinelli e Antonio Servetti, a cui va un sincero e profondo ringraziamento.

Grazie alla Cineteca del Museo Nazionale del Cinema, e in particolare a Gabriele Angelo Perrone, per aver supportato il mio lavoro.

Grazie alla professoressa Louena Shtrepi per i consigli nella fase finale della stesura.

Grazie mamma e papà, perché, con amore, mi avete insegnato la curiosità e il pensiero critico, senza mai smettere di credere in me, neanche quando avevo paura di non potercela fare.

Grazie Natalia per essere un punto di appoggio imprescindibile, una confidente e una sorella affettuosa.

Grazie Francesco perché mi hai sempre fatto sentire l'affetto e la guida di un fratello maggiore, anche se sei distante.

Grazie Andrea, sei l'amico più prezioso che ho trovato in questo percorso.

Grazie Gabriele, Davide, Luca, Giovanni e Pietro perché siete stati, e rimarrete, compagni straordinari e instancabili di avventure. Siete dei fratelli che non mi hanno mai abbandonato.

Grazie ai miei coinquilini, i tre Paoli, per la pazienza, le discussioni, le serate insieme. Nella casa che condividiamo lascerò per sempre una parte di me.

Grazie Chiara, perché nei mesi di stesura di questa tesi sei stata fondamentale più di chiunque altro. Tu per me sei luce, sei primavera.

Un ultimo ringraziamento va a Sergio Leone, Stanley Kubrick, Francis Ford Coppola, Quentin Tarantino, Denis Villeneuve, Damien Chazelle, Brian De Palma, Steven Spielberg e Ridley Scott. È grazie al loro Cinema se questa passione, che mi arde dentro, mi ha portato fino a qui.