

Quelle voci poco fa: l'intelligenza artificiale a contrastare l'eclisse delle memorie sonore

Sergio Canazza¹, Carlo Fantozzi¹, Niccolò Pretto¹, Antonio Rodà¹,
Anthony Chmiel², Emery Schubert²

1 Centro di Sonologia Computazionale
DEI, Università di Padova

2 Empirical Musicology Laboratory
University of New South Wales

{canazza,fantozzi,prettoni,roda}@dei.unipd.it
{anthony.chmiel,e.schubert}@unsw.edu.au

Abstract

Le informazioni acustiche – anche a causa della grande varietà delle tecnologie impiegate, e nonostante il loro continuo perfezionamento –, a differenza di quelle grafiche, hanno dimostrato di essere estremamente labili, soggette sia a degradazione, sia a possibilità di perdita irreparabile in tempi relativamente brevi. Conservare i supporti originali e, contemporaneamente, l'equipaggiamento necessario alla loro riproduzione è senza speranza: la comunità archivistica internazionale ha introdotto in questo senso lo slogan difendere il contenuto, non il supporto. La conservazione diventa attiva, accogliendo l'idea che è necessario digitalizzare le informazioni acustiche. Procedura complessa, che deve essere svolta da gruppi di ricerca interdisciplinari, possibilmente avvalendosi di strumenti sviluppati nel campo dell'intelligenza artificiale.

1 Introduzione

Il *Centro di Sonologia Computazionale* (CSC), laboratorio di *Sound and Music Computing* del Dipartimento di Ingegneria dell'Informazione dell'Università di Padova, è uno dei riferimenti internazionali nello studio e nella sperimentazione di nuovi strumenti informatici per la tutela e la valorizzazione di documenti audio storici [Canazza e Vidolin, 2001]. Questo campo di ricerca, scientificamente e culturalmente particolarmente complesso, ha come obiettivo quello di produrre un archivio digitale che soddisfi nel lungo termine requisiti di *affidabilità* e *autorevolezza*, ed è intrinsecamente interdisciplinare, coinvolgendo discipline che vanno dall'informatica alla chimica dei materiali, dall'archivistica, alle scienze storiche e alla psicologia della percezione.

L'innovativa metodologia definita in questi anni al CSC coinvolge sia la conservazione attiva [Bressan e Canazza, 2013], sia l'accesso e la fruizione [Canazza *et al.*, 2015; Fantozzi *et al.*, 2017] dei documenti sonori storici, con particolare attenzione ai nastri magnetici analogici ed è stata testata utilizzando registrazioni storiche di musica elettroacustica, durante diversi progetti internazionali relativi alla conservazione e al restauro di raccolte di documenti sonori di alcuni tra i più importanti archivi audio europei (tra gli altri: Fondazio-

ne Paul Sacher, Basilea; Fondazione Arena di Verona, Italia; Archivio storico del Teatro Regio di Parma, Italia; Archivio Luigi Nono, Italia). Dopo aver eseguito con cura l'eventuale restauro del supporto sonoro analogico (che il CSC opera grazie alla collaborazione con il laboratorio Chimico di Roberta Bertani, presso il Dip. di Ingegneria Industriale), è necessario trasferire nel dominio digitale le informazioni acustiche, assieme ai metadati, del documento audio originale. Le procedure di digitalizzazione tradizionali richiedono l'intervento di tecnici esperti, per: (1) documentare in maniera adeguata le caratteristiche del supporto di memorizzazione originale (presenza di danni o di annotazioni, estrazione di metadati) (2) valutare in maniera adeguata la configurazione del sistema di lettura (configurazione delle tracce, velocità di scorrimento, curva di equalizzazione). A fronte di un patrimonio culturale enorme (l'UNESCO nel 2001 stimava che esistessero 50 milioni di ore di registrazione non digitalizzate [Schüller, 2001]) e nel frattempo sono stati scoperti molti altri archivi di documenti analogici in copia unica), la necessità di un intervento decisionale umano nella catena di digitalizzazione rende difficile attuare pratiche di digitalizzazione massive, col rischio di incorrere in errori che potrebbero compromettere il risultato finale. Inoltre, anche l'analisi approfondita di una grande quantità di registrazioni da parte di uno studioso può essere ugualmente difficoltosa e a rischio di errori. Con l'obiettivo di aiutare musicologi e tecnici audio nel loro lavoro e aprire nuove possibilità di analisi nell'ambito della *computational musicology* e del *music information retrieval*, negli ultimi anni il CSC ha iniziato a sperimentare l'utilizzo di strumenti sviluppati nel campo dell'intelligenza artificiale per supportare la digitalizzazione e l'analisi dei documenti sonori storici e minimizzare la soggettività e gli errori che caratterizzano questi processi. Di seguito si presentano due esempi in cui tecniche basate su reti neurali sono utilizzate per il riconoscimento automatico di discontinuità nel nastro magnetico audio analogico (danni, giunte, discontinuità) e algoritmi di *machine learning* per il riconoscimento automatico di alcune manipolazioni del segnale sonoro.

2 Riconoscimento automatico di eventi su nastro

Il risultato della digitalizzazione di una registrazione su nastro magnetico è una copia conservativa digitale che deve poter

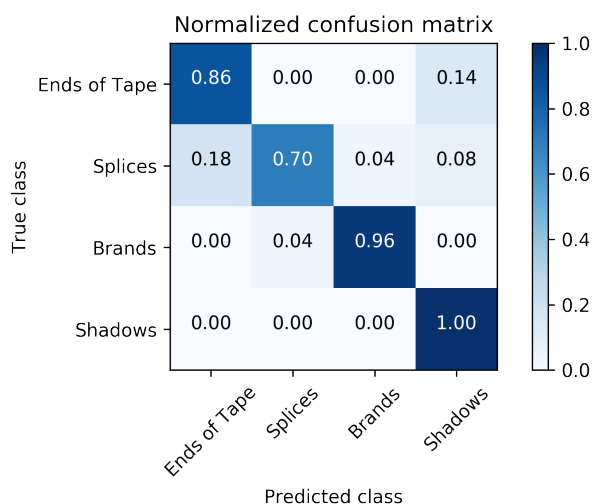


Figura 1: Matrice di confusione ottenuta dalla classificazione di discontinuità nel dataset a 7.5 ips.

sostituire il documento originale per tutte le future operazioni di accesso ai contenuti. La copia conservativa in questo senso comprende l'informazione primaria (il segnale audio), un insieme completo di metadati, e le informazioni contestuali, ovvero tutti quei dati non relativi all'informazione primaria, ma che risultano necessari per mantenere l'integrità del documento e il suo contesto. Il segnale video del nastro in scorrimento diventa l'ingresso degli strumenti per il riconoscimento automatico attualmente in fase di sviluppo presso il laboratorio, basati su tecniche di *computer vision* e su reti neurali. Questo lavoro mira a identificare e classificare le principali alterazioni nel nastro magnetico individuabili nel video (definite con il termine discontinuità). Le principali classi sono: giunte, l'unione di due pezzi di nastro tramite l'utilizzo di un apposito nastro adesivo, *corrottele*, *segnì*, *marchi*, *ombre*, *increspature*, *sporcizia*, *frame* che indicano la fine del nastro.

Per validare questi strumenti, sono stati effettuati esperimenti su diversi dataset composti da copie conservative di nastri registrati a velocità differenti (7.5 e 15 ips). La Figura 1 mostra alcuni dei risultati descritti in dettaglio in [Pretto et al., 2018], che confermano l'efficacia di questi strumenti.

3 Riconoscimento automatico elaborazioni

Il contenuto acustico del documento sonoro può fornire molte informazioni relative sia al supporto, sia alle manipolazioni subite dal segnale audio. La registrazione su nastro magnetico richiede l'utilizzo di una curva, chiamata di pre-enfasi, che deve essere compensata con una curva inversa in fase di lettura (post-enfasi). Se questa curva non è corretta, a causa di una impostazione errata del dispositivo di riproduzione, il contenuto sonoro risulta non conforme all'originale, in pratica si crea un *falso storico*, ossia, filologicamente parlando, un testimone non attendibile. Utilizzando campioni di rumore di fondo incisi nel nastro è possibile risalire alla curva di pre-enfasi, come spiegato in [Pretto et al., 2018]. Gli esperimenti proposti si sono avvalsi di algoritmi di *clustering* e classificazione su campioni composti da 13 *Mel-Frequency Cepstral Coefficients* (MFCCs). La Tabella 1 sintetizza le prestazioni

Alg.	Type	Accur.	Recall	Specif.	Prec.
DT	Medium	0.83	0.94	0.70	0.77
SVM	Gaussian	0.80	0.89	0.70	0.76

Tabella 1: Prestazioni degli algoritmi più adatti al riconoscimento delle curve di pre-enfasi.

degli algoritmi che si sono rivelati più idonei per il riconoscimento delle curve di post-enfasi. Gli ottimi risultati ottenuti incoraggiano lo sviluppo di strumenti automatici per sollevare gli addetti alla digitalizzazione dalle scelte soggettive che caratterizzano il processo.

4 Conclusioni

Esiste un enorme patrimonio di registrazioni acustiche su nastro magnetico, spesso in copia unica e di cui anche in questo istante è in corso la perdita inarrestabile: milioni di ore di registrazione di musica, suoni, voci, testimonianze di vita passata. Un'erosione della memoria collettiva dovuta all'intrinseca instabilità fisica e chimica dei supporti sonori, che si traduce in un'aspettativa di vita molto breve (qualche decennio), in netto contrasto con quella di altri beni culturali, quali i dipinti o le sculture, il cui degrado è misurabile in secoli o millenni. Presso il CSC si è lavorato molto per la soluzione dei problemi relativi alla conservazione di questi documenti, utilizzando strumenti maturati nel campo dell'intelligenza artificiale. Si confida che gli enti preposti alla salvaguardia dei beni artistici e culturali del nostro Paese dedichino attenzione e risorse adeguate alla conservazione e al restauro dell'immenso patrimonio delle informazioni di tipo acustico.

Riferimenti bibliografici

- [Bressan e Canazza, 2013] F. Bressan e S. Canazza. A systemic approach to the preservation of audio documents: Methodology and software tools. *Journal of Electrical and Computer Engineering*, 2013:21 pages, 2013.
- [Canazza et al., 2015] Sergio Canazza, Carlo Fantozzi, e Niccolò Pretto. Accessing tape music documents on mobile devices. *ACM Trans. on Multimedia Comput., Commun. Appl.*, 12(1s):20, 2015.
- [Canazza e Vidolin, 2001] S. Canazza e A. Vidolin. Introduction: Preserving electroacoustic music. *Journal of New Music Research*, 30(4):289–293, 2001.
- [Fantozzi et al., 2017] Carlo Fantozzi, Federica Bressan, Niccolò Pretto, e Sergio Canazza. Tape music archives: from preservation to access. *International Journal on Digital Libraries*, 2017.
- [Pretto et al., 2018] Niccolò Pretto, Carlo Fantozzi, Edoardo Micheloni, Valentina Burini, e Sergio Canazza. Computing methodologies supporting the preservation of electroacoustic music from analog magnetic tape. *Computer Music Journal*, 42(4), 2018.
- [Schüller, 2001] Dietrich Schüller. Preserving the facts for the future: Principles and practices for the transfer of analog audio documents into the digital domain. *Journal of Audio Engineering Society*, 49(7-8):618–621, 2001.