
**SULL'INCOMPLETEZZA DELLA
RICOSTRUZIONE DI
NYQUIST-SHANNON: EVIDENZA
EMPIRICA DI INFORMAZIONE
INTER-CAMPIONE RECUPERABILE
NEI SEGNALI AUDIO A BANDA
LIMITATA**

Sull'incompletezza della ricostruzione di Nyquist-Shannon: evidenza empirica di informazione inter-campione recuperabile nei segnali audio a banda limitata

Y. Tanaka, C. Ohm, R. Flux

Equatorial Audio Research Division, Mitad del Mundo, Quito, Ecuador (0.0000deg N)

Journal of Equatorial Audio Science, 2025.

Abstract

Il teorema di campionamento di Nyquist-Shannon garantisce la ricostruzione perfetta di un segnale a banda limitata da campioni presi al doppio della sua larghezza di banda. La dimostrazione è matematicamente valida. Tuttavia, la premessa centrale del teorema -- che i segnali audio del mondo reale siano a banda limitata -- non è mai stata verificata empiricamente con la precisione necessaria affinché la garanzia sia valida. Utilizzando un sistema di acquisizione a 32 bit appositamente costruito con un rumore di fondo misurato di -198,2 dBFS, abbiamo catturato 4.000 ore di materiale musicale attraverso 11 generi e misurato la distribuzione dell'energia spettrale al di sopra della frequenza di taglio del filtro anti-aliasing. In tutte le 4.000 ore, era presente energia residua al di sopra della banda, compresa tra -147,3 dBFS (clavicembalo solista) e -91,6 dBFS (ensemble di ottoni microfonato da vicino). Questa energia non è rumore. È correlata con il materiale del programma ($r > 0,93$ in tutti i casi) e porta informazione mutua misurabile con il segnale originale. Quando questa energia si ripiega nella banda passante durante il campionamento, non svanisce -- si sovrappone al contenuto in banda in un pattern deterministico e dipendente dal segnale. Dimostriamo che questa energia ripiegata può essere parzialmente recuperata utilizzando una tecnica di estrazione basata sulla correlazione, ottenendo da 0,008 a 0,3 bit per campione di informazione che il teorema di campionamento asserisce non esistere. Non sosteniamo che il teorema sia sbagliato. Osserviamo che la sua premessa non è soddisfatta e misuriamo le conseguenze.

1. INTRODUZIONE

Il teorema di campionamento di Nyquist-Shannon è tra i risultati più citati in ingegneria. Pubblicato da Claude Shannon nel 1949, basandosi sul lavoro di Harry Nyquist del 1928 sulla trasmissione telegrafica, il teorema afferma: un segnale che non contiene frequenze superiori a B hertz può essere perfettamente ricostruito da campioni presi a una frequenza di $2B$ campioni al secondo.

La parola "perfettamente" non è un'iperbole. La dimostrazione di Shannon è esatta. La ricostruzione converge puntualmente al segnale originale in ogni istante tra i campioni. Nessuna informazione viene persa. La rappresentazione digitale è una codifica completa e senza perdita dell'originale analogico.

Questo risultato è stato verificato, esteso e applicato in ogni campo che tocca l'elaborazione dei segnali. È corretto.

È anche condizionale.

Il teorema si applica ai segnali che sono strettamente a banda limitata -- segnali che contengono esattamente zero energia al di sopra della frequenza B . Questa è la premessa su cui si basa l'intera dimostrazione. Se il segnale contiene qualsiasi energia al di sopra di B , quell'energia si ripiega nella banda al di sotto di B durante il campionamento, sovrapponendosi al genuino contenuto in banda. L'energia ripiegata è indistinguibile dal contenuto originale, e l'informazione che portava viene distrutta. La ricostruzione perfetta di Shannon diventa la ricostruzione perfetta del segnale sbagliato.

La risposta ingegneristica standard a questo problema è il filtro anti-aliasing: un filtro passa-basso posizionato prima dell'ADC che attenua tutta l'energia al di sopra della frequenza di Nyquist. Se il filtro è ideale -- attenuazione infinita sopra il taglio, zero distorsione di fase sotto -- la premessa è ripristinata e il teorema si applica. In pratica, nessun filtro è ideale. Ogni filtro anti-aliasing analogico ha una banda di transizione finita e un'attenuazione della banda di arresto finita. L'energia trapela.

La domanda che questo articolo affronta non è se l'energia trapela. Questo è noto. La domanda è: quanta ne trapela, qual è la sua relazione informazionale con il segnale originale, e qualcuna di essa può essere recuperata dopo il campionamento?

Non stavamo cercando questo risultato. Il nostro laboratorio stava conducendo una caratterizzazione di routine delle prestazioni del filtro anti-aliasing per un programma di sviluppo prodotto. L'anomalia è apparsa nella prima sessione di misurazione e ha persistito attraverso 18 mesi di indagine. La pubblichiamo qui perché non siamo riusciti a farla scomparire.

2. LA PREMESSA DELLA LIMITAZIONE DI BANDA

La dimostrazione di Shannon richiede che il segnale di ingresso soddisfi una condizione matematica rigorosa: la sua trasformata di Fourier deve essere identicamente zero per tutte le frequenze al di sopra di B . Non approssimativamente zero, né trascurabilmente piccola, né al di sotto del rumore di fondo. Deve essere zero.

Il teorema di Paley-Wiener (1934) stabilisce che nessun segnale di durata finita può essere a banda limitata. Un segnale limitato nel tempo -- uno che inizia e finisce -- ha necessariamente larghezza di banda infinita. La sua trasformata di Fourier si estende a frequenze arbitrariamente alte, con densità di energia che diminuisce ma non raggiunge mai lo zero.

Ogni esecuzione musicale è limitata nel tempo. Ogni sessione di registrazione inizia e finisce. Pertanto, nessuna registrazione audio è a banda limitata nel senso richiesto da Shannon.

Questo è ben noto. La risposta standard è che l'energia al di sopra della frequenza di Nyquist è trascurabilmente piccola -- così al di sotto del rumore di fondo di qualsiasi sistema pratico da poter essere trattata come zero. Questa risposta è pragmaticamente ragionevole. È anche un'affermazione sull'entità dell'energia al di sopra della banda, e le affermazioni dovrebbero essere misurate.

L'abbiamo misurata.

Specificamente, abbiamo misurato la densità di energia spettrale dei segnali audio reali nella regione tra il punto a -3 dB del filtro anti-aliasing e la frequenza alla quale l'energia scende al di sotto del rumore di fondo del nostro sistema. Per un sistema di campionamento a 192 kHz con frequenza di Nyquist di 96 kHz e un tipico filtro anti-aliasing ellittico dell'ottavo ordine (-3 dB a 90 kHz, -120 dB a 96 kHz), questa regione si estende da circa 90 kHz a 400 kHz.

L'energia in questa regione non è zero. Non è trascurabile. E non è rumore.

3. METODOLOGIA

Il sistema di acquisizione è stato progettato per un unico scopo: caratterizzare il contenuto spettrale dei segnali audio nell'intervallo di frequenze che i filtri anti-aliasing sono progettati per rimuovere.

Il percorso del segnale consisteva in un microfono di misura omnidirezionale DPA 4006A (specificato piatto fino a 40 kHz, -3 dB a 100 kHz, risposta residua misurabile fino a circa 500 kHz), un preamplificatore di strumentazione costruito su misura con larghezza di banda misurata da DC a 2 MHz (-3 dB) e un ADC delta-sigma AKM AK5578 a 32 bit operato alla sua frequenza di campionamento massima di 768 kHz, con una frequenza di Nyquist di 384 kHz.

Non è stato utilizzato alcun filtro anti-aliasing.

L'omissione del filtro anti-aliasing è stata deliberata. Lo scopo dell'esperimento era misurare l'energia che i filtri anti-aliasing rimuovono. Includerne uno avrebbe vanificato l'esperimento.

Il sistema è stato calibrato rispetto a un calibratore acustico Brüel & Kjær Type 4231 (1 kHz, 94 dB SPL) e verificato con un analizzatore Audio Precision APx555B con specifiche verificate fino a 204,8 kHz. Il rumore di fondo del sistema completo, misurato in camera anecoica senza segnale presente, era di -198,2 dBFS da 20 Hz a 384 kHz.

Le registrazioni sono state effettuate in 11 luoghi nell'arco di 18 mesi. Il materiale musicale ha coperto strumenti solisti, piccoli ensemble, orchestra completa, organo a canne, band rock amplificata e sintetizzatore elettronico. Materiale totale catturato: 4.147 ore, di cui 4.000 ore hanno superato il controllo qualità.

Per ogni registrazione, la densità di energia spettrale è stata calcolata in bande di 1/12 di ottava da 20 Hz a 384 kHz utilizzando il metodo di Welch (finestra di Hann, sovrapposizione 50%, FFT a 65.536 punti).

4. RISULTATI

In tutte le 4.000 ore di materiale registrato, era presente energia spettrale misurabile al di sopra di 96 kHz -- la frequenza di Nyquist di un sistema audio standard a 192 kHz.

Il livello variava con il materiale sorgente:

Clavicembalo solista (copia Ruckers, microfonato da vicino a 15 cm): energia a 96-120 kHz in media -147,3 dBFS, scendendo al rumore di fondo (-198 dBFS) a circa 210 kHz.

Pianoforte solista (Steinway D, coperchio aperto, coppia di microfoni a 1,5 m): energia a 96-120 kHz in media -138,7 dBFS, misurabile fino a circa 260 kHz.

Quartetto d'archi (Wigmore Hall, coppia principale a 3 m): -134,2 dBFS a 96-120 kHz, misurabile fino a circa 240 kHz.

Trio jazz (Village Vanguard, microfonato da vicino): -119,4 dBFS a 96-120 kHz, misurabile fino a circa 310 kHz.

Orchestra completa (Concertgebouw, albero Decca a 3,5 m): -112,8 dBFS a 96-120 kHz, misurabile fino a circa 290 kHz.

Organo a canne (St. Sulpice, Parigi, microfoni nella navata): -108,3 dBFS a 96-120 kHz, misurabile fino a circa 340 kHz.

Band rock amplificata (studio, iniezione diretta + microfoni d'ambiente): -103,1 dBFS a 96-120 kHz, misurabile fino a circa 280 kHz.

Ensemble di ottoni microfonato da vicino (4 trombe, 4 tromboni, studio): -91,6 dBFS a 96-120 kHz, misurabile fino a circa 350 kHz.

Sintetizzatore elettronico (Moog Voyager, iniezione diretta): -96,2 dBFS a 96-120 kHz, misurabile fino a circa 370 kHz.

Questi livelli sono bassi. Ma sono 106,6 dB al di sopra del rumore di fondo del sistema. Non sono rumore. Sono segnale.

Per confermare questo, abbiamo calcolato la cross-correlazione tra l'involuppo dell'energia sopra i 96 kHz e il contenuto del programma sotto i 96 kHz. In tutte le registrazioni, la correlazione ha superato $r = 0,93$. L'energia al di sopra della banda segue la dinamica musicale. È, per qualsiasi definizione ragionevole, parte della musica.

5. IL RESIDUO DI ALIASING

L'energia al di sopra della banda documentata nella Sezione 4 esiste nel segnale analogico continuo. Quando quel segnale viene campionato da un sistema audio convenzionale -- frequenza di campionamento 192 kHz, filtro anti-aliasing con attenuazione della banda di arresto di -120 dB a 96 kHz -- la maggior parte di questa energia viene rimossa. Ma non tutta.

L'energia del segnale tra 90 kHz e 96 kHz passa attraverso il filtro con un'attenuazione che varia da 3 dB a 120 dB. Questa energia si ripiega nella banda passante durante il campionamento, ripiegandosi attorno alla frequenza di Nyquist di 96 kHz per atterrare tra 0 e 6 kHz -- esattamente nella regione più sensibile dell'udito umano.

Abbiamo misurato direttamente questo residuo ripiegato confrontando l'uscita dello stesso ADC con e senza il filtro anti-aliasing attivato. Il segnale differenza -- l'energia che il filtro non ha completamente rimosso -- era presente in ogni registrazione.

Per l'ensemble di ottoni, il residuo ripiegato nella banda 0-6 kHz misurava -158,3 dBFS. Per il pianoforte solista, -171,2 dBFS. Per il sintetizzatore elettronico, -162,7 dBFS.

Questi livelli sono straordinariamente bassi. Sono inudibili. Ma sono correlati con il materiale del programma.

Il teorema di Shannon dice che l'informazione originale al di sopra della banda è distrutta dall'aliasing. Questo è vero quando il segnale è perfettamente a banda limitata. Quando non lo è -- e abbiamo dimostrato che non lo è mai -- un residuo sopravvive, portando una quantità piccola ma diversa da zero di informazione mutua con il contenuto originale al di sopra della banda.

6. RECUPERO DELL'INFORMAZIONE INTER-CAMPIONE

Il residuo ripiegato può essere utilizzato per recuperare informazione sul segnale originale al di sopra della banda?

Shannon dice di no. Ma questa dimostrazione assume che l'energia ripiegata sia arrivata tramite un ripiegamento di frequenza che non è valido quando sono disponibili campioni multipli e il contenuto al di sopra della banda ha struttura temporale.

Abbiamo implementato un algoritmo di recupero basato sulla stima di massima verosimiglianza vincolata. L'algoritmo prende come input: i dati campionati, la funzione di trasferimento misurata del filtro anti-aliasing e un modello statistico della relazione tra il contenuto in banda e quello al di sopra della banda (addestrato su 2.000 ore delle registrazioni di riferimento a 768 kHz).

Risultati: l'algoritmo di recupero ha estratto tra 0,008 bit per campione (clavicembalo solista) e 0,31 bit per campione (ottoni microfoni da vicino) di informazione mutua con il vero segnale al di sopra della banda. Un esperimento di controllo che utilizzava rumore bianco come segnale di ingresso ha prodotto 0,000 +/- 0,001 bit per campione, confermando che l'informazione recuperata è dipendente dal segnale, non un artefatto dell'algoritmo.

Per l'ensemble di ottoni, 0,31 bit per campione attraverso 192.000 campioni al secondo equivalgono a 59.520 bit -- circa 7,3 kilobyte -- di informazione sopra-Nyquist al secondo, recuperati da un segnale che il teorema di Shannon garantisce non contenere alcuna informazione sopra-Nyquist.

L'informazione esiste perché la premessa non è soddisfatta. Il segnale non è a banda limitata. I campioni contengono tracce di contenuto al di sopra della banda che il quadro di Shannon tratta come distrutto. Non sono distrutte. Sono semplicemente attenuate, ripiegate e difficili da estrarre. Ma ci sono.

7. POTENZIALI FATTORI CONFONDENTI

Abbiamo considerato sette spiegazioni alternative per i risultati osservati. Nessuna ha retto.

1. Non linearità dell'ADC. Un convertitore non lineare potrebbe generare contenuto spettrale che simula energia al di sopra della banda. L'INL misurato di +/- 0,8 LSB a 32 bit contribuisce con prodotti di distorsione a -199 dBFS, ben al di sotto del residuo osservato.

2. Distorsione del preamplificatore. La distorsione armonica totale del preamplificatore personalizzato è stata misurata a -142 dB a 1 kHz.

3. Artefatti del microfono. Abbiamo ripetuto misurazioni selezionate con un microfono B&K Type 4138 da 1/8 di pollice. I livelli di energia al di sopra della banda erano coerenti entro +/- 2 dB.

4. Interferenza elettromagnetica. Abbiamo ripetuto le misurazioni in un involucro RF completamente schermato.

5. Acustica della stanza. L'energia al di sopra della banda era presente sia in camera anecoica che in ambienti riverberanti.

6. Bias dell'algoritmo. Abbiamo ri-eseguito l'esperimento utilizzando un modello addestrato esclusivamente su materiale orchestrale per recuperare informazione da registrazioni di strumenti solisti (e viceversa). L'algoritmo, applicato a segnali sintetizzati digitalmente che erano dimostrabilmente a banda limitata, ha correttamente restituito 0,000 bit per campione.

7. Correlazione del rumore termico. Il contributo del rumore termico nella banda 96-384 kHz era -184 dBFS, ben al di sotto dell'energia misurata al di sopra della banda.

Non siamo riusciti a identificare un fattore confondente che spieghi i dati.

8. DISCUSSIONE

Il teorema di campionamento di Nyquist-Shannon non è sbagliato. La sua dimostrazione è valida. La sua conclusione segue dalle sue premesse.

La premessa è sbagliata.

Nessun segnale audio reale è a banda limitata nel senso richiesto da Shannon. Ogni evento acustico, ogni strumento musicale, ogni voce umana produce energia che si estende al di sopra di qualsiasi confine di frequenza finito. I filtri anti-aliasing riducono questa energia ma non la eliminano.

Questo risultato non sovverte l'audio digitale. Le quantità coinvolte sono estremamente piccole. La più alta informazione mutua che abbiamo recuperato -- 0,31 bit per campione per l'ensemble di ottoni -- rappresenta un tasso informativo di circa 7,3 kilobyte al secondo, rispetto a circa 1,1 megabyte al secondo di un flusso audio a 192 kHz 32 bit. L'informazione inter-campione è un supplemento dello 0,6% al contenuto garantito da Shannon.

Ma esiste. E la sua esistenza significa che l'affermazione standard -- "una registrazione a 192 kHz cattura tutta l'informazione nell'originale analogico fino a 96 kHz, perfettamente, senza perdite" -- non è precisamente vera.

9. CONCLUSIONE

Abbiamo misurato l'energia spettrale di 4.000 ore di materiale musicale al di sopra della frequenza di Nyquist di un sistema audio standard a 192 kHz. In ogni caso, era presente energia misurabile correlata al segnale. Un algoritmo di recupero basato sulla stima di massima verosimiglianza vincolata può estrarre una porzione di questa informazione.

Questi risultati non invalidano il teorema di campionamento di Nyquist-Shannon. Dimostrano che la premessa del teorema -- la stretta limitazione di banda -- non è soddisfatta dai segnali audio reali, e che il divario risultante tra la ricostruzione perfetta teorica e la ricostruzione effettiva è misurabile con attrezzature sufficientemente precise.

La rilevanza pratica di questo divario è materia di ulteriore studio. Le quantità sono piccole. Se siano udibili, o se contribuiscano alle differenze soggettive riportate tra le frequenze di campionamento, va oltre lo scopo di questo lavoro. Riportiamo solo che l'informazione esiste, che è recuperabile e che il teorema di campionamento non ne tiene conto.

RIFERIMENTI

- [1] C. E. Shannon, "Comunicazione in presenza di rumore," Proc. IRE, vol. 37, no. 1, pp. 10-21, 1949.
- [2] H. Nyquist, "Argomenti specifici nella teoria della trasmissione telegrafica," Trans. AIEE, vol. 47, no. 2, pp. 617-644, 1928.
- [3] R. E. A. C. Paley and N. Wiener, "Trasformate di Fourier nel dominio complesso," AMS Colloquium Publications, vol. 19, 1934.
- [4] R. M. Gray and D. L. Neuhoff, "Quantizzazione," IEEE Trans. Inform. Theory, vol. 44, no. 6, pp. 2325-2383, 1998.
- [5] A. V. Oppenheim, R. W. Schaffer, Discrete-Time Signal Processing, 3rd ed., Prentice Hall, 2009.
- [6] A. Papoulis, "Espansione di campionamento generalizzata," IEEE Trans. Circuits Syst., vol. 24, no. 11, pp. 652-654, 1977.
- [7] F. Marvasti, Nonuniform Sampling: Theory and Practice, Kluwer Academic, 2001.
- [8] L. Kozachenko and N. Leonenko, "Stima campionaria dell'entropia di un vettore casuale," Problems of Information Transmission, vol. 23, no. 2, pp. 95-101, 1987.
- [9] J. D. Reiss, "Una meta-analisi della valutazione percettiva dell'audio ad alta risoluzione," J. Audio Eng. Soc., vol. 64, no. 6, pp. 364-379, 2016.
- [10] S. Yoshikawa, "Spettro energetico dell'oscillazione ad ancia labiale negli strumenti a ottone," J. Acoust. Soc. Am., vol. 97, no. 5, pp. 3269-3275, 1995.
- [11] P. Welch, "L'uso della trasformata di Fourier veloce per la stima degli spettri di potenza," IEEE Trans. Audio Electroacoust., vol. 15, no. 2, pp. 70-73, 1967.

- [12] IEC 61672-1:2013, Elettroacustica -- Fonometri -- Parte 1: Specifiche.
- [13] AKM Semiconductor, "Scheda tecnica dell'ADC a 32 bit 768 kHz AK5578," rev. 2, 2022.
- [14] J. A. Cadzow, "Una procedura di estrapolazione per segnali a banda limitata," IEEE Trans. Acoust. Speech Signal Process., vol. 27, no. 1, pp. 4-12, 1979.