Deconvolució

Marc Franco Meca

23 Març 2020

1 Deconvolució

Un cop determinats els diferents tipus de senyals d'excitació del sistema acústic, cal determinar un altre element indispensable com és la tècnica de deconvolució amb la qual obtindrem la resposta impulsional. La deconvolució és l'operació inversa a la convolució i ens permet revertir els efectes de la convolució de la sortida mesurada.

Recordem que generant un senyal d'excitació $x(t)$, reproduint-lo a través d'un sistema acústic $h(t)$, que es correspon amb el recinte acústic, i capturant-lo obtenim el senyal mesurat $y(t)$. Per obtenir la resposta impulsional del sistema acústic $h(t)$ s'ha d'aplicar una tècnica de deconvolució entre senyal mesurat i el senyal d'excitació. Aquesta operació de deconvolució consisteix en aplicar una convolució lineal del senyal de sortida $y(t)$ amb el filtre invers $f(t)$ pre-processat del senyal d'excitació, expressat matemàticament amb la següent formulació

$$h(t) = y(t) * f(t) \tag{1}$$

A continuació es mostra un diagrama de blocs per a visualitzar gràficament el procediment,

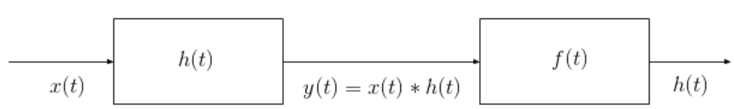

Figure 1: Diagrama de blocs del procés de la deconvolució d'un senyal

2 El Filtre Invers

Com hem vist, la deconvolució de la resposta d'impuls requereix la creació d'un filtre invers $f(t)$ amb la capacitat de convertir el senyal d'excitació inicial en la funció delta de Dirac δ retardada. Aquesta característica o propietat es pot expressar matemàticament com

$$x(t) * f(t) = \delta(t - K) \tag{2}$$

El filtre invers $f(t)$ es genera tenint en compte quin és el senyal d'excitació. En el nostre cas utilitzem el sweep logarítmic el qual és un senyal estable i causal, és a dir, depèn només dels valors presents i passats de l'entrada, no de futurs valors. Aquest filtre es dissenya mitjançant un sweep logarítmic invertit temporalment, retardat per obtenir un senyal causal i amb una modulació de l'amplitud per compensar la caiguda espectral.

Escrita d'una altra manera, si recordem que l'equació que defineix l'escombrat logarítmic és la següent,

$$s(t) = sin[\theta(t)] = sin[K \cdot (e^{\frac{t}{L}} - 1)] \tag{3}$$

on,

$$K = \frac{T \cdot \omega_1}{ln(\frac{\omega_2}{\omega_1})}, L = \frac{T}{ln(\frac{\omega_2}{\omega_1})} \tag{4}$$

on ω_1 i ω_2 són els extrems del rang freqüencial i T es correspon amb el temps de duració en segons.

Es pot definir la freqüència instantània com

$$\omega(t) = \frac{d[\theta(t)]}{dt} = \frac{K}{L} \cdot e^{\frac{t}{L}} \tag{5}$$

L'energia en qualsevol freqüència d'un senyal que varia la seva freqüència és proporcional al temps que està excitant aquella freqüència en particular. D'aquesta manera es pot definir l'energia $E(t)$ d'un escombrat logarítmic en funció del temps com,

$$E(t) \propto \frac{1}{\omega'(t)} = \frac{L^2}{K} \cdot e^{-\frac{t}{L}}, \tag{6}$$

on $\omega'(t)$ és la taxa de canvi de freqüència expressada en temps.

Anàlogament es pot expressar l'energia $E(jw)$ en funció de la freqüència mitjançant la següent expressió

$$E(j\omega) = \frac{kL^2}{K} \cdot \frac{1}{L + j\omega} \tag{7}$$

on k és una constant de proporcionalitat.

Intuitivament es pot deduir que l'augment de la freqüència suposa una disminució de l'energia. Si observem l'equació veiem que hi ha un factor $\frac{1}{\omega}$ el qual al doblar la freqüència esdevé $\frac{1}{2\omega}$ i suposa una caiguda d'energia de 3dB, $10log_{10}(\frac{1}{2})$. Dit d'una altra manera, l'espectre d'energia cau -3dB/oct tant en el senyal d'excitació original com l'invertit temporalment.

Per a compensar aquestes diferències d'energia i obtenir un espectre pla cal aplicar una modulació d'amplitud. Farina [1] enuncia que per generar el filtre invers només cal invertir temporalment el senyal d'excitació, i després aplicar-ne un factor per a reduir la amplitud amb un nivell de 6dB/oct, començant per 0dB i acabant amb $-6log_2\frac{\omega_2}{\omega_1}$. Aplicant aquesta correcció d'amplitud de 6dB/oct al senyal invertit fa que es compensi amb el senyal d'excitació.

El factor de correcció de l'amplitud es pot obtenir trobant el temps Δt en què la funció de l'escombrat logarítmic té una freqüència instantània igual a N vegades la actual. Aquesta condició es pot expressar matemàticament com

$$N \cdot \frac{d}{dt}\left[\frac{\omega_1 \cdot T}{ln(\frac{\omega_2}{\omega_1})} \cdot \left(e^{\frac{t}{T}\cdot ln(\frac{\omega_2}{\omega_1})} - 1\right)\right] = \frac{d}{dt}\left[\frac{\omega_1 \cdot T}{ln(\frac{\omega_2}{\omega_1})} \cdot \left(e^{\frac{t+\Delta t}{T}\cdot ln(\frac{\omega_2}{\omega_1})} - 1\right)\right] \quad (8)$$

i obtenint la següent expressió en funció del temps,

$$\Delta t = T \cdot \frac{ln(N)}{ln(\frac{\omega_2}{\omega_1})} \quad (9)$$

i anàlogament es pot expressar en funció de la freqüència com,

$$N = e^{\frac{\Delta t \cdot ln(\frac{\omega_2}{\omega_1})}{T}} = e^{\frac{\Delta t}{L}} \quad (10)$$

3 No Linealitat dels Sistemes Acústics

Quan el senyal mesurat $y(t)$ es convoluciona amb el filtre invers $f(t)$ explicat anteriorment obtenim gairebé una resposta impulsional perfecte amb artefactes causats per les distorsions harmòniques del sistema de reproducció del senyal en uns certs instants de temps previs a la resposta impulsional.

Els altaveus o sistemes acústics que produeixen vibracions o emeten un so es comporten de manera diferent en funció de l'amplitud. Aquesta dependència de l'amplitud és un indicador de no linealitat en el sistema que acostuma a mostrar una distorsió del senyal, especialment a amplituds altes. Un altre tipus de distorsió o no linealitat que proporciona el sistema és la distorsió harmònica. Aquest efecte de distorsió consisteix en la generació de noves components espectrals que són inexistents en el senyal d'excitació. Aquestes components es generen normalment a freqüències múltiples de la fonamental, d'aquí el seu nom distorsió harmònica. Els resultats d'aquesta distorsió depenent fortament amb les propietats de l'estímul o senyal d'excitació, com per exemple la freqüència, l'amplitud o la fase. Una altra distorsió que es produeix donada la no linealitat del sistema és la distorsió d'intermodulacio, la qual genera components a freqüències que són sumes o diferències de la freqüència de senyal d'excitació i els seus múltiples. La darrera distorsió que produeixen els altaveus és de caire aleatori i no determinista i es manifesta en forma de soroll o de distorsió del so.

Quan el senyal d'excitació $x(t)$, el qual està format per una amplitud constant i seguit d'uns instants de silenci, es reprodueix per un altaveu i la resposta impulsional es capta a través d'un micròfon, el senyal resultant conté els efectes de la reverberació de la sala, el soroll i les distorsions no lineals explicades anteriorment.

El fet d'utilitzar l'escombrat freqüencial logarítmic possibilita fer la deconvolució de la resposta impulsional del sistema o sala i simultàniament separar la resposta impulsional de les distorsions ocasionades pel sistema acústic, ignorant les distorsions harmòniques que apareixen abans de la resposta impulsional principal. Arribat a aquest punt, es lògic pensar en aplicar un enfinestrament per extreure només el segment que conté la resposta principal i descartar les distorsions, el qual es pot realitzar fàcilment situant la resposta impulsional a l'inici i eliminant tot el contingut anterior.

4 Càlcul del RT60 a partir de la resposta impulsional

Un cop hem obtingut la resposta impulsional $h(t)$ del nostre sistema o recinte acústic podem realitzar el càlcul dels paràmetres acústics especificats en apartats anteriors.

Per a extreure el paràmetre acústic del temps de reverberació el primer que cal fer és filtrar la resposta a l'impuls $h(t)$ en bandes de freqüència, obtenint així un senyal la resposta a l'impuls filtrada $h_f(t)$. Això es degut a que el temps de reverberació varia en funció de la freqüència a la qual es mesura i la posició en què es pren dins la sala. Tot i que s'assumeix el fet que el camp sonor és difús i que el decaïment del so és el mateix en cada posició de la sala, s'ha de considerar la dependència respecte les bandes freqüencials. Per això, es pot filtrar la resposta impulsional utilitzant filtres d'octava o d'un terç d'octava. L'espectre audible d'àudio en éssers humans va desde 20Hz a 20kHz, per tant, si fem ús de filtres d'octava el podem dividir en 11 bandes d'octaves, o bé en 31 bandes si utilitzem filtres d'un terç d'octava. Aleshores, per a cada banda freqüencial, la meitat de l'octava inferior i superior vé donada per la següent fòrmula $f_n^{low} = f_n/2^{\frac{1}{2}}$ i $f_n^{high} = f_n \cdot 2^{\frac{1}{2}}$ en el cas dels filtres d'octava i per $f_n^{low} = f_n/2^{\frac{1}{6}}$ i $f_n^{high} = f_n \cdot 2^{\frac{1}{6}}$ les corresponents als filtres d'un terç d'octava.

En quant al tipus de filtre el més utilitzat és el filtre de Butterworth de tercer ordre, el qual otorga una resposta en freqüència plana i una caiguda de 18dB per octava.

Obtinguda la resposta impulsional filtrada $h_f(t)$ el següent pas és suavitzar aquest senyal el màxim possible abans de convertir-lo a l'escala logarítmica. Per a fer això s'utilitza la Transformada de Hilbert H, la qual és una eina matemàtica per descriure l'envolvent complexa d'un senyal modulat per una portadora real. La transformada H d'una funció real $s(t)$ es defineix com la convolució d'aquest senyal real $s(t)$ i la resposta a l'impuls $\frac{1}{\pi t}$. L'objectiu d'aplicar aquesta transformada és crear un senyal analític

$$h_A(t) = h(t) + j\tilde{h}(t) \tag{11}$$

on $\tilde{h}(t)$ és la Transformada de Hilbert de la resposta impulsional i j és un nombre complex.

Coneixent que el senyal analític és complexe, mitjançant la representació de fasors podem determinar que aquest ve expressat per l'amplitud i la fase.

$$h_A(t) = A(t) \cdot e^{j\psi(t)} \tag{12}$$

on $A(t)$ és l'envolvent del senyal i ψ la fase instantània.

En el nostre cas només estem interessats en l'envolvent del senyal, que es correspon amb la magnitud del senyal analític. A continuació es pot visualitzar la resposta impulsional filtrada i la seva envolvent calculada mitjançant la transformada H.

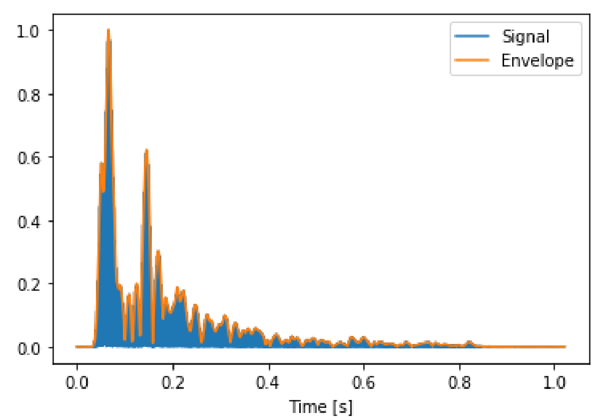

Figure 2: Representació gràfica de la resposta impulsional filtrada i l'envolvent

Tot i haver suavitzat la resposta impulsional cal una darrera suavització del senyal. En aquest cas es fa ús dels filtres de mitjana mòbil o moving average filters en anglès. Aquests són un filtre passa baix de resposta impulsional finita que agafen L mostres del senyal d'entrada, calculen la mitjana d'aquestes L mostres i produeixen una única mostra de sortida.

Es defineix l'equació discreta d'un filtre de mitjana mòbil de L mostres amb un vector d'entrada x i el vector de sortida y com,

$$y[n] = \frac{1}{L}\sum_{k=0}^{L-1} x[n-k] \tag{13}$$

Anàlogament la resposta en freqüència d'un filtre de mitjana mòbil s'expressa de la següent manera,

$$H(f) = \frac{sin(\pi f L)}{Lsin(\pi f)} \tag{14}$$

Un cop hem filtrat el senyal $h_f(t)$ amb el filtre de mitjana mòbil amb una finestra de L samples, es procedeix a passar a escala logarítmica el senyal $A(t)$ convertint-lo així en una corba d'energia $E(t)$ la qual és una versió suavitzada de l'envolvent de la Transformada de Hilbert en escala logarítmica.

La conversió a escala logarítmica es dona mitjançant l'expressió,

$$E(t) = 20log_{10}\frac{A(t)}{maxA(t)} \tag{15}$$

El següent pas consisteix en aplicar el mètode d'Integració de Schroeder a l'envolvent per a suavitzar i fer més senzills els càlculs. Aquest mètode proporciona una corba que s'obté mitjançant la integració endarrere del quadrat de la resposta impulsional. D'aquesta manera ens proporciona una corba de decaïment plana i fàcil d'implementar mitjançant la següent expressió matemàtica,

$$L(t) = 10log_{10}[\frac{\int_t^\infty h^2(\tau)d\tau}{\int_0^\infty h^2(\tau)d\tau}] \tag{16}$$

Aleshores, per a fer el càlcul del paràmetre acústic del temps de reverberació es necessari realitzar una interpolació lineal entre la corba de decaïment obtinguda amb l'integració de Schroeder i la funció lineal $L = A \cdot t + B$ en el rang que es vulgui calcular segons el paràmetre. Com s'ha explicat en apartats anteriors no és possible tenir un rang dinàmic de 60dB, per tant, el rang en què es fa la interpolació és limitat i s'extrapola el resultat.

Un cop feta aquesta regressió lineal, es pot calcular el temps de reverberació fent ús del gradient d'aquesta línia i utilitzant la següent expressió matemàtica

$$RT = \frac{-60}{A} \tag{17}$$

on A és el coeficient de la corba treta de la línia de la interpolació en dB/s.

References

[1] Angelo Farina. Simultaneous measurement of impulse response and distortion with a swept-sine technique. In *Audio Engineering Society Convention 108*. Audio Engineering Society, 2000.

[2] Guy-Bart Stan, Jean-Jacques Embrechts, and Dominique Archambeau. Comparison of different impulse response measurement techniques. *Journal of the Audio engineering society*, 50(4):249–262, 2002.

[3] T. M. Culda, V. Popa, D. Stanomir, and C. Negrescu. Reducing time in acoustic impulse response measurements using exponential sine sweeps. In *International Symposium on Signals, Circuits and Systems ISSCS2013*, pages 1–4, July 2013.

[4] Q Meng, D Sen, S Wang, and L Hayes. Impulse response measurement with sine sweeps and amplitude modulation schemes. In *2008 2nd International Conference on Signal Processing and Communication Systems*, pages 1–5. IEEE, 2008.

[5] Angelo Farina. Advancements in impulse response measurements by sine sweeps. In *Audio Engineering Society Convention 122*. Audio Engineering Society, 2007.

[6] Thatcher Ulrich. Diy capture of room impulse responses, 2017.

[7] Wolfgang Klippel. Loudspeaker nonlinearities–causes, parameters, symptoms. In *Audio Engineering Society Convention 119*. Audio Engineering Society, 2005.

[8] Viktor Wilhelmsson. Measuring loudspeaker distortion and room reverberation time using a speakerphone, 2016.