

*A cura di Diego Tuzi – ing tlc  
Valeria Di Marco – ing mtm*

 POLITECNICO DI MILANO



*RIRsim:*

# *Room Impulse Response Simulation*

*Corso: Audio and Video signals*

*A.A. 2011/2012*



## Breve descrizione del progetto

*Il progetto costituisce uno studio approfondito su un'applicazione che permetta di stimare la risposta all'impulso di una stanza.*

*Il metodo utilizzato è l'Image Source, con una particolare attenzione alla possibilità di assegnare coefficienti di riflessione delle pareti variabili in frequenza e alla gestione dei ritardi frazionari.*

*Il progetto comprende l'implementazione in MATLAB del metodo nel programma RIRsim, il quale grazie all'interfaccia grafica, permette all'utente di scegliere le features del problema, quali le dimensioni della stanza, il materiale delle pareti, il numero di riflessioni, la posizione della sorgente, ecc.*



# *Image source method*

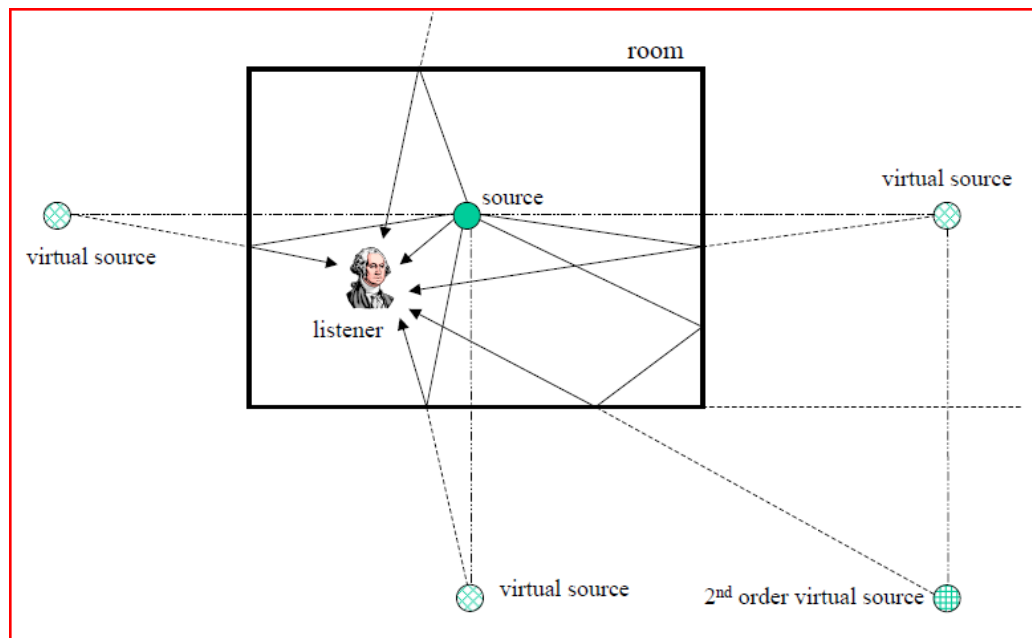
## *Assunzioni:*

- *stanza rettangolare (scatola di scarpe)*
- *no assorbimento dell'aria*
- *ogni parete si comporta come uno specchio (acustico) e dà origine ad una sorgente immagine*
- *ogni raggio emesso dalla sorgente  $\mathcal{A}$  e riflesso dalla parete si può pensare come emesso dalla sorgente immagine  $\mathcal{A}'$  le cui caratteristiche direzionali sono le simmetriche di  $\mathcal{A}$*



- *ciascun raggio subisce riflessioni multiple e si arresta quando incontra una parete completamente assorbente o esaurisce la sua energia*
- *il numero di riflessioni determina l'ordine del metodo*

*Il percorso è rappresentabile come segue:*





*L' algoritmo più diffuso per la risoluzione dell' Image Source Method è stato proposto da Allen e Berkley nel 1979.*

*Chiamando  $\beta$  i coefficienti di riflessione delle sei pareti, la risposta finita all' impulso della stanza si ottiene come segue:*

$$h(r, r_s, t) = \sum_{p \in P} \sum_{m \in M} \beta_{x_1}^{|m_x - q|} \beta_{x_2}^{|m_x|} \beta_{y_1}^{|m_y - j|} \beta_{y_2}^{|m_y|} \beta_{z_1}^{|m_z - k|} \beta_{z_2}^{|m_z|} \frac{\delta(t - \tau)}{4\pi d}$$

*$d = \|R_p + R_m\|$  è la distanza tra il ricevitore e la sorgente immagine*

*$\tau = \frac{d}{c}$  il tempo d' arrivo del suono ritardato,  $c$  la velocità del suono*

*vediamo in dettaglio...*



*Il vettore  $[x_s, y_s, z_s]$  rappresenta la posizione della sorgente,  $[x, y, z]$  la posizione del ricevitore, entrambi rispetto all'origine posta in uno degli angoli della stanza.*

*La posizione delle immagini rispetto al ricevitore, simmetriche ai muri  $x=0$ ,  $y=0$ ,  $z=0$ , è data da:*

$$\mathcal{R}_p = [ (1-2q)x_s - x, (1-2j)y_s - y, (1-2k)z_s - z ]$$

$$\mathcal{P} = \{ (q, j, k) : q, j, k \in \{0, 1\} \}$$

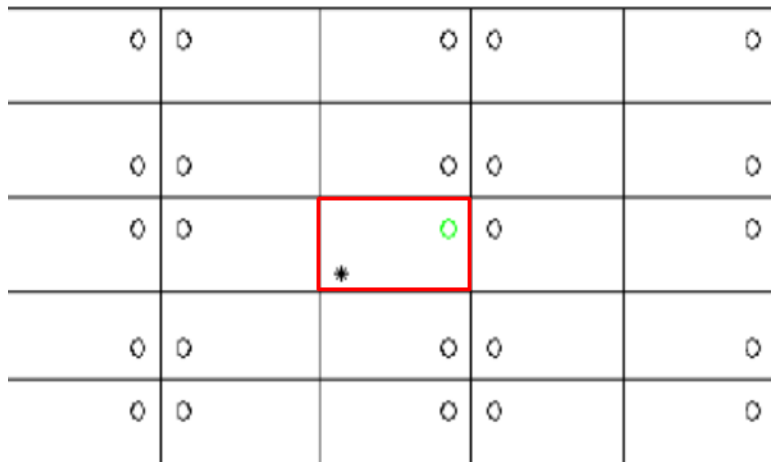
*Quando il valore di  $\mathbf{p}$  è 1 in una certa direzione, vuol dire che si sta considerando un'immagine della sorgente in quella precisa direzione.*



Data la simmetria del problema, tutte le altre immagini si ottengono trasladando nelle tre direzioni le sorgenti precedentemente calcolate, di un multiplo di due volte le dimensioni della stanza:

$$\mathcal{R}_m = [ 2m_x \mathcal{L}_x, 2m_y \mathcal{L}_y, 2m_z \mathcal{L}_z ]$$

$$\mathcal{M} = \{ (m_x, m_y, m_z) : -\mathcal{N} \leq m_x, m_y, m_z \leq \mathcal{N} \}$$

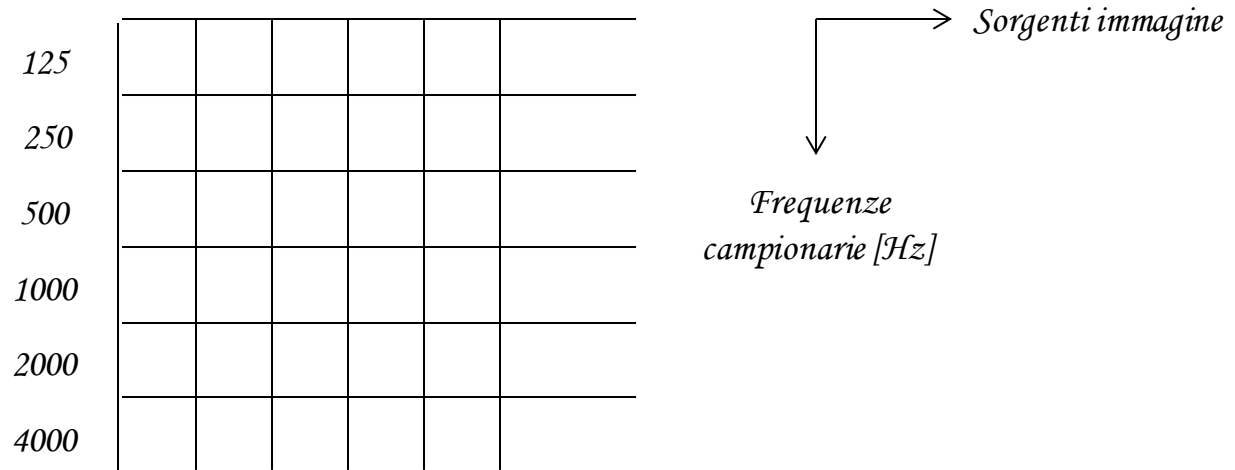


Si può ricavare anche l'ordine di riflessione della generica immagine alla posizione  $\mathcal{R}_p + \mathcal{R}_m + r$ :

$$O_{p,m} = |2m_x - q| + |2m_y - j| + |2m_z - k|$$

# ↘ Coefficienti variabili in frequenza

*Si crea una matrice che contiene per ogni sorgente immagine il coefficiente di riflessione, escludendo, da tutte le possibili combinazioni degli insiemi  $\mathcal{P}$  e  $\mathcal{M}$ , le immagini la cui distanza in campioni dal ricevitore è superiore alla lunghezza della RIR*







- *Si seleziona una colonna e l'interpolazione lineare dei valori viene ricampionata in 256 campioni. I valori sulle frequenze negative vengono ricostruiti come simmetrici coniugati sfruttando la seguente proprietà:*

$$x(t) \text{ reale} \leftrightarrow X(f) = X^*(-f)$$

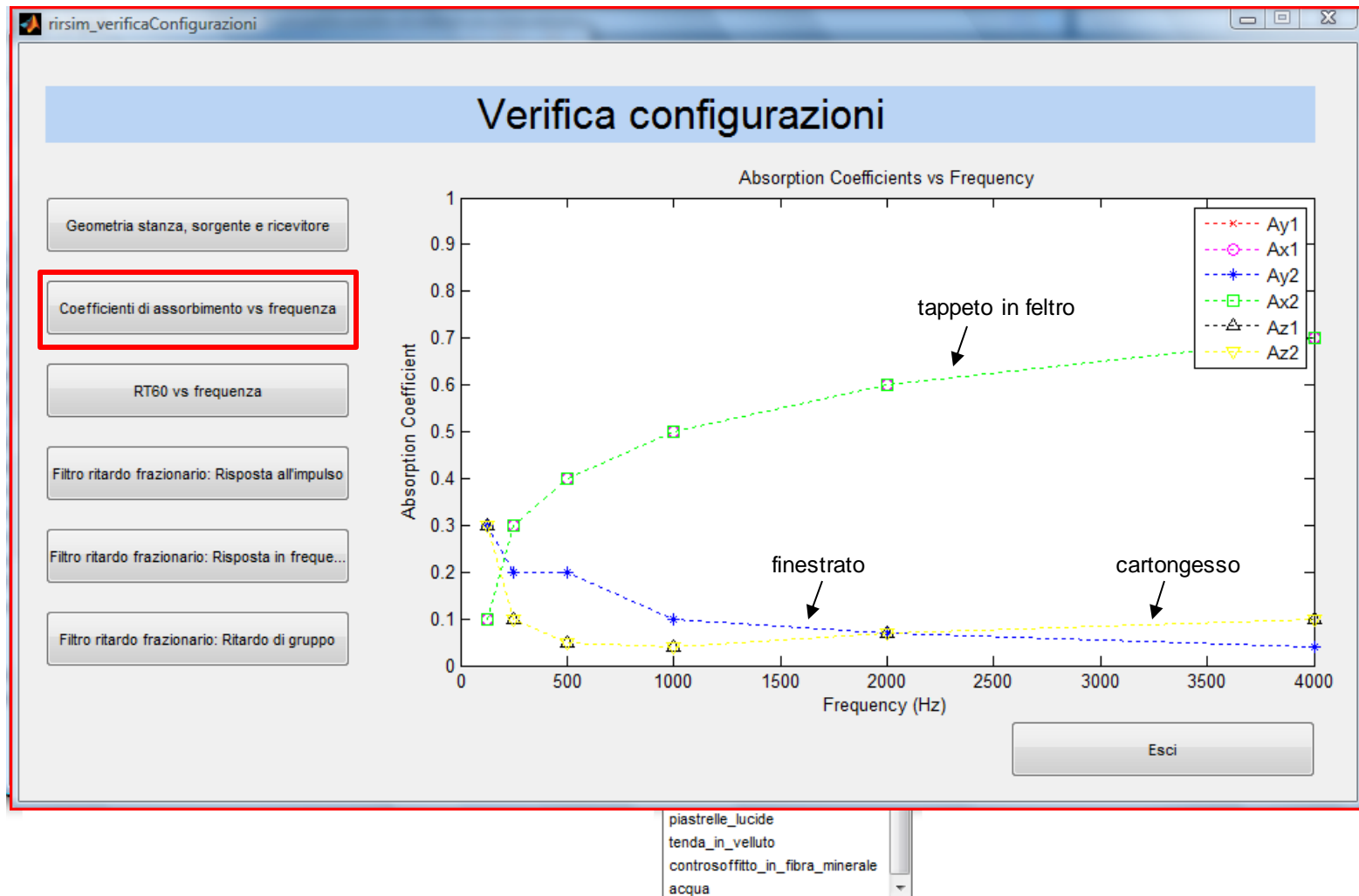
*vogliamo infatti che la RIR sia un segnale reale nel tempo*

- *Al contributo in frequenza viene applicata la IFFT per ricavare i valori nel tempo e lo si moltiplica per un Hanning window con semplice funzione di pesatura*
- *Si applica la convoluzione con il filtro per il ritardo frazionario*

*Il procedimento si itera sull'intera matrice e i contributi vengono sommati nella RIR finale*



# Coefficienti variabili in frequenza in RIRsim





# Ritardi Frazionari

*Si ha ritardo frazionario quando il contributo di ampiezza alla  $RIR$  di una sorgente immagine cade tra due istanti di campionamento.*

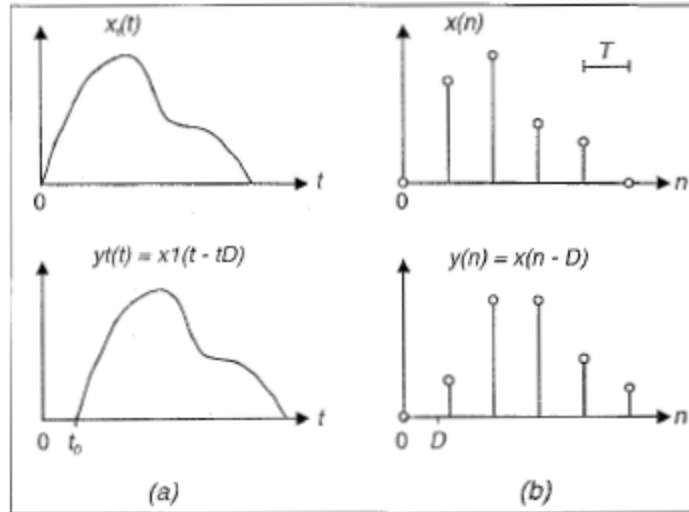
*Il filtro per il ritardo frazionario ha lo scopo di distribuire tale contributo sugli istanti di campionamento.*

*Ci sono differenti modi per realizzare un filtro per il ritardo frazionario, i filtri utilizzati nel progetto sono:*

- *Sinc + Hamming*
- *Sinc troncata*
- *General least-square*
- *Lagrange*



# Formulazione generale



*Traslare una sequenza discreta di un numero di campioni non intero, può essere visto come un processo di ricampionamento.*

*Tale operazione può essere ridotta ad una appropriata operazione di filtraggio lineare.*

*Gli obiettivi per il design del filtro sono i seguenti:*

$$H_{id}(e^{j\omega}) = e^{-j\omega D}$$

$$|H_{id}(e^{j\omega})| = 1, \quad \arg\{H_{id}(e^{j\omega})\} = \Theta_{id}(\omega) = -D\omega$$

$$\tau_{g,id}(\omega) = D, \quad \tau_{p,id}(\omega) = D$$



## Soluzione ideale

*La risposta all'impulso del filtro può essere ottenuta mediante la I-DTFT.*

$$h(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\omega}) e^{j\omega n} d\omega$$

*Ponendo  $H(e^{j\omega}) = H_{id}(e^{j\omega})$  si ottiene la risposta impulsiva ideale.*

$$h_{id}(n) = \frac{\sin[\pi(n - D)]}{\pi(n - D)} = \mathbf{sinc}(n - D)$$

Osservazioni:

- *il filtro è non causale*
- *se  $D$  non intero la lunghezza del filtro è infinita*
- *non realizzabile praticamente*



# Approssimazione con filtro FIR

*L'impossibilità pratica di realizzare il filtro ideale, genera la necessità di realizzarne un'approssimazione con un filtro FIR di ordine  $N$  e lunghezza  $L=N+1$*

$$H(z) = \sum_{n=0}^N h(n)z^{-n}$$

*Una misura interessante per la definizione degli obiettivi di design del filtro per il ritardo frazionario è l'errore di approssimazione del filtro FIR rispetto a quello ideale:*

$$E(e^{-i\omega}) = H(e^{-i\omega}) - H_{id}(e^{-i\omega})$$



# Sinc troncata

Obiettivo: minimizzare l'errore quadratico.

$$E = \frac{1}{\pi} \int_0^{\pi} |E(e^{j\omega})|^2 d\omega = \sum_{n=-\infty}^{+\infty} |h(n) - h_{id}(n)|^2$$

La soluzione si ottiene troncando la risposta ideale.

$$h(n) = \begin{cases} \mathit{sinc}(n - D) & M \leq n \leq M + N \\ \mathbf{0} & \text{altrimenti} \end{cases}$$

Errore residuo:  $E = \sum_{n=-\infty}^{M-1} |h_{id}(n)|^2 + \sum_{n=M+N+1}^{+\infty} |h_{id}(n)|^2$

Osservazioni:

- all'aumentare dell'ordine del filtro l'errore diminuisce
- all'aumentare dell'ordine aumenta la complessità del filtro
- fenomeno di Gibbs ovvero ripple sul modulo della risposta in frequenza



# *Sinc con Hamming*

Obiettivo: ridurre il fenomeno di Gibbs che si presenta con il troncamento della risposta impulsiva ideale.

L'obiettivo si raggiunge moltiplicando la risposta impulsiva con una finestra di Hamming. L'operazione di finestramento permette di addolcire le transizioni ai bordi della risposta e attenuare l'indesiderato ripple in frequenza.

$$h(n) = \begin{cases} W(n - D)\text{sinc}(n - D) & 0 \leq n \leq N \\ 0 & \text{altrimenti} \end{cases}$$

Osservazioni:

- errore di approssimazione maggiore rispetto alla sinc troncata
- ottenuto mediante semplici considerazioni e non da una formulazione matematica





# General least-square (1)

Obiettivo: minimizzare l'errore su una precisa banda di frequenze  $[0, \alpha\pi]$ , con l'ausilio di una pesatura in frequenza  $W(\omega)$ :

$$E = \frac{1}{\pi} \int_0^{\alpha\pi} W(\omega) |E(e^{j\omega})|^2 d\omega$$

Soluzione:  $\mathbf{h} = \mathbf{P}^{-1}\mathbf{p}$

$$\mathbf{P} = \frac{1}{\pi} \int_0^{\alpha\pi} W(\omega) \mathbf{C} d\omega \quad \mathbf{p} = \frac{1}{\pi} \int_0^{\alpha\pi} W(\omega) [\operatorname{Re}\{H_{id}(e^{j\omega})\}\mathbf{c} - \operatorname{Im}\{H_{id}(e^{j\omega})\}\mathbf{s}] d\omega$$
$$\mathbf{C} = \begin{bmatrix} 1 & \cdots & \cos(N\omega) \\ \vdots & \ddots & \vdots \\ \cos(N\omega) & \cdots & 1 \end{bmatrix}, \quad \mathbf{c} = [0 \quad \cdots \quad \cos(N\omega)]^T, \quad \mathbf{s} = [0 \quad \cdots \quad \sin(N\omega)]^T$$



### Osservazioni:

- *formulazione matematica valida per una generica espressione di  $H_{id}$*
- *il calcolo degli integrali di  $\mathbf{P}$  e  $\mathbf{p}$  è in forma numerica*
- *complessità computazionale elevata a causa della flessibilità di implementazione  $\rightarrow$  inversione di matrice  $O(N + 1)^3$*



## General least-square (2)

*Nel caso  $H_{id}(e^{j\omega}) = e^{-j\omega D}$  e  $W(\omega) = 1$  ci sono notevoli semplificazioni:*

- il calcolo degli integrali di  $\mathbf{P}$  e  $\mathbf{p}$  è in forma chiusa*
- la matrice  $\mathbf{P}$  è indipendente dal valore di ritardo  $D$  e quindi può essere calcolata ed invertita una sola volta*
  - inversione di matrice, algoritmo di Levinson, complessità  $O(N)^2$*

### Osservazioni:

- dipendenza dalla banda di approssimazione e della funzione peso*
- riduzione dell'errore nella banda di approssimazione, rispetto ai metodi precedenti*
- non ha controllo dell'errore al di fuori della banda di approssimazione*



# Lagrange (1)

Obiettivo: rendere massimamente piatta la funzione errore  $E(e^{j\omega})$  intorno ad una precisa frequenza  $\omega_0$ , tipicamente  $\omega_0 = 0$ , in modo tale da avere la migliore approssimazione vicino a tale frequenza.

La formulazione matematica dell'obiettivo è un sistema lineare di  $L=N+1$  equazioni:

$$\left. \frac{d^n E(e^{j\omega})}{d\omega^n} \right|_{\omega=\omega_0} = 0, \quad n = 0, 1, 2, \dots, N$$

che in termini di risposta impulsiva diviene

$$\sum_{k=0}^N k^n h(k) = D^n, \quad n = 0, 1, 2, \dots, N$$

Si nota che è possibile ottenere anche una formulazione matriciale del problema.



## Lagrange (2)

*La soluzione del problema è equivalente alla classica formula di **Interpolazione Lagrangiana**, dove i coefficienti sono ottenuti adattando il polinomio interpolatore ad un insieme di valori.*

*La soluzione può essere fornita in forma esplicita*

$$h(n) = \prod_{k=0, k \neq n}^N \frac{D - k}{n - k}, \quad n = 0, 1, 2, \dots, N$$

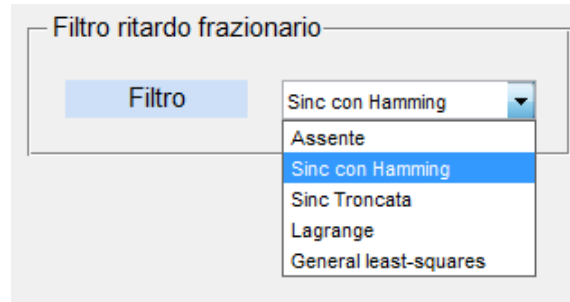
Osservazioni:

- *caso  $N = 1 \rightarrow$  interpolazione lineare*
- *filtro lunghezza pari  $\rightarrow$  fase della risposta in frequenza migliore*
- *filtro lunghezza dispari  $\rightarrow$  modulo della risposta in frequenza migliore*
- *formula esplicita  $\rightarrow$  facile da implementare*
- *migliore risposta alle basse frequenze*

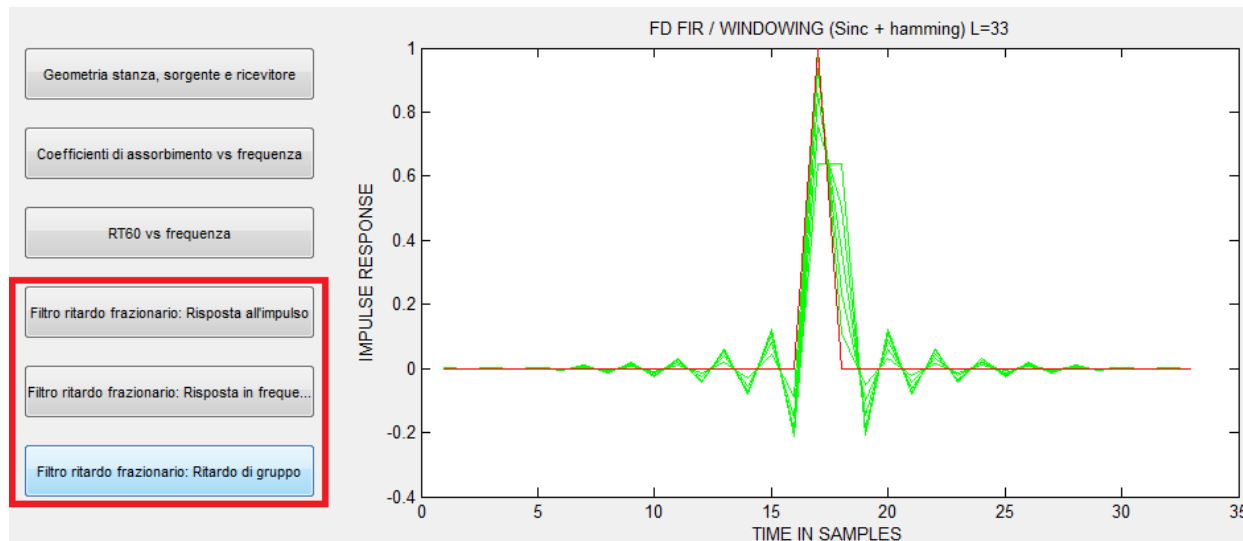


# Ritardi frazionari in RIRsim

*E' possibile scegliere il tipo di filtro per il ritardo frazionario da utilizzare.*



*Risposta impulsiva, modulo della risposta in frequenza e ritardo di gruppo.*





- Emanuel A.P. Habets, *Room Impulse Response Generator*, September 20, 2010
- John B., Allen and David A. Berkley, *Image method for efficiently simulating small-room acoustics*, Acoustics Research Department, Bell Laboratories, Murray Hill, New Jersey, 1978
- Timo I. Laakso, Vesa Valimaki, Matti Karjalainem, Unto K. Laine, *Tools for fractional delay filter design*, Ieee Signal Processing Magazine, 1996
- V. Valimaki, T.I. Laakso, *Principles of fractional delay filters*, Ieee International Conference of Acoustics, Speech and Signal Processing (ICASPP'00), Instambul, Turkey, 2000
- Matlab Toolbox ROOMSIM, by Douglas R. Campbell
- Manuale online MathWorks