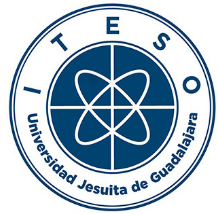
**INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE OCCIDENTE**



**MAESTRÍA EN CIENCIA DE DATOS**

**REPORTE SÍNTESIS DE ARTÍCULOS:**

**Room Response Equalization, Designing Audio Equalization Filters by Deep Neural Networks**

Director: Dr. Jorge Arturo Pardiñas Mir

Presenta: Ing. Adrián Ramos Pérez

adrian.ramos@iteso.mx

17 de noviembre de 2021

TABLA DE CONTENIDO

[Síntesis de “Room Response Equalization-A Review” 3](#_Toc88509212)

[Respuesta al Cuarto y su percepción 4](#_Toc88509213)

[Invertibilidad de la Respuesta del Cuarto 4](#_Toc88509214)

[Síntesis “Designing Audio Equalization Filters by Deep Neural Networks” 5](#_Toc88509215)

[Referencias 6](#_Toc88509216)

# Síntesis de “Room Response Equalization-A Review”

En el artículo “Room Response Equalization” se trata el estado del arte de la ecualización de la respuesta de un cuarto, detallándose que el área tiene aproximadamente 40 años (más a la fecha presente) de estudiarse, así como los métodos para abordarlo que han ido evolucionando con el tiempo y las tendencias más recientes.

Algunos puntos clave son:

* Ecualización de la respuesta de un cuarto trata de mitigar efectos indeseados como resonancias y reflexiones en el sistema conjunto cuarto-sistema de reproducción de audio.
* El sistema de reproducción de audio también puede inducir defectos no deseados como no linealidades, extensión de la banda de frecuencias entre otros.
* La idea parte de obtener la **función de transferencia** de un cuarto a través de obtener su **respuesta al impulso** e invertirla para llegar así a un ecualizador a la medida.
* Hay muchos términos para el mismo fin, entre ellos: “Room equalization”, “Room correction”, “Room compensation”, “Room inversión”, “Room dereverberation”, “Dereverberation”, “Reverberation reduction”, etc. (Es importante conocer la mayoría para unificar la investigación en el tema.)
* Algunos métodos son: mínimos cuadrados (no paramétrico), inversión directa de la respuesta en frecuencia (no paramétrico), modelado ARMA (auto-regressive moving average, paramétrico), y control del decaimiento temporal en bajas frecuencias (paramétrico) entre otros, cuyo objetivo es ecualizar la **magnitud** de la RTF.
* También se pueden clasificar estos métodos con respecto a si la fase es mínima o mezclada, actuando además de en magnitud, en el componente de la RTF que tiene **exceso de fase**.
* El artículo clasifica a los ecualizadores en dos categorías: punto único y multipunto según su relación de entradas y salidas:

|  |  |  |
| --- | --- | --- |
| **ROOM RESPONSE EQUALIZERS** | | |
| **Relación entradas/salidas** | **Enfoque** | **Descripción** |
| SISO  (una entrada – una salida) | Single point | Estiman el filtro ecualizador con base a **un único punto de medición de la función de transferencia**. Es válido para una zona limitada y pequeña (del tamaño de una fracción de la longitud de onda acústica) alrededor del punto medición. |
| MISO  (múltiples entradas – una salida) |
| MIMO  (múltiples entradas – múltiples salidas) | Multi point | Usan múltiples mediciones de la función de transferencia en diferentes puntos fijos en un tiempo dado para estimar el ecualizador, lo que agranda la zona de efectividad del ecualizador. |

* La RTF varía significativamente dependiendo la posición en el cuarto y el tiempo, por lo que el cuarto se puede considerar como un sistema “débilmente no estacionario”.

## Respuesta al Cuarto y su percepción

Las propiedades acústicas del entorno determinan la respuesta del cuarto y por lo tanto la percepción humana, lo que conlleva conocimiento de la percepción humana y la psico acústica para un correcto análisis de la respuesta al impulso y diseño del ecualizador.

Componentes de la respuesta al impulso:

* Sonido directo:
* Reflexiones tempranas:
* Reflexiones tardías:
* **El contenido espectral del sonido directo y sus reflexiones es diferente**.
* Las paredes absorben típicamente frecuencias altas de las reflexiones, dejando así las reflexiones tardías o reverb con mucha menor energía en los agudos.
* Las muescas que aparecen en la respuesta son determinadas por los patrones de interferencia causados por el sonido directo y sus reflexiones, se vuelven más densos entre más aumenta la frecuencia.
* Arriba de la frecuencia de Schroeder la respuesta en frecuencia se vuelve muy irregular.

La percepción de las altas frecuencias en específico se ve afectada por la resolución del sistema auditivo humano, la cual es no lineal ni uniforme, siendo esta casi logarítmica. Por ello se utilizan escalas de frecuencia psicoacústica como lo son:

* Escala de Mel
* Escala de Bark (tasa de banda crítica)
* Escala ERB (Ancho de banda rectangular equivalente)

Se debe considerar en el diseño del ecualizador, la baja sensibilidad de la escucha humana a picos y valles a altas frecuencias.

Existe el enmascaramiento entre sonidos, donde un sonido agudo impide que se escuche otro que se escucharía aisladamente, le hay simultánea y no simultáneamente.

El enmascaramiento es más fuerte si el contenido espectral del sonido directo y el de sus reflexiones coincide.

## Invertibilidad de la Respuesta del Cuarto

El primer paper data de 1979 y se atribuye a Neely y Allen, quienes mostraron que, si la reflectividad de una pared es menor al 36% la RTF es de fase mínima y por lo tanto invertible, si es mayor (70 a 90%) la respuesta no es de fase mínima, aunque aun así puede ecualizarse la parte de fase mínima de la respuesta factorizando la función de transferencia del cuarto en un producto de un término de fase mínima y un filtro estable pasa-todo .

Tomando el inverso de la transformada z del recíproco del espectro de .

De lo anterior surgen una serie de problemas que pueden afectar y deben ser tratados, los más importantes:

* Si la respuesta del cuarto no es de fase mínima, **no se puede implementar una respuesta inversa con una sola fuente de sonido**, esto porque la respuesta es inestable o acausal.
* La respuesta depende fuertemente del posicionamiento de los reproductores y el micrófono de medición.
* Una ecualización exacta sólo es posible en un lugar y la zona ecualizada es una fracción de la longitud de onda acústica. (En frecuencias altas puede ser menor a 18 cm, la distancia inter-aural de los oídos).
* La respuesta del cuarto varía lentamente en el tiempo afectada por la humedad y la temperatura.
* El ecualizador debe preservar el roll-off natural de los parlantes en frecuencias graves y agudas.

# Síntesis “Designing Audio Equalization Filters by Deep Neural Networks”

El artículo en su primer apartado también menciona las reflexiones y reverberaciones indeseadas en un entorno donde se reproduce audio, haciendo así hincapié en la necesidad de ecualizar este para atenuar o mitigar esos efectos indeseados y mejorar la calidad de la reproducción y escucha del audio, enfocándose particularmente en el escenario de la cabina de un automóvil con un gran número de bocinas. Se hace mención de que la respuesta al impulso en un determinado punto es la suma de la señal, sus reflexiones y así mismo, se ve condicionada por la respuesta en frecuencia de los reproductores de audio, lo que se intenta atacar aplicando un filtro lineal a la señal antes de ser transducida y reproducida por los parlantes, invirtiendo la respuesta al impulso.

Se hace mención de la complejidad generada por el número de fuentes y micrófonos de medición, por lo que, si bien existen enfoques de optimización lineal y algoritmos de inversión, entra la **necesidad de abordar nuevas técnicas no lineales** como son los algoritmos evolutivos, aprendizaje máquina y redes neuronales. Como ejemplo referencia casos de uso de ***optimización por enjambre de partículas***, donde se obtienen mejores resultados que los métodos de *mínimos cuadrados promediados* y *mínimos cuadrados recursivos*, como antecedente uno donde se usa este algoritmo para obtener polos y ceros óptimos para un filtro IIR, así como otro donde se hace uso de algoritmos genéticos para ecualización de canal adaptativa para reducción de interferencia inter-simbólica en un canal. Sin embargo, se toca un punto clave que es **las diferencias de implementación contra los sistemas de comunicaciones**, donde se aplica el ecualizador al final de la cadena, contra el caso de uso de reproducción de audio, donde se utiliza antes de ser reproducida una señal, teniendo así la problemática de asegurar una buena calidad de audio para múltiples puntos de escucha, lo que difiere a los sistemas de telecomunicación que buscan minimizar la tasa de error por símbolo. Otro punto clave es que en el audio se asume la **invarianza en el tiempo**, por lo que se ataca la respuesta al impulso en un cuarto de manera estática.

Entrando en materia, se hace referencia a artículos que utilizan los siguientes algoritmos:

* *Time Delay Neural Network* para ecualización tomando como entrada la misma entrada retrasada una unidad de tiempo y como salida del sistema la señal recapturada del micrófono.
* *K-Nearest Neighbour* para ecualización de timbre basado en preferencias del usuario, usando esto como data de entrenamiento.
* *Dilated Residual* *Network* para automatizado de ecualización de la resonancia. *Convolutional Neural Network* para una ecualización punto a punto, sin conocimiento previo de los parámetros del filtro (ganancia, frecuencia de corte, factor de calidad, etc.).
* *Particle Swarm Optimization* para diseño de filtro IIR **obteniendo los parámetros del filtro con una respuesta plana en magnitud y fase lineal.**
* *Gravitational Search Algorithms* para
* *Artificial Inmune Algorithm* contrastado con *Touring Ant Colony Optimization* y *Tabu Search* para diseño de filtro IIR

Una aportación valiosa que se hace es la referencia 19, la cual hace la comparación en desempeño de arquitecturas de redes neuronales:

* Radial Basis Function (RBF)
* General Regression Neural Networks (GRNN)
* Radial Basis Exact (RBE)
* Back Propagation Neural Network (BPNN)
* Multilayer Perceptron (MLP)

# Referencias

1. Cecchi, S; Carini, A.; Spors, S. Room Response Equalization-A Review. Appl. Sci. 2018,8,16.
2. Pepe, G. Gabrielli, L, Squartini, S, Cattani, L. Designing Audio Equalization Filters by Deep Neural Networks. Appl. Sci. 2020, 10, 2483.