

Atténuation du bruit dans des signaux audios a travers l'implémentation de méthodes classiques et une méthode hybride

1st José Miguel Galindo Barco
Ecole d'ingénieurs
Génie de systèmes
Université EAFIT
jmgalindob@eafit.edu.co

2nd Juan José Tamayo Acevedo
Ecole de sciences
Génie mathématique
Université EAFIT
jjtamayoa@eafit.edu.co

3rd Salomón Cardeño Luján
Ecole de sciences
Génie mathématique
Université EAFIT
scardenol@eafit.edu.co

Résumé

L'atténuation du bruit est un domaine d'étude actif dans des milieux comme le traitement digital et analogique des signaux. Dans ce domaine il existe une grande variété d'algorithmes et/ou méthodes classiques et modernes qui sont utilisés dans ce but. Dans ce travail de recherche, nous réaliserons une étude théorique du filtre analogique passe-bas avec une approximation de Butterworth, le filtre de Weiner et la méthode dite hybride de DSP/Deep Learning RNNbruit. Puis nous développerons une analyse de ces trois méthodes où nous comparerons les résultats de signaux lorsque nous filtrons un signal propre, un signal avec addition de bruit blanc, un signal avec addition de bruit rose et un signal avec addition de bruit rouge. Les enregistrements audios résultant de chaque méthode de filtrage seront comparés qualitativement et quantitativement, d'où la nécessité de tenir en compte la perception auditive du récepteur. Finalement, nous proposerons des modifications sur la mise en œuvre ou implémentation de ces méthodes en considérant le contexte et le bon fonctionnement dans ce dernier.

I. INTRODUCTION

HAY QUE PONER ALGO AQUÍ

II. CONCEPTS FONDAMENTAUX

Signal : un signal est la représentation physique de l'information, qu'il convoit de sa source à sa destination.

Bruit : Un bruit correspond à tout phénomène gênant la transmission ou l'interprétation d'un signal

Traitement du signal : Le traitement du signal est une discipline qui étudie des méthodes de traitement, d'analyse et d'interprétation des signaux. Pour cela, dans ce domaine on utilise des méthodes tels que le contrôle, le filtrage, la compression et la transmission de données, la réduction du bruit, la déconvolution, la prédiction, l'identification, la classification.

Filtrage : On appelle "filtrage" toute application qui, à un signal $f_e(t)$, dit "signal d'entrée", associe un signal $f_s(t)$, dit "signal de sortie", tel que le module du spectre de $f_s(t)$ soit inférieur ou égal au module du spectre de $f_e(t)$, pour ω réel quelconque.

A. Signal sonore et note audio

Le son ou signal sonore est une perturbation mécanique d'un état d'équilibre, qui se propage à travers d'un milieu solide élastique. Il est possible de le définir subjectivement comme ceux qui sont perçus par l'ouïe, mais cela reviendra à restreindre la définition même d'un signal. De même, la lumière est perçue par la vue mais il existe certaines longueurs d'ondes lumineuses qui ne sont pas perçues par l'œil humain. Il existe donc des domaines de fréquences des signaux sonores qui ne sont pas perçus par l'ouïe humaine.

Dans un premier temps, en s'appuyant sur le point de vue physique, les ondes sonores se propagent à travers l'air ou autres milieux (solides et élastiques) sous forme de longueur d'ondes, où la vibration mécanique qui décrit l'onde se propage dans la même direction que la propagation de celle-ci. L'air peut être compris comme un milieu élastique sur lequel les particules, ou tout simplement les couches d'air, ont un comportement semblable à un ressort, c'est-à-dire qu'elles "se poussent" et "s'attirent" les unes avec les autres.

Plus clairement, une onde sonore est une variation périodique ou sinusoidale de la pression d'un milieu pendant un temps précis dans un endroit précis.

Dans la figure ci-dessous, dans un premier temps nous pouvons voir l'air dans son équilibre (figure xA), les variations dans l'air lorsqu'une onde sonore pure se déplace dans ce milieu (figure xB) et l'allure de l'onde qui se déplace dans l'air.

B. Caractéristiques physique et description mathématique d'un signal sonore

Il est important de remarquer que dans figure précédente, plus précisément dans la partie C de cette figure, s'agit d'une autre représentation du signal sonore illustré dans la partie B. En effet, dans cette partie nous pouvons voir une courbe sinusoïdale, cette courbe représente la variation de la pression du signal sonore, celui-ci se répète de façon périodique. La distance entre 2 pics de la courbe est nommée longueur d'onde et se représente par la lettre λ .

Lorsqu'un signal sonore ou onde sonore se déplace dans l'air, la longueur d'onde prend un certain temps pour aller d'un point A à un point B de l'espace, cette durée de temps est appelée période et notée T.

De plus, pendant un intervalle de temps équivalent à une seconde, un certain nombre de longueurs d'onde passe d'un point A à un point de l'espace. Ce phénomène est connu comme fréquence du signal. Dans d'autres mots, la fréquence est le nombre de périodes par unité de temps ce qui correspond à l'inverse de la période : $f = 1/T$ ou f est la fréquence en Hertz (Hz ou s^{-1}) et T la période en seconde (s).

Il est important de remarquer que les signaux sonores qui ont de hautes fréquences, ont des périodes courtes, alors que les signaux sonores qui ont des basses fréquences, ont de longues périodes. De plus, l'intervalle de fréquences que l'être humain peut percevoir se trouve entre les 20Hz et les 20kHz. Il existe une propriété physique qui permet de classer les signaux sonores en fonction de la perception physiologique des êtres humains, cette propriété est connue comme le ton. En plus, les signaux ont une vitesse de déplacement, celle-ci est connue comme vitesse de l'onde (S), elle est obtenue par une relation entre la fréquence ou la période et la longueur d'onde, comme montré ci-dessous :

$$S = f * \lambda = \frac{\lambda}{T}$$

Le déplacement d'une onde sonore dans une dimension (signal sonore plat) est décrit mathématiquement par l'équation général du mouvement des ondes, qui peut être écrit de la façon suivante :

$$y(x, t) = A \sin(2\pi(ft - \frac{x}{\lambda}))$$

L'amplitude (A) d'une onde correspond à la hauteur maximale atteinte par l'onde par rapport à sa position n au repos.

De plus, grâce à l'amplitude nous pouvons déterminer l'intensité de l'onde, qui est perçue par l'ouïe sous le nom de volume. L'intensité acoustique (I), est définie comme le taux de transmission d'énergie par unité d'aire perpendiculaire à la direction de propagation de l'onde. La relation existante entre l'intensité et l'amplitude est la suivante :

$$I = \frac{A^2}{2\rho S}$$

Avec :

- ρ : densité de l'air à l'équilibre (en kg/m^3).
- S vitesse de l'onde (en m/s).

L'intensité I a pour unité les watt par mètre carré (W/m^2). Sous des "conditions atmosphériques standards, on a :

$$\rho = 10^5 Pa = 10^5 \frac{N}{m^2}$$

L'amplitude minimum de variation de pression que l'ouïe humaine peut percevoir est de $10^{-5}Pa$ et l'amplitude maximale est de $10Pa$.

C. Audition et perception du son chez les êtres humains

D'après ce qui a été décrit précédemment, il est convenable de remarquer les faits suivants :

- Entre 20Hz et 20kHz se trouve l'intervalle de fréquence perçue par l'être humain.
- L'amplitude d'une onde sonore permet de déterminer l'intensité de l'onde.
- L'amplitude minimum de variation de pression que l'ouïe humaine peut percevoir est de $10^{-5}Pa$ et l'amplitude maximale est de $10Pa$.

D. Echelles des décibels

L'échelle des décibels est une échelle qui permet de classer les ondes sonores en fonction de leur intensité sonore. Cette échelle est décrite par l'équation suivante :

$$I_{dB} = L = 10 \log\left(\frac{I}{I_0}\right)$$

Avec L qui représente les décibels, I l'intensité de l'onde et $I_0 = 10^{-12} W/m^2$ est l'intensité de référence.

Plus précisément, L'échelle des décibels est logarithmique, ce qui signifie qu'une augmentation du niveau sonore de $3dB$ représente déjà un doublement de l'intensité sonore. Par exemple, le volume d'une conversation normale peut être d'environ $65dB$ et, pour quelqu'un qui crie, ce chiffre peut atteindre environ $80dB$. La différence est seulement de $15dB$, mais le cri représente une intensité trente fois supérieure.

Le décibel (dB) est une unité utilisée pour mesurer l'intensité des sons et celle d'autres grandeurs physiques. Un décibel équivaut à un dixième de bel (B), une unité qui doit son nom à Graham Bell, l'inventeur du téléphone. Son échelle logarithmique permet de représenter le spectre auditif de l'être humain dans son ensemble.

Le champ auditif de l'être humain (en vert) est limité par une courbe qui nous fournit la limite inférieure et une autre courbe qui nous donne la limite supérieure de la perception sonore. A chaque fréquence, entre $20Hz$ et $20kHz$, le seuil de notre sensibilité est différent. L'intervalle le plus large de perception a lieu aux alentours des $2kHz$ et commence à partir des $0dB$. Dans cet intervalle dit "moyen" les dynamiques sensorielles sont les plus aptes possibles et peuvent avoir un équivalent de $130dB$.

E. Tonalité pure

Lorsque nous utilisons une fonction comme celle de la vitesse de déplacement d'une onde, nous pouvons dire que cette fonction représente une tonalité pure avec une fréquence f . Ce nom dérive du fait que cette fonction décrit une onde sonore avec une seule composante en fréquence, c'est-à-dire qu'il s'agit d'un cas idéal, parce que en faisant analogie avec la lumière, dans la nature nous ne trouverons jamais une lumière monochromatique (d'une seule couleur, c'est-à-dire d'une seule fréquence), nous ne trouverons jamais une onde sonore avec cette caractéristique.

Afin de donner une définition plus précise d'une tonalité pure, nous devons nous appuyer sur la définition donnée en psychoacoustique. En effet, en psychoacoustique, une tonalité pure, ou encore son pur ou note pure (en anglais : pure tone) est un son avec une forme d'onde sinusoïdale ; c'est-à-dire une onde sinusoïdale de n'importe quelle fréquence, phase et amplitude. En audiologie clinique, les tons purs sont utilisés pour l'audiométrie tonale (en) afin de caractériser les seuils d'audition à différentes fréquences.

F. Descomposición del sonido en tonos puros

Proprement dit, toute onde sonore résulte de l'addition de plusieurs tonalités pures de fréquence différente. La quantité de

Formellement, cualquier sonido f es una suma de tonos puros a diferentes frecuencias. La cantidad de cada frecuencia requerida para formar el sonido f es el contenido frecuencial de f . Cualquier sonido puede ser reconstruido partiendo de su contenido frecuencial.

Una implicación de lo anterior es que, en general, cualquier sonido puede ser considerado como una función, y por más complejo que sea el sonido, puede ser representado por la suma de tonos puros o funciones sinusoidales (senos y cosenos) lo cual no es sorpresa que esté relacionado con el concepto matemático de series de Fourier.

Adicionalmente, lo anterior también implica que el sonido f puede ser almacenado al guardar simplemente su contenido frecuencial, como una alternativa de guardar a f como tal.

G. Espectro del sonido

El espectro del sonido muestra las diferentes frecuencias presentes en un sonido, es decir, muestra el contenido frecuencial que caracteriza el sonido en particular y con qué amplitud está presente, respectivamente. Usualmente es presentado como una gráfica de frecuencia versus intensidad, presión o potencia en dB . Cabe aclarar que si es evidente que hay una frecuencia dominante (de mayor amplitud) se le llama frecuencia fundamental y al resto armónicos. A continuación observe el ejemplo de una flauta que toca la nota G4 equivalente a aproximadamente $400Hz$; note que esta es la frecuencia fundamental o predominante con mayor amplitud y el resto son sus armónicos. Si se comparara esta misma nota en otro instrumento también se obtendría la misma frecuencia predominante, pero sus armónicos serían distintos y esto es lo que hace que los seres humanos percibamos cada instrumento de forma distinta aún tocando la misma nota. Esta percepción auditiva se conoce como el timbre.

H. Sonido digital

L'échantillonnage d'un signal continu est l'opération qui consiste à prélever des échantillons du signal pour obtenir un signal discret, c'est-à-dire une suite de nombres représentant le signal, dans le but de mémoriser, transmettre, ou traiter le signal.

L'échantillonnage intervient dans l'opération de conversion analogique-numérique, par exemple dans un dispositif de numérisation du son ou de l'image. Un autre exemple d'échantillonnage est celui que l'on fait pour obtenir la représentation graphique d'une fonction à une ou deux variables. D'une manière générale, l'échantillonnage intervient dans toute opération de conversion continu/discret.

Le théorème de l'échantillonnage de Shannon, qui permet de savoir à quelle fréquence minimale il faut échantillonner un signal pour ne pas perdre l'information qu'il contient.

Soit $u(t)$ une fonction représentant un signal continu. On considère un échantillonnage périodique défini par :

où k est un entier. T_e est la période d'échantillonnage. $f_e = 1/T_e$ est la fréquence d'échantillonnage.

Le théorème de Shannon ([1]) concerne les signaux dont le spectre possède une fréquence maximale f_{max} , que l'on appelle des signaux à bande limitée. Par exemple, si $u(t)$ est un polynôme trigonométrique, la fréquence maximale est celle de la plus grande harmonique.

Théorème de Shannon : pour que le signal puisse être entièrement reconstruit à partir des échantillons, il faut et il suffit que :

La fréquence d'échantillonnage doit être strictement supérieure à deux fois la plus grande fréquence présente dans le spectre du signal continu (condition de Nyquist-Shannon). Si cette condition est vérifiée alors :

où la fonction sinus cardinale est définie par :

Cette relation montre que le signal peut être reconstruit à partir des échantillons, ce qui signifie que toute l'information présente dans le signal original est conservée dans les échantillons. Nous verrons plus loin comment l'opération de reconstruction est effectuée en pratique.

La moitié de la fréquence d'échantillonnage est appelée la fréquence de Nyquist f_n et la condition de Nyquist-Shannon s'écrit donc $f_{max} < f_n$.

Lorsque la condition n'est pas vérifiée, on dit qu'il y a sous-échantillonnage. On parle de sur-échantillonnage lorsque la fréquence de Nyquist est beaucoup plus grande que f_{max} .

Las mediciones normalmente se les conoce como *muestras*. El tiempo entre mediciones sucesivas se le llama *periodo de muestreo* y usualmente es denotado por T_s . La longitud del vector usualmente se asume como N , e indexado desde 0 a $N - 1$. El proceso de convertir una onda de sonido (audio análogo) a sonido digital se le llama cuantización del audio en donde a cada una de las muestras anteriores se les asocia un valor de amplitud (usualmente entre -1 y 1). El rango de posibles niveles de amplitud, en otras palabras, el número de particiones equiespaciadas del intervalo de valores de amplitud $[-1, 1]$ se le conoce como *profundidad de bits*, por ejemplo

$$8 - bit = 2^8 = 256 \text{ posibles valores}$$

$$16 - bit = 2^{16} = 65536 \text{ posibles valores}$$

Para visualizar el efecto de la cuantización del audio y la influencia de la frecuencia de muestreo y la profundidad de bits observe la figura a continuación.

Así como lo sugiere la figura anterior, la frecuencia de muestreo y la profundidad de bit juegan un papel fundamental en la determinación de la calidad del audio resultante.

I. Almacenamiento de un sonido digital en el computador

En un computador usualmente se utiliza el método de modulación por impulsos codificados (PCM, Pulse-code modulation en inglés) para transformar la señal análoga en una secuencia de bits (señal digital). Este método fue inventado por Alec Reeves en 1937 y es la forma estándar de audio digital en las computadoras. (ref)

Para que un computador pueda reproducir un sonido digital en el computador las muestras obtenidas deben de ser guardadas en un archivo o en la memoria del computador; en el caso del PCM se almacena en formato RAW o en otros como WAV, AIFF, AU, entre otros. Sin embargo, como los archivos de audio son grandes, lo normal es trabajar con formatos comprimidos que reducen el tamaño del archivo : (ref)

- **Formatos con pérdida (lossy)** : son formatos que comprimen el archivo de audio original al eliminar partes de él de acuerdo a un algoritmo de compresión específico, por lo que al descomprimir el archivo no se obtiene el original. Algunos de ellos son : MP3, Opus, Vorbis, WMA, entre otros.
- **Formatos sin pérdida (lossless)** : son formatos que comprimen el archivo de audio original al organizar la información de forma eficiente, por lo que al descomprimir el archivo se obtiene el original. Algunos de ellos son : FLAC, APE, WV, ALAC, entre otros.

J. Bruit et types de bruit

Comme vu précédemment, le bruit dans un signal sonore correspond à une perturbation de celui-ci entraînant des pertes de d'information. Il existe plusieurs types de bruit qui ont leurs propres domaines de fréquences et qui altèrent ou modifient l'information contenue dans un signal de façon différente. En effet, entre les différents types de bruits nous pouvons remarquer 2 groupes très importants : Les bruits dit naturels et les bruits dit industriels.

Dans un premier temps, en regardant les bruits naturels nous pouvons nous rendre compte que ces bruits ne sont pas produits de façon artificielle et existe par différentes raisons. Dans cette famille de bruit nous pouvons trouver le bruit ambiant qui correspond au bruit total existant dans une situation donnée pendant un intervalle de temps donné, ainsi que les bruits qui le compose telles que le bruit résiduel et le bruit particulier. De plus, nous trouvons aussi des bruits dit colorés telles que

le bruit blanc ou le bruit rose, qui feront objet de notre étude de filtrage. Tout d'abord, il faut remarquer que le bruit blanc est un bruit composé par la somme de tous le bruit coloré, c'est à dire l'addition en fréquence du bruit rose, bruit rouge ou brownien, bruit bleu ou azur, bruit violet et bruit gris. Chacun d'entre eux aillant un domaine de fréquence particulier et des caractéristique propre. Par exemple :

- **Le bruit blanc** es ruido aleatorio que se caracteriza por tener densidad espectral plana, es decir, su componente frecuencial está compuesto por todas las frecuencias a una misma intensidad. Su percepción se asemeja a escuchar sonidos de la naturaleza como estática de televisores o radios.
- **Le bruit rouge ou brownien** correspond dans une première approximation, dans les domaines qui utilisent des définitions précises, la terminologie « bruit rouge », « bruit brownien » ou « bruit brun » fait référence au son ayant une puissance sonore qui décroît de 6 dB par octave lorsque la fréquence augmente (densité proportionnelle à $1/f$) sur un intervalle de fréquence n'incluant pas de DC (Qui dans un sens général, n'inclut pas de composante constante, ou de valeur pour $f=0$).
- D'un autre point de vue, dans les domaines qui utilisent des définitions plus approximatives, le « bruit rouge » correspond à tout son dont la densité de puissance diminue lorsque la fréquence augmente³.
- Plus exactement Le bruit rouge, peut être obtenu en utilisant un algorithme simulant le mouvement brownien ou par intégration mathématique du bruit blanc.
- **Le bruit rose** se caracteriza por su densidad espectral decreciente de -3dB/Octava . Su percepción se asemeja a escuchar sonidos de la naturaleza como lluvia constante.
- **Le bruit bleu** qui correspond à un bruit qui augmente sa puissance sonore de 3 dB par octave lorsque la fréquence augmente (densité proportionnelle à f), et ce jusqu'à une fréquence infinie.
- Dans le domaine de l'informatique graphique, le terme « bruit bleu » est parfois utilisé d'une façon plus approximative pour décrire tout son de puissance sonore minimale à basse fréquence et ne présentant aucun pic lorsque la fréquence augmente (croissance constante).
- Il est convenable de connaitre et différencier aussi le bruit rose des autres types de bruits coloré, dans la mesure que ce type de bruit fait partie de notre étude. Donc le bruit rose, ce type de bruit est caractérisé par une puissance égale sur des bandes proportionnelles en largeur. Une bande correspond en fait à un changement de fréquence (hausse ou baisse à exprimer en pourcentage de l'une des extrémités de l'intervalle). Par exemple, la puissance d'un bruit rose est la même sur les intervalles allant de 40 à 60 Hz et de 4 000 à 6 000 Hz car ces intervalles sont proportionnels (ils correspondent à une hausse de 50 pourcent de la fréquence).

Dans un deuxième temps, nous voyons l'existence d'une deuxième famille de bruits que nous choisisant de nommé des bruits industriels. En effet, comme son nom l'indique les bruits industriels sont des perturbations sonores produits par l'industrie ou par des effet industrie. Dans cette famille nous trouvons des bruits tels que le bruit de route qui est un bruit normalisé. Il est une référence pour le bruit des trafics routiers et ferroviaires. Son spectre est enrichi en basses fréquences et appauvri dans les aigües par rapport à un bruit rose. Mais aussi nous trouvons des bruits comme les suivants :

- **Le bruit d'impact** : c'est le bruit transmis par une paroi mise en vibration par un choc (bruit de pas, déplacement de meubles, chute d'objet, enfoncement d'un clou dans un mur...).
- **Le bruit aérien** : c'est le bruit propagé dans l'air (bruit de voix, bruit de télévision, bruit de circulation...).
- **Le bruit solide** : c'est le bruit propagé dans les milieux solides comprenant (le bruit d'impact transmis par les éléments solides, le bruit d'équipement (chaufferie, ascenseurs, ...).

III. MÉTHODES

Le but de ce projet est de implementer le filtrage dans les notes audios dans l'objectif d'attenuer le bruit present. Pour cela nous allons implementé le filtrage analogique avec une approximation de Butterworth, la méthode RNN et un autre qui reste définir.

A. Filtre de Butterworth

B. Filtre de Wiener

Le filtre de Wiener est une des meilleur filtres linaires de minimum carré. Il peut être utilisé pour prédire, estimer, interpoler, filtrer un signal et le bruit, entres autres. Pour créer ce type de filtre, il faut connaitre avoir une connaissance préalable du signal que le système aura comme entrée, c'est à dire le signal qui sera filtré. Le plus grand problème c'est que cette information est difficile à obtenir, c'est pour cela que les filtres adaptatifs sont très importants, parce qu'ils utilisent les données d'entrée pour pouvoir déterminer les données statistiques requis. En général, on a un signal $f(k)$, une réponse désiré $d(k)$ et un filtre linéaire de réponse impulsionnelle $h(k)$. Ce filtre est alimenté par $f(k)$ et produit une sortie $g(k)$. La différence entre le signal de sortie $g(k)$ et le signal désiré $d(k)$, est connue comme erreur d'estimation $e(k)$, la figure suivante illustre ce fait :

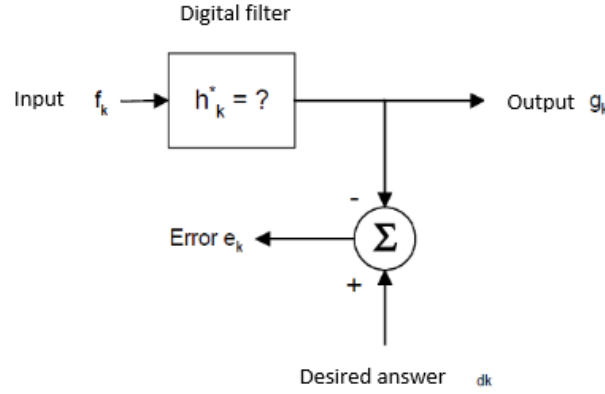


FIGURE 1. Filtre de Wiener

L'objectif du filtre de Wiener est de déterminer la réponse impulsionnelle $h(k)$ dans le but de produit une erreur $e(k)$ le plus petit possible. Le critère utiliser pour réaliser cela est la minimalisation du valeur quadratique moyen de l'erreur.

Le filtre digital montrer dans l'image a un signal d'entre et produit un signal de sortie. Le filtre sera un filtre de Wiener si sa réponse impulsionnelle est choisie dans l'objectif de minimiser l'erreur quadratique moyen. L'erreur est définie comme la différence entre la sortie et la réponse désiré et est décrit par l'équation suivante :

$$e_k = d_k - g_k$$

Lorsque nous travaillons avec le filtre de Wiener, généralement la réponse désirée existe dans la forme théorique. Cependant, dans ce cas on travaille avec un réponse désiré connue. La réponse impulsionnelle du filtre de Wiener est obtenue

Si nous appliquons le carré a l'équation de l'erreur, nous avons :

$$e_k^2 = d_k^2 + g_k^2 - 2d_k g_k$$

Et on sait que :

$$g_k = \sum_{l=0}^{\infty} f_{k-l} h_l$$

Si on substitue dans l'équation de l'erreur, on a :

$$e_k^2 = d_k^2 + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m f_{k-l} f_{k-m} - 2 \sum_{l=-\infty}^{\infty} f_{k-l} h_l d_k$$

Si, on prend la valeur moyenne à chaque côté, on trouve une expression pour l'erreur quadratique moyenne

$$E[e_k^2] = E[d_k^2] + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m E[f_{k-l} f_{k-m}] - 2 \sum_{l=-\infty}^{\infty} h_l E[f_{k-l} d_k]$$

$$= \phi_{dd}(0) + \sum_{l=-\infty}^{\infty} \sum_{m=-\infty}^{\infty} h_l h_m \phi_{ff}(l-m) - 2 \sum_{l=-\infty}^{\infty} h_l \phi_{fd}(l)$$

Avec ϕ_{dd} et ϕ_{ff} l'autocorrélation et ϕ_{fd} la corrélation croisée entre les signaux f et d . Si on dérive en fonction de h et on cherche la solution de l'équation lorsque celle-ci tend vers 0, on a :

$$\sum_{l=-\infty}^{\infty} h_l^* \phi_{ff}(j-l) = \phi_{fd}(j)$$

i

Cette équation est l'équation de Wiener-Holpf, qui peut être écrit de la façon suivante, grave au produit de convolution :

$$h_l^* * \phi_{ff}(k) = \phi_{fd}(k)$$

Et si on applique la transformé en Z a chaque partie de l'équation, on a :

$$H^*(z) \phi_{ff}(z) = \phi_{fd}(z) \text{ ó } H^*(z) = \frac{\phi_{fd}(z)}{\phi_{ff}(z)}$$

Avec la solution de Weiner, on peut trouver la fonction de transfert du filtre $H(z)$ a partir de la transformé en Z de la fonction d'autocorrélation d'un signal d'entrée, de la corrélation croisée du signal d'entrée et la réponse désirée. Si on remplace cette équation dans l'expression de l'erreur, on a la valeur minimal MSE :

$$E(e_k^2)_{min} = \phi_{dd}(0) - \sum_{l=-\infty}^{\infty} h_l^* \phi_{fd}(l)$$

C. RNNbruit : Méthode hybride DSP/Deep Learning

La méthode RNNbruit est une méthode créée par Jean-Marc Valin de Mozilla Corporation. Il s'agit d'une méthode hybride pour atténuer le bruit, en termes généraux elle fonctionne de la façon suivante : elle utilise le Deep Learning dans des cas où la réduction du bruit demande réglage ou une configuration minutieuse, alors que pour les cas qui ne le demandent pas, la méthode emploie des procédés d'élimination du bruit classiques. Pour mieux comprendre le fonctionnement de ce filtre, nous allons étudier plus attentivement les concepts qui ont un rôle important dans cette méthode, tels que réseau de neurones artificiels, Deep Learning et réseau de neurones récurrentes.

1) *Réseau de neurones artificiels*: Pour pouvoir comprendre de quoi il s'agit un réseau de neurones, il est convenable de commencer par un élément de ce réseau : un neurone artificiel. Il existe plusieurs modèles de neurones artificiels, cependant, il suffit de définir le plus bas niveau, le perceptron, et le modèle le plus utilisé dans les industries modernes, le sigmoïde.

a) *Perceptron*: Les perceptrons ont été développés entre les années 1950 et 1960 par le scientifique Frank Rosenblatt, inspirée du travail réalisé par Warren McCulloch et Walter Pitts en 1943.

Un perceptron prend plusieurs entrées binaires (uniquement de 0 et de 1) x_1, x_2, \dots et produit une unique sortie binaire, par exemple :

Dans ce cas le perceptron a 3 entrées, mais en général il peut avoir une quantité arbitraire d'entrées, Rosenblatt propose une règle simple pour calculer la sortie. En effet, il introduit des poids w_1, w_2, \dots qui représentent des nombres réels illustrant l'importance de chaque entrée respectivement avec la sortie. La sortie du neurone (0 ou 1) résulte de la somme pondérée des entrées, si cette somme est supérieure à une valeur limite, la sortie sera égale à 1, mais si elle est inférieure à cette valeur limite la sortie sera 0.

La valeur limite est aussi un nombre réel. Il s'agit d'un paramètre du neurone. Par notation, on simplifie la description précédant en considérant w et x comme des vecteurs des poids et des entrées, respectivement, donc la somme pondérée devient :

Finalement, on définit le biais noté b comme , alors on a :

Il faut comprendre le biais comme un objet de mesure qui permet de déterminer à quel point il est facile de faire que la sortie du perceptron soit égale à 1. En termes biologiques, le biais est une mesure qui permet de savoir à quel point il est facile de faire que le perceptron se déclenche ou s'active.

Même si les perceptrons sont très utilisés pour modéliser une grande variété de problèmes, ils ont un grand désavantage. En effet, un petit changement dans le vecteur des biais ou dans celui des poids, peut entraîner un grand changement dans la sortie, ce qui peut être problématique, même dans le cas d'un algorithme de bas niveau d'apprentissage.

b) *Neurones Sigmoïdes*: Elles se représentent de la même façon que les perceptrons :

Elles ont un nombre arbitraire d'entrées x_1, x_2, \dots , mais les valeurs des entrées prennent des valeurs comprises entre 0 et 1. Elles possèdent aussi un poids pour chaque entrée, w_1, w_2, \dots , et un biais b . Dans ce cas, les sorties ne sont pas des valeurs de 0 ou 1, mais une fonction $g(w \cdot x + b)$, où g est appelé la fonction sigmoïde ou en général, fonction d'activation, et est définie par :

Où graphiquement est définie par :

Dans ce cas, on observe que la sortie prend des valeurs entre 0 et 1, ceux qui revient à dire que la sortie peut prendre une infinité de valeurs. Usuellement, lorsqu'on emploie des sigmoïdes, l'intervalle est divisé en partitions et en fonction de la valeur obtenue, on fait une interprétation particulière de la sortie. Par exemple, en traitement d'image si on veut savoir si l'image d'entrée possède le numéro 9, on pourrait employer une cote de 0.5

Autres modèles de neurones utilisés dans le modèle La méthode RNNbruit emploie des réseaux de neurones avec de neurones sigmoïdes mais il peut employer deux autres modèles de neurones : Tanh et RELU. Le comportement et la description de ces modèles est exactement le même, avec un changement de la fonction d'activation g.

2) *Notation dans l'architecture des réseaux neuronales*: Les entrées sont encerclées et symbolisées :

Considérons un réseaux neuronal arbitraire comme celui-ci-dessous :

La colonne de gauche qui contient toutes les entrées est connue comme "couche d'entrée" (input layer en anglais). La colonne de droite qui contient toutes les sorties est connue comme "couche de sortie" (output layer en anglais). Les autres colonnes qui contiennent les neurones sont connues comme "couches cachées" (hidden layers en anglais). Dans le cas de l'image ci-dessus, on a un réseau neuronal avec 6 entrées (input features en anglais), 2 couches cachées et une sortie. Il est important de remarquer que ce type de réseau neuronal, qui est composé d'une couche d'entrées, 2 couches cachées et une couche de sortie, est connue sous le nom de réseau neuronal pré-alimenté.

3) *Réseau de neurones récurrents (Recurrent Neural Networks, RNN)*: Les réseaux de neurones récurrents sont un autre type de réseau de neurones. Il se caractérisent par le fait que la sortie donnée par une neurone est la sortie générale du réseau et au même temps, cette sortie est l'entrée pour le neurone suivante, comme l'illustre l'image ci-dessous :

Le réseau de neurones récurrent est symbolisé comme le réseau de gauche avec une boucle de retro alimentation, cependant, il est plus de comprendre son fonctionnement en la "déroulant" (unroll en anglais) comme nous pouvons voir dans le réseau de neurones équivalent de droite. Dans ce cas, les neurones sont représentés par des rectangles en vert et la lettre A, les entrées sont représentées par des cercles bleus et la lettre X avec un indice, et les sorties sont représentées par des cercles violets et la lettre h avec un indice. Il faut remarquer que dans ce cas la couche d'entrée correspond à la rangée qui contiens toutes les entrées en bleus, la couche de sortie correspond à la rangé supérieure qui contiens toutes les sorties en violet et la couche restante est la seule couche cachée.

IV. RÉSULTATS ET COMPARAISON DES MÉTHODES

Los métodos fueron implementados de forma cuantitativa y cualitativa. Para ambos casos se utilizó el software Matlab R2020a en el SO Microsoft Windows 10 Home Single Language (v. 10.0.19041). Para los casos de implementación se generó ruido blanco, rosa y café utilizando el software Audacity 2.4.2. Luego en Matlab se tomó el audio original y se crearon 3 audios adicionales al mezclar el audio original con 30% de ruido blanco, 30% de ruido rosa y 30% de ruido café, respectivamente.

A. *Résultats quantitatifs*

En la tabla I se muestra el porcentaje de ruido resultante al aplicarle 30% de ruido blanco al audio original y filtrarlo con cada uno de los métodos.

Méthodes	Blanc (30%)
Sans méthode	80.658
Wiener	73.621
Butter T	10
RNNbruit	9.9668

TABLE I
RÉSULTATS DES MÉTHODES POUR 30% DU BRUIT BLANC

En la tabla II se muestra el porcentaje de ruido resultante al aplicarle 30% de ruido rosa al audio original y filtrarlo con cada uno de los métodos.

Méthodes	Rose (30%)
Sans méthode	39.442
Wiener	65.359
Butter T	35
RNNbruit	11.493

TABLE II
RÉSULTATS DES MÉTHODES POUR 30% DU BRUIT ROSE

En la tabla II se muestra el porcentaje de ruido resultante al aplicarle 30% de ruido café al audio original y filtrarlo con cada uno de los métodos.

Méthodes	Marron (30%)
Sans méthode	53.969
Wiener	7.7496
Butter T	35
RNNbruit	7.7496

TABLE III
RÉSULTATS DES MÉTHODES POUR 30% DU BRUIT MARRON

Premièrement, en étudiant les résultats obtenus par l'implémentation du filtre passe-bas avec une approximation de Butter Worth, nous confirmant les hypothèses théoriques vues précédemment. En effet, compte tenu que l'approximation de Butter Worth permet d'avoir une précision en amplitude. Cette précision est centrée sur des basses fréquences. Nous voyons cela évidente dans le filtrage des bruits ajoutés à notre signal. Lorsque nous ajoutons du bruit blanc et nous filtrons le signal résultant, nous constatons une diminution du pourcentage du bruit présent dans le signal. Alors que lorsque nous effectuons le même procédé pour le bruit rose et le bruit marron, le pourcentage de bruit ne change pas. Donc cette méthode avec cette approximation particulière est bien plus convenable pour filtrer le bruit blanc présent dans les enregistrements audios.

Dans un deuxième temps, le cas de la méthode hybride du RNNbruit, nous observons que cette méthode est belle et bien convenable pour les différents types de bruit additionnés. En effet, nous constatons une atténuation du bruit considérable dans chacun des cas. De plus, nous remarquons que cette méthode ne gêne pas la transmission du signal sonore, en d'autres termes cette méthode lise et "purifie" le signal original. Nous pouvons expliquer ce comportement par la nature adaptative de la méthode, dû au fait de l'implémentation de réseau de neurones récurrents pour faire face au bruit non harmonique. Cependant, la partie traditionnelle de cette méthode dite hybride se charge du bruit harmonique comme nous l'avons vue précédemment dans la partie théorique.

Finalement, lorsque nous analysons les résultats obtenus en utilisant le filtre adaptatif de Wiener, nous apercevons que cette méthode de filtrage numérique a des meilleurs résultats sur du bruit blanc. Or, en effectuant une analyse auditive, nous pourrions penser que ce filtrage est bien plus efficace avec le bruit marron ou le bruit rose, nous pourrions affirmer théoriquement que c'est avec le bruit blanc qu'il est plus efficace. En effet, en employant cette méthode nous introduisons des nouveaux types de bruits que nous ne considérerons pas dans notre analyse. Par exemple, dans le filtrage de l'enregistrement audio avec du bruit marron, nous pouvons avoir l'impression que le résultat est nettement plus compressible qu'avec du bruit blanc. Cependant, en

regardant les résultats de l'analyse nous nous rends compte que l'audio résultant a perdu de l'intensité sonores, c'est-à-dire que l'amplitude de son gain en fréquence a considérablement diminuer. Cette méthode, théoriquement parlant, nous est convenable dans la mesure ou elle permet un filtrage du bruit blanc, mais elle rencontre ses limites avec les 2 autres types de bruit étudiés.

B. Résultats qualitatifs

En la tabla IV se muestra la calificación que se le dio a cada uno de los métodos para el audio original mezclado con 30% de ruido blanco.

Méthodes	Blanc (30%)
Sans méthode	
Wiener	
Butter T	
RNNbruit	

TABLE IV
QUALIFICATION DES MÉTHODES POUR 30% DU BRUIT BLANC

En la tabla V se muestra la calificación que se le dio a cada uno de los métodos para el audio original mezclado con 30% de ruido rosa.

Méthodes	Rose (30%)
Sans méthode	
Wiener	
Butter T	
RNNbruit	

TABLE V
QUALIFICATION DES MÉTHODES POUR 30% DU BRUIT ROSE

En la tabla VI se muestra la calificación que se le dio a cada uno de los métodos para el audio original mezclado con 30% de ruido café.

Méthodes	Marron (30%)
Sans méthode	
Wiener	
Butter T	
RNNbruit	

TABLE VI
QUALIFICATION DES MÉTHODES POUR 30% DU BRUIT MARRON