

Análisis de señales de audio utilizando la transformada de Gabor

Giraldo C, Daniel* Quintero M, Olga L.†

* Universidad EAFIT, Grupo de Modelado Matemático . Medellín
Colombia (e-mail: dgiral32@eafit.edu.co).

† Universidad EAFIT, Departamento de Ciencias Básicas. Medellín
Colombia (e-mail: oquinte1@eafit.edu.co).

Resumen: La transformada de Gabor es un caso especial de la transformada en tiempo corto de Fourier (STFT por sus siglas en inglés) la cual permite estimar las frecuencias de una señal en un determinado intervalo de tiempo. Para poder acotar la señal en estos intervalos, se utilizan funciones conocidas como funciones ventanas las cuales están determinadas en el dominio temporal y posibilitan el cálculo del espectrograma el cual es la gráfica de la densidad de energía de una señal definida en un plano tiempo-frecuencia. Cuando la ventana usada es una función Gaussiana, esta se conoce como la transformada de Gabor. El problema con este espectrograma es que dependiendo de la ventana escogida, los resultados podrían no ser claros y las conclusiones tomadas dependiendo de esta gráfica podrían ser erróneas. El interés de este trabajo es comparar algunas ventanas de tiempo con la intención de determinar las características que hacen buena o no a las ventanas. Además, se aplicará este espectrograma a una señal de audio con el objetivo de analizar las características que se pueden extraer de esta.

Palabras Clave: Espectrograma, ventanas de tiempo, frecuencias, atenuación, lóbulos.

1. INTRODUCCIÓN

Las señales son uno de los fenómenos que más se presentan dentro del mundo cotidiano, especialmente en el campo de la física, las finanzas, los procesos químicos, etc. Estas señales muestran una gran cantidad de características las cuales pueden ser analizadas e interpretadas utilizando diferentes tipos de herramientas matemáticas. En este caso, el trabajo se centrará en señales temporales (especialmente las señales de audio), ya que sobre estas se han construido diferentes métodos analíticos que permiten extraer este tipo de características. Una de las herramientas más comunes y usadas es la transformada de Fourier (TF) la cual pasa una señal del dominio temporal y la lleva al dominio frecuencial, lo que permite analizar las diferentes frecuencias presentes dentro de la señal. La transformada de Fourier se define como:

$$\hat{f}(\omega) = \int_{-\infty}^{\infty} f(t)e^{-i2\pi\omega t} dt \quad (1)$$

donde:

$f(t)$: Señal en el dominio temporal

$\hat{f}(\omega)$: Señal en el dominio frecuencial

Pero la gran debilidad de dicha transformación es que no puede hacer un análisis localmente, es decir, no se puede hacer un análisis de las frecuencias en algún intervalo de tiempo Arango (2009). Esto se puede ver en (1) ya que el resultado de la TF depende únicamente de la frecuencia ω . Debido a esto, se desarrolló un método el

cual permite realizar un análisis tanto en tiempo como en frecuencia. Este se basa en multiplicar a la señal original por otra llamada función ventada la cual acota la señal en cierto intervalo de tiempo y esto permite hacer la relación tiempo-frecuencia (esta multiplicación se hace dentro de la TF). Esta transformación se llama *transformada de Fourier en tiempo corto* (STFT) por sus siglas en inglés. Existen muchas funciones ventanas que permiten realizar este análisis:

- Rectangular
- Hann
- Hamming
- Blackman
- Gaussiana
- Triangular
- Bartlett
- Kaiser

... entre otras. La STFT está definida como:

$$G(\omega, \tau) = \int_{-\infty}^{\infty} f(t)g(t - \tau)e^{-i2\pi\omega t} dt \quad (2)$$

donde $g(t)$ es la función ventana. Cuando la función ventana $g(t)$ es Gaussiana, esta transformada se llama *transformada de Gabor* y además, esta transformada permite analizar la densidad espectral de energía de la señal la cual se define como $S(\omega, \tau) = G(\omega, \tau)^2$. La gráfica en el plano tiempo-frecuencia de la función $S(\omega, \tau)$ se conoce como espectrograma, el cual ha sido una herramienta muy utilizada por los programas editores de sonido ya que permite visualizar los cambios frecuenciales dentro de las señales de audio en tiempo real Lukin and Todd (2006).

Pero el método de *ventaneo* de una señal está fuertemente ligado con el principio de incertidumbre de Heisenberg

y en este caso se refiere a que aumentar o disminuir la longitud de la ventana (disminuir o aumentar la resolución en tiempo respectivamente) implica que la resolución en frecuencia aumentará o disminuirá en ese sentido, es decir, entre mas se quiera mejorar la calidad en tiempo, se perderá la calidad en frecuencia y viceversa como se describe en Arango (2009). Esta relación tiempo-frecuencia se determina como

$$\sigma_t \sigma_\omega \geq \frac{1}{2} \quad (3)$$

donde:

- σ_t : Dispersión de la ventana en el tiempo.
- σ_ω : Dispersión de la ventana en el frecuencia.

Esto indica que hay una relación inversamente proporcional de la dispersión en tiempo con la dispersión en frecuencia y que como mínimo, esta relación sera igual a 0.5. Cuando la ventana de tiempo es Gaussiana, el valor de (3) es el menor posible, es decir $\sigma_t \sigma_\omega = 1/2$.

Para poder trabajar con señales de audio, es indispensable discretizar esta transformada ya que este tipo de señales se presentan en forma digital, es decir, en tiempo discreto y es aquí donde varios autores han diseñado diferentes algoritmos para estimar eficientemente la STFT, como en Peter (2009) donde desarrolla un algoritmo para calcular esta transformación con la restricción que la ventana de tiempo tenga la misma longitud que la señal. En Bastiaans and Geilen (1996) se utiliza la transformación de Zak para representar la transformada de Gabor y la expansión de Gabor conjuntamente con el objetivo de generar una estructura de ambas herramientas para que su calculo algorítmico sea de menor costo computacional. En conclusión, la idea principal de este trabajo es comparar diferentes ventanas de tiempo con la ventana Gaussiana usando la TF y la STFT y determinar en que casos es mejor el uso de la una que la otra.

El artículo se organiza de la siguiente manera:

- En la sección 2 se plantea el problema a trabajar.
- En la sección 3 se explica las funciones ventanas a comparar.
- En la sección 4 se muestran los resultados de las ventanas y los espectrogramas trabajados.
- En la sección 5 muestran las conclusiones.

2. PLANTEAMIENTO DEL PROBLEMA

Para poder analizar las frecuencias presentes dentro de una señal a lo largo del eje temporal o espacial, es necesario utilizar la STFT definiendo que tipo de ventana se usará. La escogencia de esta ventana determina el buen cálculo de la STFT ya que cada una de éstas exhibe diferentes características que pueden reducir el error de estimación del espectrograma.

Como se vio en (3), existe una relación inversamente proporcional entre la dispersión en tiempo σ_t y la dispersión en frecuencia σ_ω . Estos dos valores están fuertemente ligados al tamaño de la ventana.

Pero también existen otros parámetros importantes dentro del espectro de frecuencias de cada ventana que determinan que tan exacto es la medida del espectrograma,

estos se conocen como *lóbulo principal* y *lóbulos laterales*. El primer parámetro se determina por el ancho de banda del lóbulo principal de la transformada de Fourier de la ventana y mide la capacidad de dicha ventana para detectar cambios de frecuencia de la señal a través del tiempo y determina la buena resolución en frecuencia siempre y cuando este ancho de banda $\Delta\omega$ sea estrecho o pequeño Smith et al. (2006) y Flórez et al. (2009). Pero al ser este valor pequeño, se disminuye la exactitud en el cálculo de la amplitud (el valor de (2)).

La medida de los lóbulos laterales indica que tantos *dB* decaen los primeros lóbulos laterales que se encuentran a lado y lado del lóbulo principal. Si los lóbulos principales caen o tienen valores muy bajos, estos atenúan la señal, es decir, reducen la medida de los picos dentro del espectrograma evitando ver frecuencias con altos valores debido a cambios abruptos de frecuencia, pero cuando estos lóbulos tienen valores muy bajos, el ancho de banda del lóbulo principal aumenta. Smith et al. (2006); National-Instruments (2012); Flórez et al. (2009).

3. MARCO TEÓRICO

3.1 Transformada de Gabor

Para el análisis y comparación entre ventanas se utilizarán la ventana Hann, Kaiser y la Gaussiana ya que la primera es una de las ventanas mas mencionadas dentro de la literatura, la segunda ventana al igual que la gaussiana son ventanas paramétricas que permiten mayor control sobre los parámetros a analizar dentro del espectro de frecuencias Smith et al. (2006).

Estas tres ventanas se definen como:

$$H(n) = \frac{1}{2} \left(1 - \cos \left(2\pi \frac{n}{N-1} \right) \right) \quad (4)$$

$$K(n) = \begin{cases} \frac{I_0 \left(\pi\beta \sqrt{1 - \left(\frac{2n}{N-1} - 1 \right)^2} \right)}{I_0\pi\beta} & 0 \leq n \leq N-1 \\ 0 & \text{en otro caso} \end{cases} \quad (5)$$

$$G(n) = e^{-\frac{1}{2} \left(\alpha \frac{n}{N/2} \right)^2} \quad (6)$$

En (4) y (5) n se encuentra en el intervalo $[0, N]$ y en (6) n se encuentra en $[-\frac{N}{2}, \frac{N}{2}]$ donde N es la longitud de la ventana en las tres funciones.

Además, β definido en la ventana Kaiser es un parámetro que se encuentra en función de la atenuación deseada de los lóbulos laterales y I_0 es la función de Bessel modificada de orden 0 del primer tipo.

El parámetro α definido en la ventana Gaussiana, es un parámetro inversamente proporcional a la varianza de esta y se define como $\alpha = \frac{N-1}{\sigma}$ donde σ es la desviación estándar de la función gaussiana.

Para la etapa de resultados, se harán diferentes simulaciones variando la longitud de la ventana para observar que características de ellas cambian y así determinar que tan bueno o que tan malo es tener ventanas de tamaño grande o pequeño. Además, teniendo en cuenta que tanto la ventana Kaiser y la Gaussiana están en función de un parámetro independiente de su longitud, este se hará variar también para analizar las características que estos exhiben.

4. RESULTADOS

Todas la ventanas tendrán una frecuencia de muestreo de $f_s = 8000Hz$ y además, los algoritmos fueron diseñados en MATLAB® y los cálculos del ancho de banda del lóbulo principal y el decaimiento de los lóbulos laterales fue hecho con el toolbox *wvtool*.

4.1 Ventana Hann

En la Fig. 1 se observa la forma de la ventana Hann. En las esquinas de la función se puede ver que está suaviza por completo el tramo de la señal a la cual se le aplica la TF.

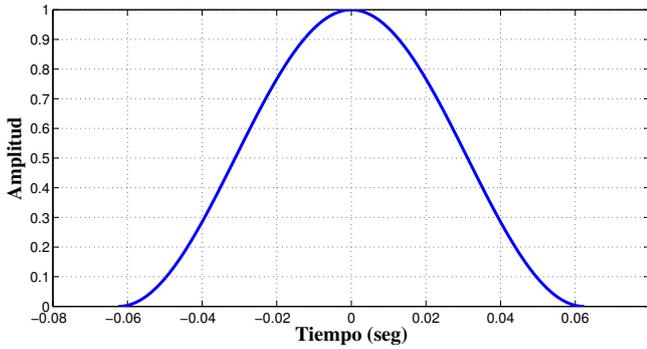


Fig. 1. Ventana Hann

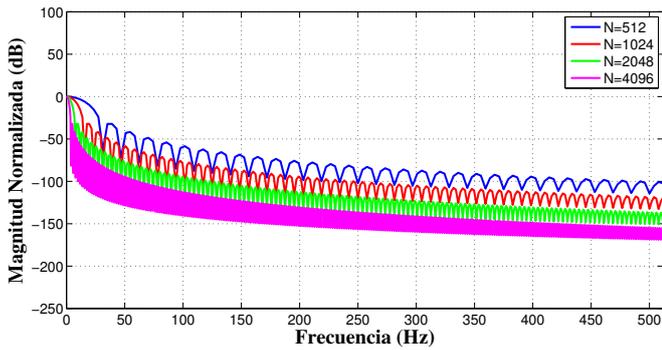


Fig. 2. TF de la ventana Hann con diferentes N .

En la Fig. 2 se observa que a medida que aumenta el valor de N , el ancho de banda $\Delta\omega$ del lóbulo principal disminuye pero la altura de los lóbulos laterales se mantiene constante. Esta información se puede ver en la Tabla 1.

Esto implica que al aumentar la longitud de la ventana, se aumenta la resolución en frecuencia y se puede detectar mejor los cambios abruptos de frecuencia pero se disminuye la resolución en tiempo, es decir, se hace muy difícil determinar en que tiempo aproximadamente aparece cierta frecuencia.

Tabla 1. Características de la ventana Hann.

Valor de N	$\Delta\omega$ (Lóbulo principal)	Altura del lóbulo lateral
512	22.461Hz	-31.5dB
1024	10.742Hz	-31.5dB
2048	5.371Hz	-31.5dB
4096	2.686Hz	-31.5dB

4.2 Ventana Kaiser

En la Fig. 3 se observan las variaciones en la ventana Kaiser en función de su parámetro β . El aumento de esta parámetro, concentra la energía de la ventana en su centro.

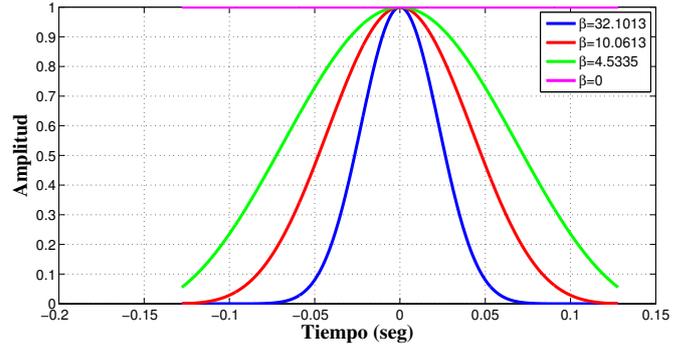


Fig. 3. Ventana Kaiser con diferentes valores de β .

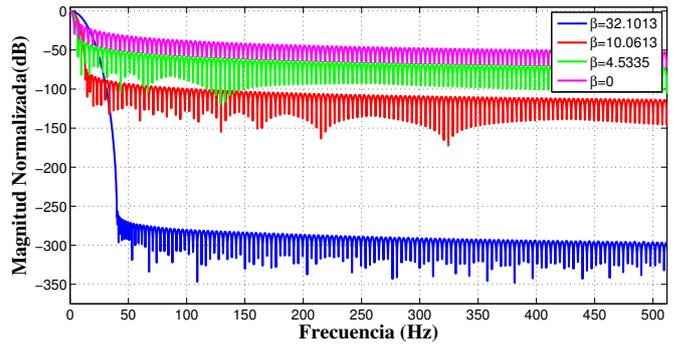


Fig. 4. TF de la ventana Kaiser variando β .

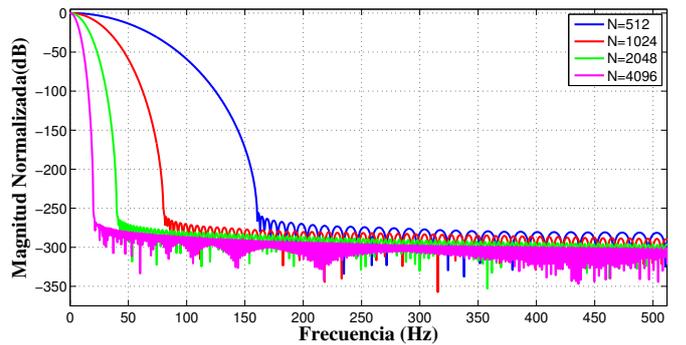


Fig. 5. TF de la ventana Kaiser variando N .

Tabla 2. Características de la ventana Kaiser variando β .

Valor de β	$\Delta\omega$ (Lóbulo principal)	Altura del lóbulo lateral
32.1013	11.719Hz	-260.7dB
10.0613	6.836Hz	-74.6dB
4.5335	4.883Hz	-33.6dB
0	3.418Hz	-13.3dB

Ya en la transformada de Fourier de estas funciones que se ven en la Fig. 4 se puede concluir que el ancho de banda de los lóbulos principales es directamente proporcional el valor de β y la altura de los lóbulos laterales disminuye a medida que crece dicho parámetro. Los resultados se pueden visualizar en la Tabla 2 y son ventanas de longitud $N = 2048$.

Ya al variar el tamaño de la ventana, ocurre el mismo efecto que en la ventana Hann, al aumentar el valor de N , disminuye el ancho de banda del lóbulo principal y la atenuación se mantiene relativamente constante como se ve en la Fig. 5 donde se tomó un $\beta = 32.1013$. Esta información se muestra en la Tabla 3.

Tabla 3. Características de la ventana Kaiser variando N .

Valor de N	$\Delta\omega$ (Lóbulo principal)	Altura del lóbulo lateral
512	46.875Hz	-256dB
1024	23.438Hz	-260.9dB
2048	11.719Hz	-260.7dB
4096	5.859Hz	-260.6dB

4.3 Ventana Gaussiana

Como se mencionó anteriormente, esta ventana depende de un parámetro α el cual es inversamente proporcional a la varianza y en la Fig. 6 se puede observar como cambia esta al variar dicho parámetro.

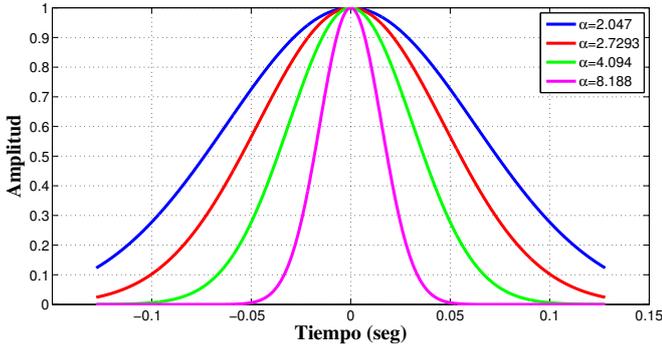


Fig. 6. Ventana Gaussiana con diferentes valores de α .

Luego, la Fig. 7 muestra la transformada de Fourier de dichas ventanas variando α y se puede notar como decaen los lóbulos laterales y por ende aumenta la atenuación pero al mismo tiempo se amplía el ancho de banda del lóbulo principal. En la Tabla 4 se muestra la relación anteriormente dicha.

Tabla 4. Características de la ventana Gaussiana variando α .

Valor de α	$\Delta\omega$ (Lóbulo principal)	Altura del lóbulo lateral
2.047	4.395Hz	-32.6dB
2.7293	5.37Hz	-48.5dB
4.094	8.301Hz	-90.6dB
8.188	16.602Hz	-313.7dB

Finalmente, la Fig. 8 muestra la misma relación que se vio en las ventanas Hann y Kaiser, es decir, el aumento del tamaño de la ventana repercute en la disminución del tamaño del lóbulo principal y un valor constante en la altura de los lóbulos laterales. Esta información es resumida en la Tabla 5.

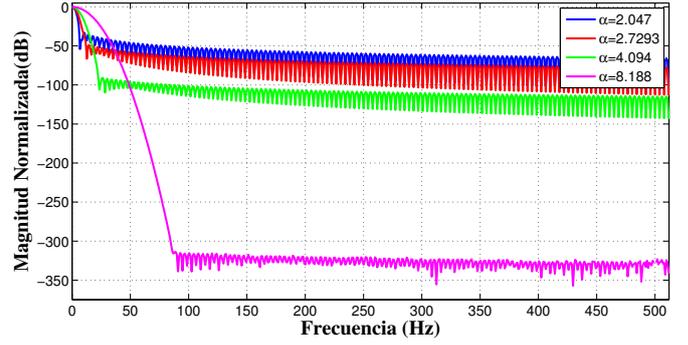


Fig. 7. TF de la ventana Gaussiana variando α .

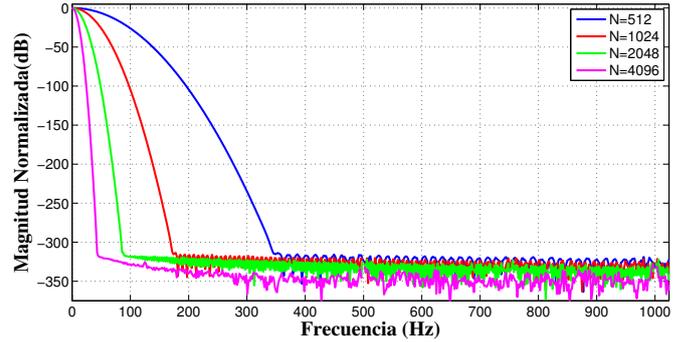


Fig. 8. TF de la ventana Gauss variando N .

De lo visto anteriormente, se puede concluir que el aumentar el tamaño de las ventanas solo permite disminuir el ancho de banda del lóbulo principal sin variar la atenuación posible de la ventana lo que implica mejoras en cuanto a resolución en frecuencia pero pérdida en la resolución en tiempo mientras que las ventanas paramétricas como son el caso de la Kaiser y la Gaussiana tienen más grados de libertad y permiten controlar de mejor manera el lóbulo principal y los laterales. Es por esto que solo se utilizarán estas dos ventanas en la próxima sección.

Tabla 5. Características de la ventana Gaussiana variando N .

Valor de N	$\Delta\omega$ (Lóbulo principal)	Altura del lóbulo lateral
512	67.383Hz	-313.6dB
1024	33.203Hz	-313.7dB
2048	16.602Hz	-313.7dB
4096	8.301Hz	-313.1dB

4.4 Espectrograma

Para evaluar las características de las ventanas de tiempo, se utilizarán dos señales, una es una señal sintética a conveniencia con el objetivo de poder ver todas las características en detalle. La segunda señal es una señal de audio la cual tiene un *beep* de 440Hz en la mitad de ella.

Señal sintética: La señal $f(t)$ se creó de la siguiente manera:

Sea $f_{aux}(t) = \sin(1800\pi t) + \sin(100\pi t^2 + 3000\pi t)$, entonces:

$$f(t) = \begin{cases} f_{aux}(t), & 0 \leq t \leq 7.5 \\ f_{aux}(t) + \sin\left(\frac{32}{3}\pi t^3\right) & 7.5 < t \leq 15 \end{cases} \quad (7)$$

Además, la frecuencia de muestro $f_s = 8000Hz$. Para simular la señal, se toma el periodo de muestro $\Delta t = \frac{1}{f_s}$ y el tamaño de la señal es $N = 120000$.

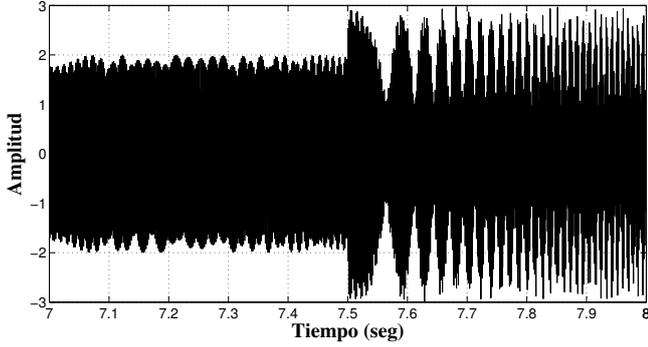


Fig. 9. Señal sintética en el tiempo.

La señal (7), la cual se puede ver en la Fig. 9, consta de un *beep* a $900Hz$, un *chirp* lineal con una frecuencia instantánea de $100Hz$ trasladado $1500Hz$ y un *chirp* cuadrático con una frecuencia instantánea de $16t^2$ que aparece a la mitad de la señal.

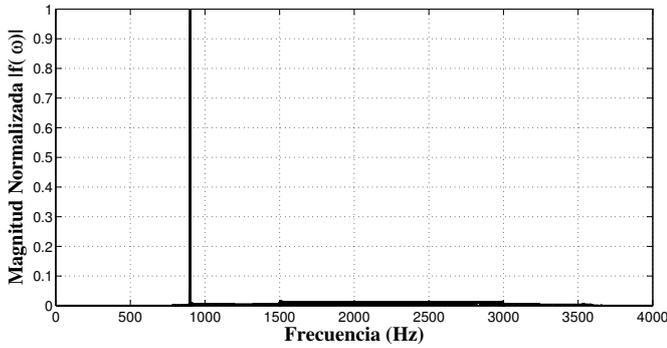


Fig. 10. Transformada de Fourier de la señal sintética.

Al analizar la Fig. 10, se observa un pico en $\omega = 900$ el cual hace referencia al *bib* introducido de $900Hz$ y además, hay unos picos de menor tamaño en un rango de frecuencias de $1500Hz - 3000Hz$ pero no se puede concluir nada más de ellos.

Para poder analizar la señal más a fondo y ver de que está compuesta, se calculará el espectrograma de $f(t)$ usando cuatro configuraciones de ventanas.

Primero que todo, el dominio del espectrograma se encuentra determinado por el plano tiempo-frecuencia $[t_0, t_f] \times [0, f_s/2]$, donde t_0 y t_f son el tiempo inicial y final de la señal respectivamente y f_s es la frecuencia de muestro de dicha señal. El hecho de que el espectrograma y la transformada de Fourier tengan un ancho de banda no superior a la mitad de la frecuencia de muestro, se fundamenta en el teorema del muestro de Nyquist Mallat (1999). Los colores determinan la potencia de las frecuencias presentes de la señal, entre más alto el valor del espectrograma en un punto (ω_k, τ_n) , implica que en una vecindad de τ_n , existen amplitudes altas a una frecuencia ω_k .

Tanto en el espectrograma de la Fig. 11 y de la Fig. 12 se puede notar el *bib*, el *chirp* lineal y el *chirp* cuadrático que componen la señal sintética, pero en la primera se observa

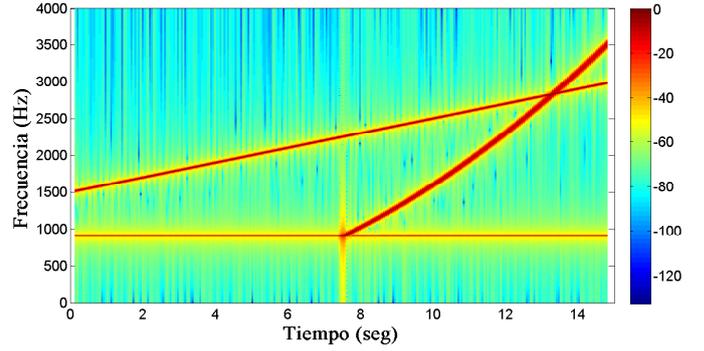


Fig. 11. Espectrograma usando la ventana Kaiser con $N = 2048$ y $\beta = 2.7829$.

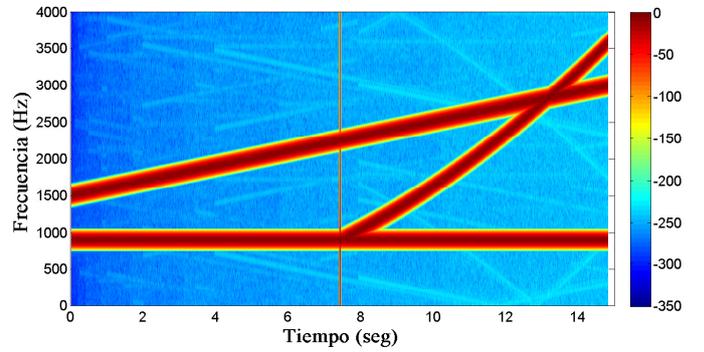


Fig. 12. Espectrograma usando la ventana Kaiser con $N = 512$ y $\beta = 32.1013$.

que hay buena resolución en frecuencia ya que el tamaño de la ventana es grande y además el lóbulo principal de esta ventana es inferior a $5Hz$; esto también implica que los cambios abruptos de frecuencia serán detectados. Esto se puede ver entre la unión del *beep* y el *chirp* cuadrático en el instante $t = 7.5seg$ ya que no hay un derramamiento de energía grande alrededor de este punto, pero los puntos cercanos a las líneas que delimitan las frecuencias presentes en la señal tienen magnitudes cercanas a las de estas frecuencias debido a que los lóbulos laterales son muy altos y la ventana no es capaz de atenuar estos puntos circundantes. Además, el tamaño de la ventana hace que la resolución en tiempo no sea tan buena.

En la Fig. 12 se puede ver la gran diferencia de magnitudes entre las frecuencias presentes en la señal y las que no, debido a que la atenuación de la ventana usada para este espectrograma es de $-256dB$. También la resolución en tiempo mejoró considerablemente debido a que se disminuyó el tamaño de la ventana, pero por este mismo hecho y recordando el principio de incertidumbre descrito por (3), la resolución en frecuencia empeora y se hace más difícil saber que frecuencias hay dentro de la señal. Además, según la Tabla 3, el lóbulo principal tiene un ancho de banda grande, lo que implica que cambios abruptos en la señal podrán acumular mucha energía dentro del espectro de frecuencias y esto se puede ver en la intersección entre el *beep* y el *chirp* que se describió anteriormente, ya que en este punto se está marcando un línea que indicaría er-

róneamente que hay una componente de la señal en $t = 7.5$ la cual se encuentra presente en todas las frecuencias.

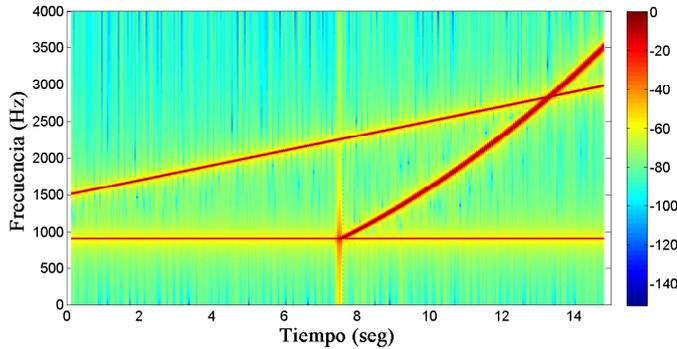


Fig. 13. Espectrograma usando la ventana Gaussiana con $N = 2048$ y $\alpha = 2.7293$.

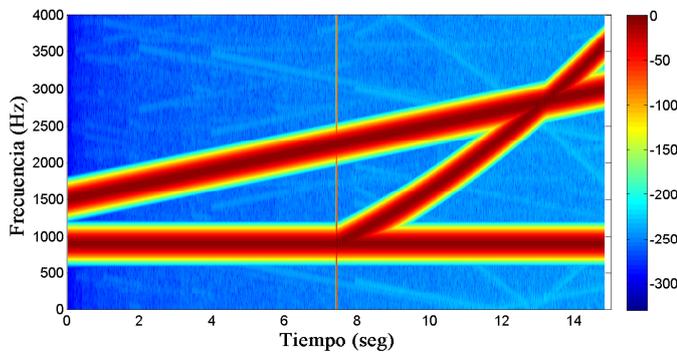


Fig. 14. Espectrograma usando la ventana Gaussiana con $N = 512$ y $\alpha = 8.188$.

En la Fig. 13 se muestra el espectrograma de (7) usando una ventana Gaussiana.

Se puede observar en el espectrograma de la Fig. 14, hay una dispersión de energía en las frecuencias debido a que el lóbulo principal de la ventana Gaussiana usada es mucho mas ancho que el del espectrograma de la Fig. 12 usando la ventana Kaiser.

Señal de audio Esta señal de audio es una conversación entre dos personas la cual tiene un *beep* de $440Hz$. La Fig. 15 corresponde esta señal y la Fig. 16 muestra su transformada de Fourier. En esta última se puede observar un pico alto cercano a los $500Hz$ el cual corresponde al beeb que se describió anteriormente aunque esta deducción solo se hace por el conocimiento de la existencia de dicho *beep*.

En la Fig. 17 y en la Fig. 18 se puede observar el espectrograma de la señal de audio usando la ventana Kaiser con diferentes valores de β mientras que la Fig. 19 y la Fig. 20 muestran el espectrograma usando ventanas Gaussianas con dos valores diferentes de α .

Inicialmente, las frecuencias que aparecen cerca a los dos, cuatro, ocho, diez, doce y catorce segundos dan cuenta de la conversación de las personas que hay en la señal y las líneas que se derivan de ellas y cambian en la misma

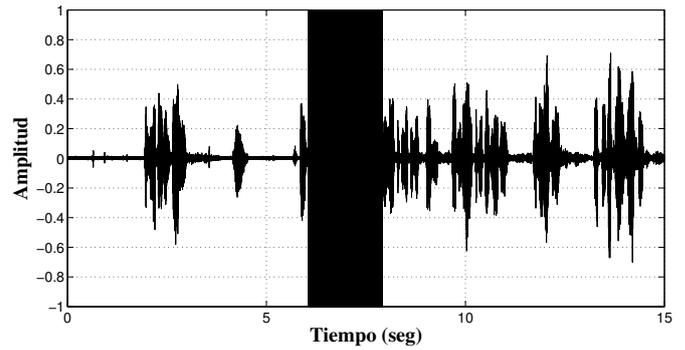


Fig. 15. Señal de audio en el tiempo.

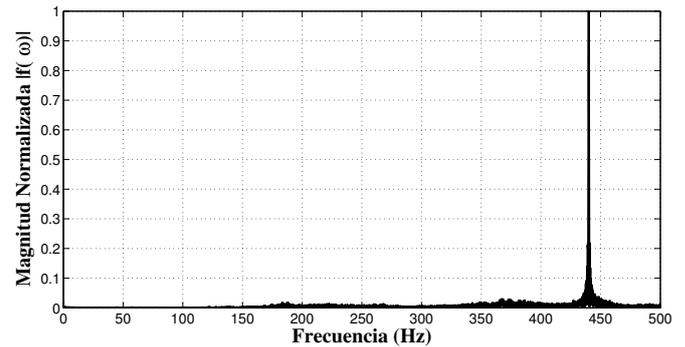


Fig. 16. TF de la señal de audio en el tiempo.

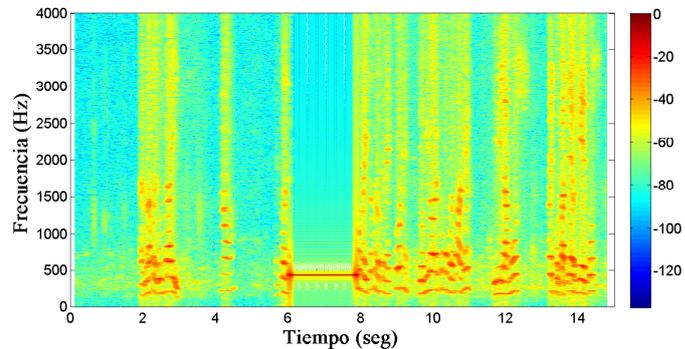


Fig. 17. Espectrograma de la señal de audio. Ventana Kaiser con $N = 2048$ y $\beta = 4.5335$.

dirección del eje de frecuencias son las formantes de la voz las cuales son las frecuencias a las cuales vibra el tracto vocal. Ya la línea intermedia aproximadamente entre los seis y ocho segundos es el *beep*.

Además, se puede observar ya sea en al Fig. 18 o en la Fig. 20 que en las frecuencias superiores a los $440Hz$ entre el intervalo de tiempo en el cual se encuentra el *beep*, el espectrograma mide la menor potencia posible en todo el dominio de la señal y como en ese instante la única frecuencia es la del beeb, el color azul que rodea esta frecuencia indica que no hay otro tipo de frecuencias dentro de este intervalo y es por esto que los valores de color amarillo limitados por los instantes de tiempo donde hay conversación no son valores nulos sino frecuencias de baja amplitud y como en esos intantes de tiempo no hay habla de las personas, esto indica la presencia un ruido de baja amplitud en estos sectores.

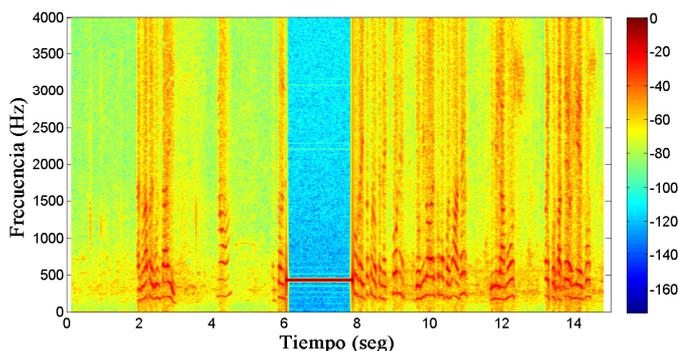


Fig. 18. Espectrograma de la señal de audio. Ventana Kaiser con $N = 2048$ y $\beta = 32.1013$.

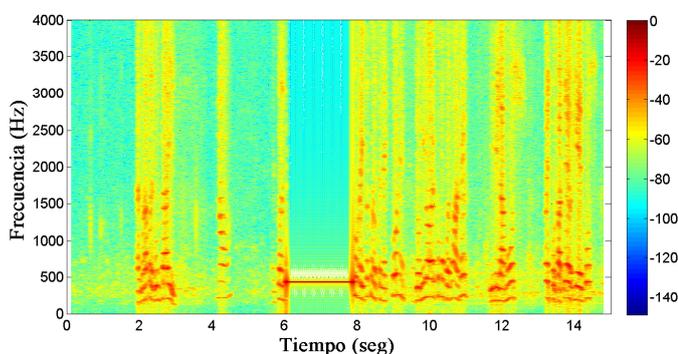


Fig. 19. Espectrograma de la señal de audio. Ventana Gaussiana con $N = 2048$ y $\alpha = 2.7293$.

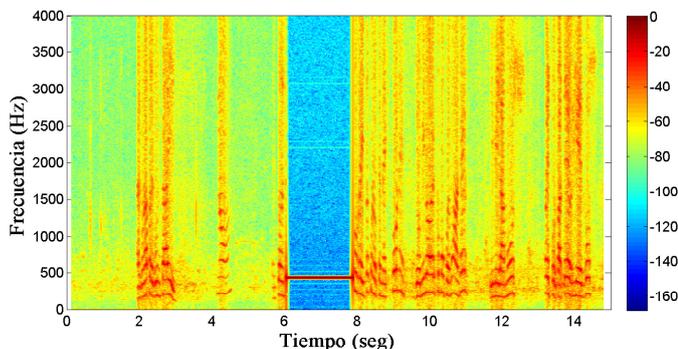


Fig. 20. Espectrograma de la señal de audio. Ventana Gaussiana con $N = 2048$ y $\alpha = 5.3222$.

Estos mismos ruidos llegan al límite del espectrograma ($4000Hz$) y en casos de presencia de ruido blanco, estos están presentes en todas las frecuencias y por esto que se obser dentro de la señal de audio un problema de aliasing que es cuando la frecuencia de muestreo es menor al doble de la frecuencia máxima dentro de una señal y para este caso, la frecuencia de muestreo es $f_s = 8000Hz$, lo cual solo permitiría ver frecuencias menores o iguales a $4000Hz$ y el ruido supera esta cota. Esta característica es muy a la hora de tratar señales.

Asimismo, se observa que en las cuatro figuras se puede concluir en si que no hay un cambio significativo com-

parando el espectrograma usando la ventana Kaiser y la ventana Gaussiana con $\beta = 4.5335$ y $\alpha = 2.7293$ respectivamente y tampoco entre estas mismas con $\beta = 32.1013$ y $\alpha = 5.3222$ respectivamente.

Esto implica que el hacer un análisis inicial de la señal usando el espectrograma, este puede dar los mismos resultados usando dos ventanas de tiempo diferentes y con esto último, se ve que a pesar de que las ventanas Gaussianas tienen la mejor relación tiempo-frecuencia descrita por la ecuación (3), existen otros parámetros determinados por la ventana de tiempo que permiten mejor la calidad y el entendimiento de este tipo de herramientas.

5. CONCLUSIONES

Se utilizaron diferentes ventanas de tiempo analizando sus características en frecuencia para observar como estas repercuten en el análisis del espectrograma.

Se determinó que la escogencia de una ventana para analizar la densidad espectral de energía $S(\omega, \tau) = G(\omega, \tau)^2$ de una señal no solo depende de las características en frecuencia que dicha ventana tiene, sino también de la misma señal a analizar y es por esto que no se puede decir que existe una ventana mejor que la otra en términos generales, sino que existen ventanas mejor que otras para cierto tipos de señales.

Las ventanas paramétricas permiten controlar mejor las variables que determinan el buen resultado del cálculo del espectrograma ya que este tipo de ventanas posibilitan controlar tanto el ancho de banda del lóbulo principal y el decaimiento de los lóbulos laterales. Esto deja analizar las señales de manera en que se puedan detectar las frecuencias que componen dichas señal extrayendo la información que se necesita.

El hecho de utilizar una ventana con alta atenuación, no implica que se mejore las medición de amplitudes o densidad de energía de una señal a través del espectrograma ya que se puede observar en los últimos ejemplos, dos ventanas con una diferencia de atenuación grande, permiten ver un espectrograma casi idéntico. Esto se da debido a las amplitudes de la señal.

REFERENCIAS

- Arango, H. (2009). *Análisis de Señales con la transformadas de Fourier, Gabor y Ondita*. Serie de textos académicos Instituto Tecnológico Metropolitano, Medellín, primera edition.
- Bastiaans, M.J. and Geilen, M.C. (1996). On the discrete gabor transform and the discrete zak transform. *Signal Processing*, 151–166.
- Flórez, E., Cardona, S., and Jordi, L. (2009). Selección de la ventana temporal en la transformada de fourier en tiempos cortos utilizada en el análisis de señales de vibración para determinar planos en las ruedas de un tren. *Rev. Fac. Ing. Univ. Antioquia N*, 145–158.
- Lukin, A. and Todd, J. (2006). Adaptive time-frequency resolution for analysis and processing of audio. *Audio Engineering Society*, 20–23.
- Mallat, S. (1999). *A Wavelet tour of signal processing*. Academic Press, San Diego, second edition.

National-Instruments (2012). *Characteristics of Different Smoothing Windows*.

Peter, L. (2009). An efficient algorithm for the discrete gabor transform using full length windows. *Sampta*.

Smith, O., Julius, and Serra, X. (2006). Parshl: An analysis/synthesis program for non-harmonic sounds based on a sinusoidal representation. Technical report, Stanford University.

Appendix A. CÓDIGO DE LA STFT

```
%% =====  
% Código: Transformada en tiempo corto de Fourier  
% Nombre: Daniel Giraldo Cuartas  
% Código: 201017500101  
% Ult. Modificación: Mayo 15 de 2014  
%% =====  
  
function [S U w P]=STFT(f_t,Fs,Window,Solap,NFFT)  
N=length(f_t);  
t = (0:N-1)/Fs;  
  
f_t=f_t(:);  
Window=Window(:);  
  
n=length(Window);  
  
%% Cálculo del número de ventanas sobre la señal  
Rate = n-Solap;  
M=floor((N-n)/Rate)+1;  
S=NaN(M,NFFT/2+1);  
  
meanU = mean(t(1:length(Window)));  
T_time = t(Rate);  
U=(0:M-1)*T_time + meanU;  
  
I=(1:n)-Rate;  
  
%% Vector de frecuencias en Hz  
w = Fs/2*linspace(0,1,NFFT/2+1);  
for u=1:M  
  
    %% Señal multiplicada por la ventana de tiempo  
    I=I+Rate;  
    G=f_t(I).*Window;  
  
    if n>NFFT  
        Gcir=buffer(G,NFFT);  
        G=sum(Gcir,2);  
    end  
  
    %% Uso de la FFT  
    fu_w = fft(G,NFFT);  
  
    S(u,:)=fu_w(1:NFFT/2+1);  
end  
%% Densidad Espectral  
K=1/(Fs*norm(Window)^2);  
P=(abs(S)).^2;  
P=[K*P(:,1) 2*K*P(:,2:end-1) K*P(:,end)];
```