

Bandas críticas y el cálculo de potencia espectral de barks en rango de frecuencia lineal y por tercios de octava

Critical bandwidths and the calculation of power bark spectrum within linear frequency range and one-third octave bands

Recibido: julio de 2012
Arbitrado: septiembre de 2012

Marcelo Herrera*, Alexander Ortega**, Carlos Vargas G.***

Resumen

En el presente artículo, calculamos la Potencia Espectral de Barks de una forma correcta y precisa. El presente trabajo calcula la Potencia Espectral de Barks con un error pequeño, el cual es producido cuando se calcula la Transformada Discreta de Fourier (DFT). Los investigadores tienden a olvidar las suposiciones y condiciones que son inherentes al cálculo. Una corta introducción sobre las Bandas Críticas y la Escala Bark son presentadas.

Palabras clave

Densidad de potencia espectral, bandas críticas, transformada discreta de Fourier.

Abstract

In the present paper we calculate so-called Bark Power Spectrum in a correct and precise way. Present works calculate the Bark Power Spectrum with a small error account which is produced because when calculating the Discrete Fourier Transform (DFT), researchers tend to forget the correct assumptions and conditions which are inherent to the calculation. A brief introduction about the Critical Bandwidths and the Bark Scale is also presented.

Keywords

Power bark spectrum, critical bandwidths, discrete Fourier transform.

I. Introducción

Científicos en áreas como fisiología, electrónica y psicoacústica investigan acerca de los mecanismos y principios que nos permiten tener la sensación auditiva. Las propiedades de transducción a nivel del oído interno, así como el acople de impedancias a nivel de los huesecillos que están adheridos al tímpano, y las reacciones químicas entre las concentraciones de los líquidos que tienen lugar en las células ciliadas internas y externas en la membrana basilar explican parcialmente los fenómenos que llevan a la sensación final que denominamos

* Ingeniero Electrónico, Magister en Radioelectrónica y Doctor en Acústica de la Universidad Técnica de Praga. Profesor Titular de la Facultad de Ingeniería de la Universidad de San Buenaventura, Bogotá. Líder del Semillero de Investigación de «Sistemas de compresión perceptual de audio». mherrera@usbog.edu.co

** Físico de la Universidad del Valle (Cali - Colombia), y Magister en Acústica y Vibraciones de la Universidad Austral de Chile (Valdivia - Chile). Profesor de la Facultad de Ingeniería de la Universidad de San Buenaventura, Bogotá. Integrante del Grupo De Investigación Acústica Aplicada, del programa de Ingeniería de Sonido. aortega@usbog.edu.co

*** Ingeniero Electrónico, Universidad Distrital Francisco José de Caldas. Especialista en Teleinformática Universidad Distrital Francisco José de Caldas. Docente Investigador Facultad de Ingeniería, Universidad Manuela Beltrán, sede Bogotá. E-mail: carlos.vargas@docentes.umb.edu.co

audición. El mapeo de frecuencia a espacio, explica la dependencia de la frecuencia de la señal con su respectivo lugar de excitación en un lugar particular en la cóclea, en la membrana basilar. Los componentes de baja frecuencia viajan a través de toda la membrana para localizar su «lugar de procesamiento», y los componentes de alta frecuencia serán procesados en el comienzo de la membrana. Los mecanismos de sincronización, así como la sincronización de tasas de descarga neuronales, también explican la variedad de mecanismos que explican el problema.

La herramienta más conocida en audio para el análisis tiempo-frecuencia es el Espectrograma. Sin embargo, desde los años 90, varios trabajos han sido realizados para desarrollar mejores técnicas para el análisis sonoro, con el fin de considerar propiedades psicoacústicas. De esta manera, en la actualidad varios algoritmos, diseñados para procesamiento de audio y de habla hacen referencia a transformaciones psicoacústicas. Una limitación fuerte del espectrograma es la descripción de valores de amplitud de la señal en decibeles, perdiendo información acerca de la impresión subjetiva de la señal. De esta manera, fueron al principio introducidas variables subjetivas como los sones y los fones. Expresiones para la descripción de sonoridad específica y sonoridad total fueron así implementadas. Estos descriptores de sonoridad pueden ser interpretados a partir del cálculo de densidad de potencia espectral de la señal de audio.

II. Marco teórico

Existen varias implementaciones robustas de la variable subjetiva sonoridad. Una de las descomposiciones más sofisticadas es el modelo de sonoridad de Zwicker usado en PEAQ.

La potencia espectral de cada campo es ponderada por la respuesta en frecuencia del oído externo y medio, derivado del modelo. Las energías de potencia espectral son agrupadas en bandas críticas, espaciadas 0.25 barks. Una compensación

es sumada a las energías de banda crítica para compensar el ruido interno generado en el oído.

Una función de dispersión triangular (en dB) es usada para implementar la función de dispersión en el dominio de la frecuencia.

El modelo de Moore y Glasberg es diferente al de Zwicker en que la escala de frecuencia auditorial usada es el ancho de banda equivalente rectangular (ERB).

2.1 Bandas críticas

Consideremos una respuesta de la periferia auditorial con respecto a la señal compuesta por dos tonos. Alineemos todas las neuronas para incrementar la CF (característica en frecuencia) de izquierda a derecha. El eje generado es el lugar *tonotópico*.

Debido a que las neuronas del campo auditorial tienen funciones de filtro asimétricas, los patrones de excitación no son simétricas con respecto a los lugares CF.

La idea del filtro de bandas críticas es que el análisis a través de la coordenada tonotópica causa que los dos componentes de la señal son procesados por canales independientes. Estos canales son llamados los *filtros de bandas-críticas*, y los anchos de los canales son los *anchos de bandas críticas*.

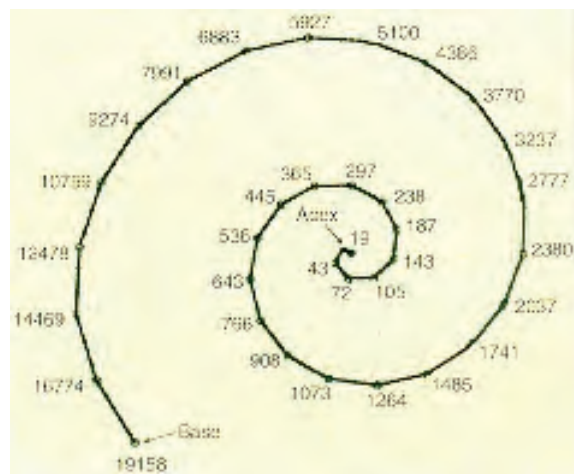


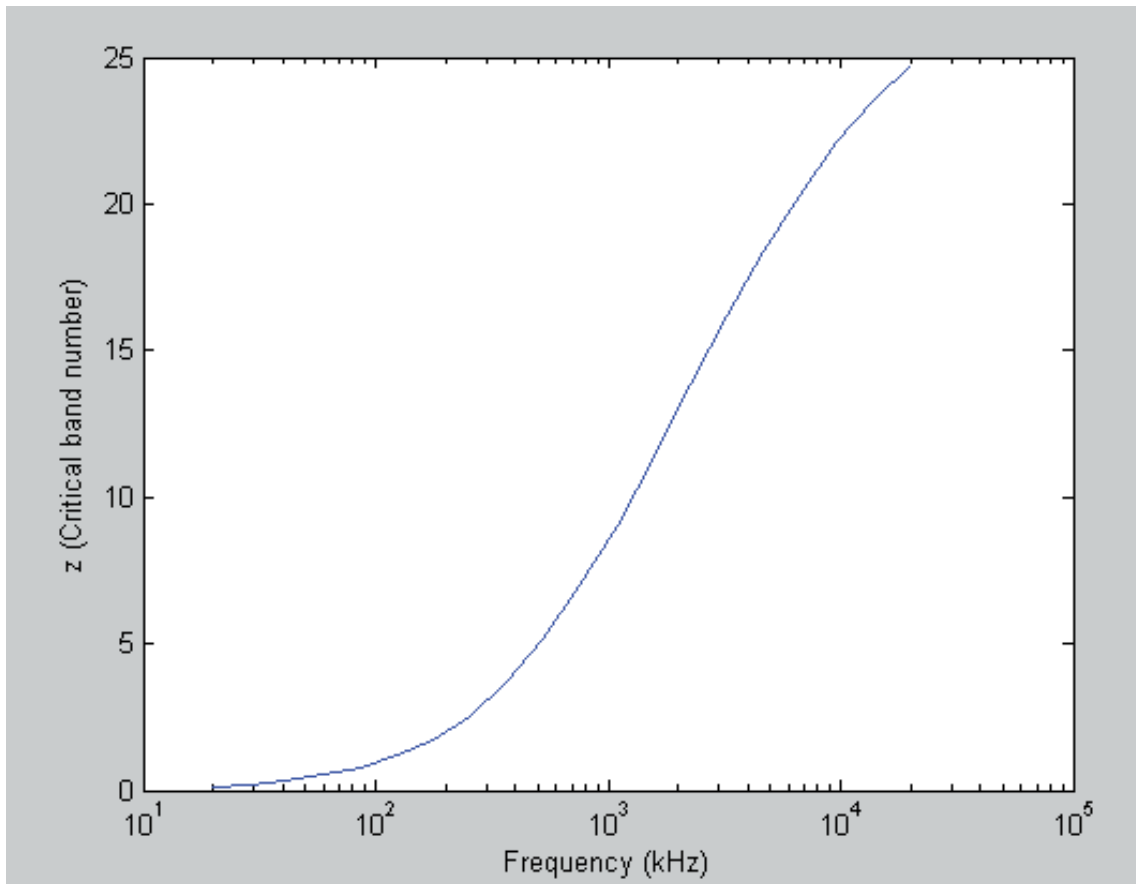
Figura 1. Distribución tonotópica dentro de la Membrana Basilar.

La siguiente gráfica es llamada las bandas críticas de «Munich» que fueron diseñadas y explicadas por Zwicker in 1961 [1]. Las gráficas se originan a partir de las siguientes expresiones aproximadas para el ERBs en unidades de Hz:

$$B_M \approx 25 \pm 75(1 \pm 1.4F^2)^{0.69} \quad (1),$$

y

$$B_c \approx 24.7(1 \pm 4.37F) \quad (2)$$



Gráfica 2. Cálculo de Barks.

Las bandas críticas tienen aproximadamente 1/3 de octava de ancho. Los bancos de filtro de un-tercio de octava son usados comúnmente en la industria de audio como una idealización conveniente de los filtros del auditorio humano. La gráfica de tercios de octava está dada por:

$$B_{1/3} \approx 2^{1/6} f \Psi 2^{n/6} f \approx 0.232 f \approx 232F \quad (3)$$

El número de banda-crítica es una escala absoluta para la frecuencia basado en el análisis frecuencial del oído. El número de bandas-críticas se representa por el símbolo z.

$$z(f) \approx 9.26 \ln(0.00437f \pm 1) \approx 21.3 \log(0.00437f \pm 1) \quad (4)$$

Esta expresión proporciona una escala de frecuencia psicoacústica basada en resolución de frecuencia, así como es determinada en experimentos de ruido gradual. Posiblemente esto se correlaciona con el mapeo de frecuencias en la cóclea.

La escala de Barks es una escala absoluta de frecuencias basada en los anchos de banda de Munich. La unidad fue nombrada en honor a Heinrich Barkhausen.

La tabla n.º 1 presenta y explica la Escala Bark.

Banda crítica (Bark)	Frec. central (Hertz)	Ancho de banda (Hertz)	Frec. min. (Hertz)	Frec. max. (Hertz)
1	50	-	-	100
2	150	100	100	200
3	250	100	200	300
4	350	100	300	400
5	450	110	400	510
6	570	120	510	630
7	700	140	630	770
8	840	150	770	920
9	1000	160	920	1080
10	1170	190	1080	1270
11	1370	210	1270	1480
12	1600	240	1480	1720
13	1850	280	1720	2000
14	2150	320	2000	2320
15	2500	380	2320	2700
16	2900	450	2700	3150
17	3400	550	3150	3700
18	4000	700	3700	4400
19	4800	900	4400	5300
20	5800	1100	5300	6400
21	7000	1300	6400	7700
22	8500	1800	7700	9500
23	10500	2500	9500	12000
24	13500	3500	12000	15500
25	18775	6550	15500	22050

Tabla n.º 1. División de Barks en bandas de frecuencia.

2.2 Consideraciones acerca de la Transformada de Fourier cuando se calcula la Potencia Espectral de Barks

La Transformada de Fourier es una generalización de la serie compleja de Fourier, debido a que el límite tiende a infinito, $L \rightarrow \infty$.

Debido a que la Serie de Fourier es una expansión de una función periódica $f(x)$ en términos de una suma infinita de senos y cosenos. Para una función periódica en $[-L/2, L/2]$, la Serie de Fourier puede ser expresada como:

$$f(x) = \sum_{n=-\infty}^{\infty} A_n e^{j(2\pi n x / L)} \quad (5)$$

$$A_n = \frac{1}{L} \int_{-L/2}^{L/2} f(x) e^{-j(2\pi n x / L)} dx \quad (6)$$

Entonces, cuando realizamos la transformación de una variable discreta a una continua, obtenemos la Transformada de Fourier. Obtenemos, de esta manera, las expresiones de la Transformada de Fourier, es decir el aparato matemático usado para calcular la Densidad de Potencia Espectral,

$$f(x) \approx \int_{-\Delta/2}^{\Delta/2} M(k) e^{j2\pi kx} dk \quad (7)$$

$$F(k) \approx \int_{-\Delta/2}^{\Delta/2} M(x) e^{j2\pi kx} dx \quad (8)$$

Sin embargo, MATLAB, un ambiente computacional discreto, computa la Transformada de Fourier en señales discretas. Esto conlleva a la Transformada de Fourier Discreta, en donde, cada uno de los valores de amplitud singulares, pertenece a un Δf definido. Cuando calculamos la Potencia Espectral de Barks, es importante tener esto en cuenta, porque es posible que un valor de amplitud determinado de la Transformada de Fourier pertenezca a dos bandas críticas diferentes, o *barks*, como explica el gráfico.

El cálculo del Espectro de Potencia de Barks antes del cómputo del valor crítico de la banda crítica está dada por la expresión:

$$P_{beforeBark} \approx \dots \ni Px \ni Py_1 \quad (9)$$

$$P_{afterBark} \approx Py_2 \ni Pz \ni \dots \quad (10)$$

La expresión para el cálculo de la Potencia Espectral del Espectro de Fourier de la señal está dado por,

$$P \approx F(w)F^*(w) \quad (11)$$

donde $F^*(w)$ es el conjugado complejo de la Transformada de Fourier de la señal.

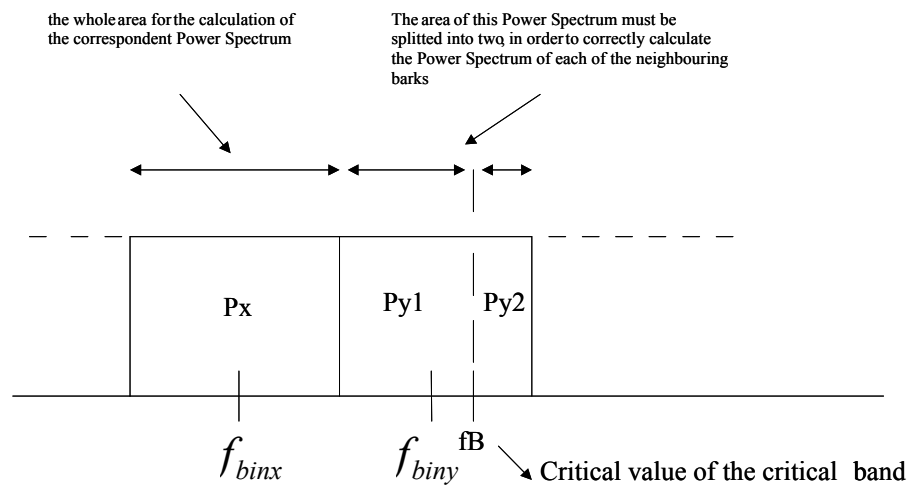


Figura 3. Cálculo de Densidad Espectral de Potencia con la corrección de bins espectrales.

III. Algoritmo para el cálculo de la potencia de densidad espectral

El algoritmo para el cálculo de la Potencia Espectral en bandas de octava, en tercios de octava o en Barks, es realizado en el ambiente de programación MATLAB.

Primero se realiza la lectura del archivo de audio. Posteriormente procedemos a calcular el espectro de la señal con el algoritmo FFT (Transformada Rápida de Fourier), y finalmente distribuimos la energía de la señal en cada una de las bandas de octava y tercios de octava.

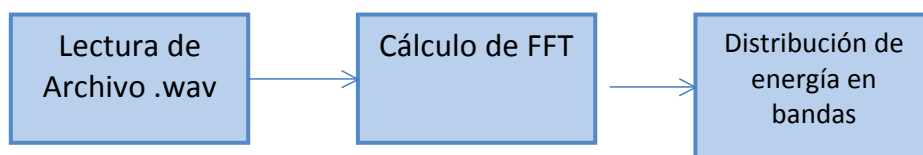


Figura 4. Diagrama de bloques del algoritmo para el cálculo de Potencia Espectral

3.1 Lectura de archivo .WAV

La aplicación del presente proyecto se realiza con señales de audio. Si cargamos una señal estereofónica puede realizarse, en cuyo caso se tomará un solo canal. En otras palabras, la aplicación realiza el cálculo de la energía espectral por bandas de señales monofónicas. El comando de lectura es el siguiente:

```
xx=wavread('C:\Users\musicamarcelo\Desktop\
castanets2.wav');
```

3.2 Cálculo de FFT

Se realiza el cálculo de la Transformada de Fourier con el algoritmo básico de la FFT, teniendo en cuenta que se calcula el valor absoluto de esta, se divide entre la longitud de la señal y posteriormente recordando que la energía de una señal es el cuadrado de los valores de amplitud de esta, tenemos que realizar el cálculo de la siguiente manera:

```
xfft=(abs((fft(x))/l)). ^ 2;
```

3.3 Distribución de energía de potencia espectral en cada una de las bandas de octava y tercios de octava

Se realiza la suma de energía dentro de cada una de las bandas de octava o de tercios de octava, para realizar la distribución de energía dentro de cada una de ellas.

```
for i=2:25
B_XX(i-1)=sum(xfft(bark(i-1,2)<=f & f<bark(i,2)));
end
```

IV. Conclusiones

El presente trabajo realiza el cálculo de la distribución de energía espectral dentro de cada una de las bandas de octava, de tercios de octava, y barks de una señal de audio monofónica.

El cálculo se realiza con la Transformada Rápida de Fourier, FFT. Posteriormente se suman las contribuciones de energía por bandas de octava, por tercios de octava y barks. Finalmente se realiza la corrección del algoritmo, para tener en cuenta el proceso de discretizado de la señal en MATLAB, en donde no se evalúan frecuencias instantáneas, sino rangos o anchos de banda muy angostos pero con ancho diferente de cero.

V. Referencias bibliográficas

- [1] Gelfand, S.A., Hearing: An Introduction to Psychological and Physiological Acoustics, Marcel Dekker, 1998.
- [2] Wonho, Y., Enhanced modified bark spectral distortion (EMBSD):An objective speech quality measure based on audible distortion and cognition model, Ph.D. thesis, Temple University, Ft. Washington, USA, 1999.
- [3] Appell, J., et al., 'Review of loudness models for normal and hearing-impaired listeners based on the model proposed by Zwicker', Audiologische Akustik, 40, No.(2), 2002.
- [4] Zwicker, E., Fastl, H., and Dallmayr, C., 'BASIC program for calculating the loudness of sounds from their 1/3 oct band spectra according to ISO 532 B', Acustica 55, 1984, pp. 63-67.
- [5] Quast, H., 'Absolute Perceived Loudness of Speech', Proceedings of the 7th Joint Symposium on Neural Computation, USC, 2000.
- [6] ITU Recommendation, P.861 Objective Measurement of Telephone Band (300-3400Hz) Speech Codecs (PSQM)
- [7] Thiede, T., Perceptual Audio Quality Assessment using a Non-Linear Filter Bank, Ph.D. thesis, Technische Universität Berlin, Berlin, Germany, 1999.
- [8] Moore, B., Glasberg, B., Baer, T., 'A model for the prediction of thresholds, loudness and partial loudness', J. Audio Eng. Soc.45, 1997, pp. 224-240.
- [9] <http://www.auditory.org/postings/2002/565.html>