



# UNIVERSIDAD DE GRANADA

SISTEMAS MULTIMEDIA

## **Trabajo voluntario: Sonido**

*Comparación de resolución, frecuencias de muestreo y bitrates de códecs*

*Javier Sáez de la Coba*

Curso 2018-2019

29 de abril de 2019

# Índice

<b>1. Introducción</b>	<b>2</b>
<b>2. Comparativa de frecuencias de muestreo y resolución de muestra</b>	<b>3</b>
<b>3. Comparativa de códecs</b>	<b>4</b>

# 1. Introducción

En este trabajo vamos a comprobar como afectan las distintas resoluciones y calidades al audio. Aunque originalmente se proponía establecer la calidad CD (44.1 KHz / 16 bits) como calidad máxima vamos a subir el listón y la calidad máxima a comparar va a ser 96 KHz / 24 bits por muestra, también conocida como calidad de estudio o audio de alta definición.

Las fuentes de sonido usadas son dos:

- Joseph Haydn: String Quartet In D, Op. 76, No. 5 - Finale - Presto grabado por el estudio 2L <sup>1</sup>
- Grabación propia de voz narrando la sinopsis del libro ' *El nombre del viento* ' de Patrick Rothfuss grabado con un micrófono AKG de condensador y una interfaz de sonido Scarlett 2i2 en alta resolución.

Se ha elegido la grabación del conjunto de cuerdas ya que representa todas las frecuencias que se pueden generar en una orquesta, teniendo el contrabajo como instrumento más grave (partiendo desde los 8.18 Hz) hasta el violín, capaz de alcanzar en sus armónicos superiores más de 15KHz. Esto acompañado con la acústica de la sala permite que se note la diferencia entre las distintas resoluciones del audio.

La voz humana sin embargo tiene un registro mucho más limitado y no cuenta con tantos armónicos, por lo que a priori podemos decir que no se detectaran diferencias cualitativas significativas entre las distintas calidades.

Todos los archivos de sonido, así como el material complementario a este trabajo se puede encontrar en <https://jscoba.com/sm-sonido>

Para las pruebas subjetivas se han usado unos auriculares aislantes sobre una población de 5 personas entre 18 y 23 años con formación musical.

Todos los trabajos de grabación de remuestreo se ha hecho usando el software Audacity <sup>2</sup> y la reproducción se ha hecho usando el software VLC <sup>3</sup>. Además el trabajo se ha hecho sobre una plataforma Linux con un kernel en tiempo real y el sistema de audio JACK para evitar cualquier tipo de interrupción de la reproducción/grabación por los distintos procesos del sistema operativo.

---

<sup>1</sup><http://www.2l.no/hires/index.html>

<sup>2</sup><https://www.audacityteam.org/>

<sup>3</sup><https://www.videolan.org/vlc/index.es.html>

## 2. Comparativa de frecuencias de muestreo y resolución de muestra

Vamos a hacer una comparativa para averiguar la importancia de la frecuencia de muestreo y la resolución del audio. Además gracias a esta comparativa podremos averiguar si realmente merece la pena llevar la calidad más allá de la calidad CD.

Dividimos la comparativa en dos partes: un estudio cuantitativo y otro cualitativo.

En el estudio cuantitativo vamos a ver el tamaño de los archivos generados. Todos los archivos de esta comparativa de frecuencias y resoluciones están en formato WAV PCM. La muestra **cuerdas** tiene una duración de 1 minuto 27 segundos. La muestra **voz** tiene una duración de 33 segundos.

### Archivo cuerdas

Frecuencia/resolución	Tamaño (MB)	Porcentaje con máximo
96 KHz – 24 bits	50.4 MB	100.00 %
44.1 KHz – 16 bits	15.4 MB	30.56 %
16 kHz – 16 bits	5.6 MB	11.11 %
8 kHz – 16 bits	2.8 MB	5.56 %
8 kHz – 8 bits	1.4 MB	2.78 %

### Archivo voz

Frecuencia/resolución	Tamaño (MB)	Porcentaje con máximo
96 KHz – 24 bits	9.5	100.00 %
44.1 KHz – 16 bits	2.9	30.53 %
16 kHz – 16 bits	1.1	11.58 %
8 kHz – 8 bits	0.266	2.80 %

Podemos comprobar que el ahorro de memoria entre la calidad de estudio y la calidad CD es más que sustancial. Ahora veamos los datos del estudio cualitativo de percepción para saber si merece la pena ese gasto de memoria extra.

La metodología seguida en este experimento ha sido la siguiente: Se le han presentado a los sujetos los distintos archivos de forma aleatoria para que hicieran una primera escucha. Después se le han ido mostrando de forma aleatoria y con repeticiones los distintos archivos y se les ha pedido que valoren del 1 al 5 la calidad de escucha, siendo 1 muy mala calidad y 5 calidad excelente. En las celdas en las que hay varios valores separados por se presentan las distintas valoraciones que los sujetos han hecho del mismo archivo, debido a las repeticiones aleatorias de la reproducción.

### Archivo cuerdas

Frecuencia/resolución	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5
96/24	5/4	3/4/4	4	5/4	5
44/16	4/5	4	4	4/5	5/4
16/16	3	3	2	2	2
8/16	2	2	2	1	2
8/8	1	2/1	1	1	1

### Archivo voz

Frecuencia/resolución	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5
96/24	4	4	5	4/5	5/5
44/16	5	4	5	4/4	4/5
16/16	3	3	4	3/3	2/2
8/8	2	1	2	2/1	1

Podemos ver que los sujetos no han sido capaces de distinguir la calidad de estudio de la calidad CD. Además podemos ver que en el caso de la grabación de voz la frecuencia de muestreo a 16 kHz no ha obtenido resultados tan malos como con el archivo del cuarteto de cuerdas. Esto se debe a que el rango frecuencias de la voz humana se puede muestrear sin muchos problemas a 16 kHz. Posiblemente si hubiese sido una grabación de una voz femenina el resultado hubiese sido diferente debido a la diferencia de timbre de la voz aguda.

Todos los sujetos daban una valoración muy mala de forma instantánea a los archivos muestreados a 8 kHz.

Con estos datos podemos afirmar que **no hay diferencias apreciables a nivel de percepción entre la calidad CD y la calidad de estudio.**

## 3. Comparativa de códecs

Ahora vamos a proceder a la comparativa de usar distintos códecs con diferentes perfiles. El archivo base del que partimos es la versión FLAC del archivo de cuerdas muestreado a 44.1 kHz con una resolución de 16 bits por muestra. Los distintos archivos se han generado mediante el programa Audacity a excepción de la versión OPUS, codificada mediante la herramienta opusenc disponible en los repositorios de Ubuntu en el paquete opus-tools. Los formatos probados han sido:

- FLAC
- WAV

- MP3 con un bitrate de 32 kbps
- MP3 con un bitrate de 96 kbps
- MP3 con un bitrate de 196 kbps
- MP3 con un bitrate de 320 kbps
- AAC calidad 1
- AAC calidad 100
- AAC calidad 500
- WMA con un bitrate de 32 kbps
- WMA con un bitrate de 128 kbps
- WMA con un bitrate de 320 kbps
- OGG/Vorbis con calidad 1
- OGG/Vorbis con calidad 10
- OPUS con un bitrate de 96 kbps
- OPUS con un bitrate de 256 kbps

La comparativa de tamaños entre los distintos archivos de sonido de una duración todos de 1:27 es la siguiente:

#### **Archivo cuerdas**

Formato	Tamaño	Porcentaje respecto al original
FLAC (original)	7.7 MB	100.00 %
WAV	15.4 MB	200.00 %
AAC 500	5.4 MB	70.13 %
AAC 100	4.8 MB	62.34 %
AAC 10	4.8 MB	62.34 %
MP3 320	3.5 MB	45.45 %
MP3 192	2.1 MB	27.27 %
MP3 96	1.1 MB	14.29 %
MP3 32	0.352 MB	4.57 %
OGG 10	5.1 MB	66.23 %
OGG 1	0.795 MB	10.32 %
OPUS 256	3 MB	38.96 %
OPUS 96	1.2 MB	15.58 %
WMA 320	6 MB	77.92 %
WMA 128	1.5 MB	19.48 %
WMA 32	0.405 MB	5.26 %

Vemos varias cosas interesantes de estos datos: La diferencia entre el tamaño de AAC 100 y ACC 1 no es significativa según el compresor de Audacity. Tanto MP3 como WMA permiten comprimir el audio mucho, pudiendo alcanzar un 4.57% del tamaño original. OPUS y Vorbis también son capaces de comprimir el audio de manera significativa. En el campo de la compresión sin pérdidas (FLAC y WAV) FLAC es capaz de comprimir el audio de manera significativa, por lo que podemos decir que como formato de audio sin pérdidas es mucho más eficiente que el audio sin comprimir de WAV.

Ahora vamos a ver las diferencias en la calidad de percepción entre los distintos formatos, para ver cual obtiene una mejor relación percepción/tamaño.

La metodología seguida en este experimento ha sido la siguiente: a los sujetos se les han reproducido las distintas muestras de forma aleatoria y con repeticiones. Han sabido cual era la calidad original (FLAC) y luego se les ha ido pidiendo una valoración del 1 al 5, siendo 1 calidad muy baja y 5 calidad muy buena de cada muestra. Además se les ha preguntado explícitamente por dos comparaciones: MP3 32 kbps contra WMA 32 kbps y FLAC original contra OPUS 96 kbps.

Los resultados han sido los siguientes:

#### Archivo cuerdas - Test de percepción

	Prueba 1	Prueba 2	Prueba 3	Prueba 4	Prueba 5
FLAC	4	5	4	5	5
MP3 320	4	5	4	4	5
MP3 192	4		4	4	5
MP3 96	4	4	3	3	4
MP3 32	1	1	1	2	2
AAC 500	2	4		4	5
AAC 100	4	4		4	5
AAC 10	3/2	3	3	4	5
OPUS 265	4	4	3	5	5.
OPUS 96	4	3	4	4	5
OGG 1	3	3	3	3	3
OGG 10	3	4		4	4/4.5
WMA 320	4	5		4	5
WMA 128	2	4		4	
WMA 32	1 mejor que mp3 32	1	1 mejor que mp3 32	1	2 mejor que mp3 32

Como podemos inferir de los resultados OPUS demuestra unos resultados similares a la calidad original. WMA 32 kbps se muestra como mejor opción a MP3 32 kbps aunque ambos casos presentan una calidad de reproducción baja. OGG en calidad baja es bastante inferior a OGG en calidad alta. AAC demuestra un desempeño similar MP3 aunque ocupa más espacio en disco.

Viendo los resultados del análisis cualitativo y el test de percepción podemos afirmar que en el campo de los códecs de compresión de audio con pérdidas **OPUS presenta una mejor relación**

**calidad/tamaño.** Además podemos decir que **no existe la necesidad** de usar las versiones de 320 kbps de WMA y MP3 con respecto a sus versiones de 128 y 192 kbps, ahorrando espacio en disco (o ancho de banda) innecesario para obtener la misma calidad de reproducción.

Queda agradecer a Pedro, Antonio, Nahia, Miguel Ángel y Alba por prestarse a ser conejillos de indias en este experimento.