

UNIVERSIDAD POLITECNICA DE VALENCIA

ESCUELA POLITECNICA SUPERIOR DE GANDIA

GRADO EN ING. SIST. DE TELECOM., SONIDO E IMAGEN



UNIVERSIDAD
POLITECNICA
DE VALENCIA



ESCUELA POLITECNICA
SUPERIOR DE GANDIA

“Estudio de codecs de compresión de audio utilizando diferente estilos de música”

TRABAJO FINAL DE GRADO

Autor/a:

Rubén Tortosa Zamora

Tutor/a:

Jaime Lloret Mauri

GANDIA,

Resumen

En pocos años la industria musical ha experimentado una continua evolución fruto de su digitalización. Este proceso ha afectado al funcionamiento de toda la industria, creando cambios en dos grandes niveles. Por una parte, la digitalización ha afectado directamente a la manera en que la música se produce y compone, permitiendo que hoy en día el proceso de grabación, mezcla y masterización pueda realizarse mediante un ordenador personal, así como la composición de música directamente por ordenador, mediante instrumentos virtuales. Por otra, ha habido una evolución en las formas de almacenamiento, reproducción y distribución de la música. En pocos años se ha pasado de utilizar formatos de reproducción sonora como los discos de vinilo, seguidos por los cassettes y posteriormente el CD hasta llegar a la actualidad, cuando la música se consume mayoritariamente en formato digital, pudiéndose comprar y descargar directamente de Internet. En este escenario, existen actualmente un gran número de códecs de audio con diversas características. El presente estudio pretende identificar los diferentes géneros musicales a partir de las canciones más representativas de cada uno de ellos, y evaluar si existen códecs de audio que por sus características se adecuen mejor a unos géneros musicales o a otros.

Keywords : Música, Audio Codecs, Compresión, Calidad de audio, Generos musicales.

Extract

In a few years the music industry has experienced a continuous evolution result of digitization. This process has affected the workflow of the entire industry, creating changes in two levels. On the one hand, digitization has directly affected the way music is produced and composed, allowing today the process of recording, mixing and mastering can be performed by a personal computer, as well as the composition of computer music directly by virtual instruments. Furthermore, there has been an evolution in the forms of storage, reproduction and distribution of music. In a few years it has gone from using reproduction sound formats from vinyl records, followed by cassettes and later CD up to the present, when the music is mostly consumed in digital format, and can be purchased and downloaded directly from the Internet. In this scenario, there are currently a large number of audio codecs with different characteristics. This study aims to identify the different music genres from the most representative songs of each genre, and evaluate whether there are audio codecs which by its nature fits better on one genre or other.

Keywords: Music, Audio codecs, Compression, Audio Quality, Music Genres.

Agradecimientos:

*A mi ojos, por haberme permitido
copiar esos miles de números sin
quedarme ciego.*

Y a la gente que conozco, Gracias.

Índice de contenido

1 - Introducción:	6
1,1 – Introducción:	6
1,2 – Objetivos:	8
1,3 - Estructura del proyecto:	8
2 - Introducción teórica:	9
2,1 - Frecuencias umbral:	9
2,2 – Enmascaramiento:	9
2,3 - Umbral auditivo:	11
2,4 - Modo de compresión con pérdidas:	11
3 – Preparación previa y metodología empleada:	12
3,1 – Banco de pruebas:	12
3,1,1 – Hardware:	12
3,1,1 – Software:	12
3,1,2 – Códecs:	13
3,2 – Lista de canciones:	14
3,3 – Precauciones tomadas:	15
3,4 – Procedimiento utilizado:	15
4 – Resultados:	17
4,1 – Resultados por Códec:	17
4,1,1 – MP3:	17
4,1,2 – Nero AAC:	20
4,1,3 – Ogg Vorbis:	24
4,1,4 – WMA pro v.10:	27
4,2 – Tiempos de compresión por género musical:	30
4,3 – Casos de interés:	31
5 – Conclusiones:	33
6 – Línea de estudio posterior:	36

Índice de Figuras

Figura 1: Enmascaramiento frecuencial.....	10
Figura 2: Ratio compresión Mp3 a 32 y 40kbps.....	18
Figura 3: Ratio compresión Mp3 64 y 80kbps.....	19
Figura 4: Ratio compresión Mp3 a 128 y 160 kbps.....	19
Figura 5: Ratio compresión Mp3 a 256 y 320kbps.....	20
Figura 6: Ratio compresión AAC a 32 y 40kbps.....	21
Figura 7: Ratio compresión AAC a 128 y 160kbps.....	22
Figura 8: Ratio compresión AAC a 64 y 80 kbps.....	22
Figura 9: Ratio compresión AAC a 256 y 320 kbps.....	23
Figura 10: Ratio compresión OGG a 32 y 40kbps.....	25
Figura 11: Ratio compresión OGG a 64 y 80kbps.....	25
Figura 12: Ratio compresión OGG a 128 y 160kbps.....	26
Figura 13: Ratio compresión OGG a 256 y 320kbps.....	26
Figura 14: Ratio compresión WMA pro v10 a 32 y 48kbps.....	28
Figura 15: Ratio compresión WMA pro v10 a 96 y 128kbps.....	28
Figura 16: Ratio compresión WMA pro v10 a 160 y 192kbps.....	29
Figura 17: Ratio compresión WMA pro v10 a 256kbps.....	29
Figura 18: Tiempos de compresión medios (s).....	30
Figura 19: Factor de compresión a 32kbps.....	33
Figura 20: Factor de compresión a 128kbps.....	34

Índice de Tablas

Tabla 1: Porcentajes de compresión Lame Mp3.....	17
Tabla 2: Porcentajes de compresión obtenidos mediante el codec NeroAAC.....	21
Tabla 3: Porcentajes de compresión obtenidos mediante el códec OGG Vorbis.....	24
Tabla 4: Porcentajes de compresión resultantes de utilizar el códec WMA pro 10.....	27

1 - Introducción:

1,1 – Introducción:

Desde el inicio de los tiempos la transmisión y almacenamiento de información ha sido una necesidad y un problema. Con la aparición de los primeros hombres la información debía transmitirse de manera directa de una persona a otra y, tal vez a una tercera persona, éste tipo de tratamiento de la información, que a día de hoy se sigue utilizando, presenta varios inconvenientes como la erosión del mensaje, cada individuo lo tenía que recordar y transmitir nuevamente por lo que el mensaje se iba deteriorando con cada transmisión. Además estas gentes almacenaban la información en su cabeza, con lo poco fiable que eso resulta. Por ese motivo las canciones populares de la época pre-escritura se han ido perdiendo a lo largo de los tiempos.

Más adelante con la invención de la escritura los eruditos de la época podían conservar de manera fidedigna todo rastro de cultura o saber plasmándolo en pergaminos manuscritos que eran almacenados en grandes salas, esta manera mucho más adecuada de conservar información tenía el problema de necesitar grandes espacios, para conservar el saber y música popular eran necesarias habitaciones, o incluso edificios. Además existía el problema de la baja seguridad dado que un sólo incendio podía acabar con la cultura de generaciones.

Épocas mas tarde, con la llegada de la imprenta se mejoró de manera notable el tema de la seguridad mediante la redundancia de ejemplares, era muy posible que si desaparecía un ejemplar hubiese otro en alguna otra parte, también esta redundancia permitió que la información fuese mucho más accesible dado el incremento de obras circulantes entre la sociedad. Estos libros precisaban también de mucho espacio para ser almacenados, con lo que si bien se solucionaba un problema el otro se mantenía.

Con la invención de la gramola y más tarde el vinilo se empezó a registrar de manera precisa la música de la sociedad, en un tamaño más o menos compacto se podía almacenar una cantidad aceptable de canciones y con una calidad sólo superada por las técnicas digitales más modernas. Estos vinilos tenían el inconveniente de ser delicados y una vez rallados eran inutilizables y la información que tenían ya no se podía recuperar.

Por otra parte la invención del telégrafo y después el teléfono permitió atajar el problema de la transmisión de la información se podían transmitir mensajes de manera clara y casi instantánea, en unos segundos el mensaje era recibido. Ya no era necesario esperar unos días o meses para recibir la información.

Entonces llegaron los primeros ordenadores, la era digital empezaba, aplicaciones y juegos necesitaban de sonidos con los que interactuar con el usuario para proporcionarle una experiencia más completa. Pero los ordenadores tenían una capacidad equivalente a una caja de cerillas y dentro de ella no cabía un vinilo, sólo cabían unos 'bip' 'bup' como los que utilizaban los primeros hombres. Era necesario, igual que lo había sido siempre, encontrar un sistema que permitiese almacenar y transmitir de manera adecuada la información mediante nuestros ordenadores.

Fue entonces cuando se crearon los primeros códecs de audio, dado que digitalizar la información tal cual era inviable debido a la cantidad de información existente, se optó por recortar la cantidad de información existente y obtener algo que fuese lo más parecido a la música que queríamos digitalizar, o bien, por otra parte se creó un sistema de tablas llamado MIDI, que con proporcionarle una partitura era capaz de recomponer de manera más o menos acertada una melodía tocada por instrumentos musicales.

Una vez los bits y kilobits dejaron paso a los megas y gigas, los hercios dejaron paso a los gigahercios, internet apareció en escena y se implantó con éxito en la sociedad. Ya no eran necesarias las tablas MIDI y los códecs con pérdidas habían evolucionado lo suficiente para poder almacenar en nuestro disco duro o CD's varias discografías con una calidad algo por debajo de un CD musical. En un CD de audio cabían 74 minutos de música sin comprimir y la tecnología permitía recuperar la información obtenida aunque tuviese alguna ralla en su superficie. Los códecs seguían siendo necesarios los gigas no abundaban y la tasa de transferencia de datos de internet estaba muy limitada. Era, por lo tanto, prioritario enfocar el desarrollo a optimizar el uso del ancho de banda necesario para transmitir la información de manera satisfactoria, tanto para voz IP como para streaming de música on-line.

Hoy en día el espacio de almacenamiento no suele ser un problema, tener discografía almacenadas sin pérdidas, en formato FLAC por ejemplo, no te deja sin espacio. Pero sin embargo cada día se transmiten más y más datos a través de la red. Cada bit transmitido conlleva un coste que hay que limitar para obtener el mayor beneficio posible, más, teniendo en cuenta la futura convergencia digital donde todos los servicios de voz, televisión e internet, estos servicios también conocidos como triple play services, se unirán en la llamada red de nueva generación. Ésta red de nueva generación debe ser capaz de soportar todos los servicios de voz, video y datos con la calidad adecuada [1].

La cantidad de servicios que pueden ser ofrecidos por internet es muy amplia. De acuerdo con Mueller et al [2] hay muchas plataformas multimedia y protocolos que son usados en diferentes campos, desde el entretenimiento hasta la enseñanza. Si asumimos que los requisitos como la tolerancia a fallos, escalabilidad, la calidad del servicio y la seguridad son requisitos que ya han sido alcanzados por nuestras redes de datos, debemos considerar las características de la transmisión de

medios por streaming.

Si vamos a transmitir audio por internet o una red de área amplia, tenemos que reducir la cantidad de información enviada, esto es más crítico cuando transmitimos por Wi-Fi [3], ya que el ancho de banda del punto de acceso es compartido por todos los usuarios [4]. Entonces para mejorar la eficiencia en el almacenaje y la transmisión es necesario comprimir los archivos previamente.

Por lo tanto, como hemos visto comprimir el audio es, ha sido y probablemente será, una necesidad y es por ello que es importante estudiar cuál de los diversos códecs de audio es más eficiente tanto en el sentido estricto referente a los tamaños de salida como a la calidad sonora aparente para utilizar de la manera más eficaz y eficiente los recursos de almacenamiento y ancho de banda disponibles. Tanto para las redes actuales como para las de nueva aparición, como son las redes Ad Hoc entre Vehículos [5] o las redes inalámbricas multimedia de sensores [6].

1,2 – Objetivos:

El objetivo principal de éste proyecto es comprobar la eficacia de diversos códecs de audio según el género de música a comprimir. Con esto se pretende evaluar qué códec comprime mejor, entendiéndose por mejor en éste proyecto como menor tamaño del fichero de salida, según el género musical y así ayudar a la elección de un códec determinado para, por ejemplo, una emisora de audio on-line, o bien, a la hora de comprimir nuestra biblioteca musical.

1,3 - Estructura del proyecto:

El proyecto se estructura en cinco secciones o capítulos, acto seguido se detallará brevemente el contenido de cada una de esas secciones.

En la segunda sección tendremos una breve explicación teórica de como los codecs de compresión con pérdidas basados en la percepción psicoacústica funcionan. Se explicarán conceptos como el enmascaramiento y cómo estos codecs utilizan estos perfiles para la eliminación de datos que el oído medio humano no percibe en esas circunstancias.

En la siguiente sección, la número tres, tendremos todos los datos previos al test como el banco de pruebas utilizado, procesador, velocidad, etc, como la lista de pistas de audio utilizadas y

porqué. También se enumeran los distintos codecs empleados así como algunas precauciones que se tomaron para garantizar las mismas condiciones durante todo el test.

Dentro del cuarto bloque se expondrán y analizarán los resultados obtenidos en tres apartados. En el primero de ellos se expondrán los datos resultantes como resultado de todos los géneros y así tener una eficiencia global de los codecs. En el segundo apartado ya sí se mostrarán y analizarán los datos resultantes del estudio por géneros. Y para finalizar la sección se expondrán los resultados de interés, aquellos resultados que resultan interesantes debido a ser más particulares.

Para finalizar tendremos el apartado de conclusiones, donde señalaremos las impresiones resultados más destacables que ha aportado el estudio. También, en un siguiente apartado, se concluirá en un último apartado con cuales deberían ser las siguientes líneas de estudio para continuar con el análisis musical por géneros.

2 - Introducción teórica.

En esta sección se revisarán conceptos importantes para la codificación de audio con pérdidas como, por ejemplo, el enmascaramiento temporal y frecuencial, umbral de silencio y frecuencias audibles.

2,1 - Frecuencias umbral:

Debido a las limitaciones fisiológicas del oído humano los humanos por lo general tienen un rango auditivo que va desde los 20Hz hasta los 20.000Hz. Éste rango auditivo varía a lo largo de nuestra vida reduciéndose de manera más destacada en las altas frecuencias.

Éste límite de auditivo es aprovechado por los codecs de audio para recortar los datos de audio por debajo de los 20Hz y por encima de los 20kHz con el consiguiente ahorro de datos que ello conlleva. Oídos entrenados como pueden ser los de músicos y otras personas aficionadas a la música pueden notar la falta de esas frecuencias.

2,2 – Enmascaramiento:

En el oído humano se presentan dos fenómenos auditivos llamados enmascaramiento temporal y frecuencial.

El enmascaramiento frecuencial es un fenómeno que se produce en el oído ante la presencia de dos sonidos de diferente nivel pero frecuencias similares. Como se puede apreciar en la Fig.1 ante la presencia de un tono a 1kHz se produce una curva de enmascaramiento sobre las frecuencias cercanas. Un sonido que no produzca un nivel superior al indicado por la curva no será audible para el individuo medio.

Mediante el estudio estadístico de sujetos de prueba sometidos a diferentes tonos se han realizado las curvas de enmascaramiento similares a la de la figura 1 éstas curvas varían con la edad son también conocidas como modelo Psicoacústico. Añadir que una curva válida para un adolescente puede no ser válida para un sujeto de mediana edad.

Éste efecto es aprovechado por los diferentes codecs para eliminar sonidos que permanezcan por debajo del umbral de enmascaramiento. Como siempre toda información eliminada conlleva una reducción importante en el tamaño del fichero resultante.

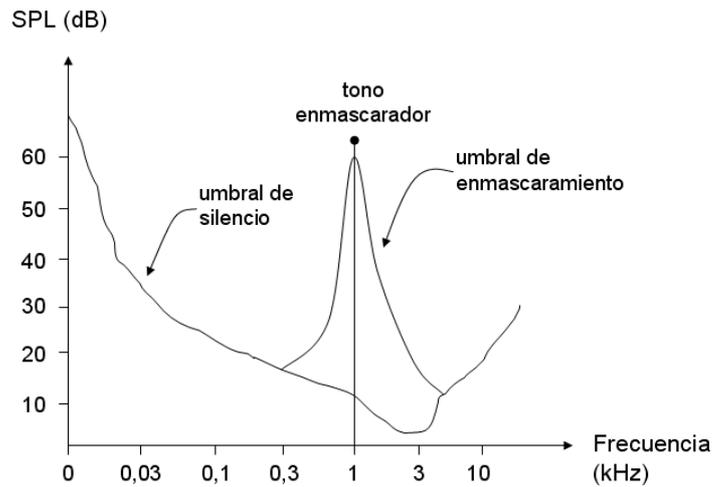


Figura 1: Enmascaramiento frecuencial

Por otro lado tenemos el fenómeno del enmascaramiento temporal, éste fenómeno ocurre cuando llegan a nuestro oído dos sonidos cercanos en el tiempo. El primer sonido, el enmascarante, producirá que el sonido posterior no sea percibido, sonido enmascarado.

Existen dos tipos pre-enmascaramiento y post-enmascaramiento. Produciéndose primero un estímulo suave y posteriormente un tono intenso, este último enmascarará igualmente al de menor amplitud, siempre y cuando estén separados en el tiempo por una diferencia menor de entre 5 y 10 ms. Dado que este fenómeno se presenta aun antes de que el tono enmascarante aparezca, implica que sea más problemático que el post-enmascaramiento.

La explicación de esta «anticipación» se basa en que la información que llega a la corteza auditiva del cerebro humano se procesa por ráfagas. Asimismo, es sabido que los sonidos intensos son procesados de forma más rápida por el cerebro que los sonidos suaves, facilitando de esta forma el fenómeno psicoacústico del pre-enmascaramiento.

El post-enmascaramiento lo encontramos cuando es el tono de mayor amplitud el que aparece primero en el tiempo seguido del de menor amplitud, percibiendo tan sólo el primer estímulo. Este fenómeno se produce cuando ambos sonidos llegan al oído humano separados en el tiempo por un intervalo mínimo de entre 30 y 60 ms aproximadamente. Esto se debe a que, percibido el tono fuerte, el oído necesita un cierto tiempo de adaptación.

Éste fenómeno no es utilizado por los codecs pero en igual de interesante.

2,3 - Umbral auditivo:

Mediante los estudios psicoacústicos se tienen gráficas sobre el nivel necesario para que un sonido sea audible. Como se puede ver en la Fig1. La curva umbral de silencio ilustra el nivel sonoro necesario para que un sonido sea audible. Ésto puede ser utilizado por los codecs para asignar una resolución menor, bitrate, a las bandas de frecuencia menos sensibles.

2,4 - Modo de compresión con pérdidas:

En éste apartado describiremos de manera genérica como un códec de audio con pérdidas comprime.

En primer lugar el audio será convertido al dominio frecuencial y las frecuencias mayores a 20kHz y menores a 20Hz serán eliminadas puesto que no serán audibles y codificarlas sería poco útil. Éstos márgenes de frecuencia pueden variar dependiendo el nivel de compresión requerido o el tipo de audio a enviar. Por ejemplo si enviamos voz podemos recortar el margen desde los 500Hz a los 5kHz y la voz será reconocible sin esfuerzo.

Acto seguido el audio se dividirá en bandas de frecuencias, la cantidad varía dependiendo de la calidad asignada,. Dentro de estas bandas se aplicará un criterio de compresión en base al enmascaramiento presente en cada una.

También se pueden aplicar criterios de predicción y otras compresiones estadísticas para ahorrar en bits.

Decir que cada códec tiene sus particularidades y su forma propia de manipular el audio para conseguir un resultado óptimo según el grupo de profesionales encargado del códec.

3 – Preparación previa y metodología empleada.

En esta sección tendremos todos los datos previos al test como el banco de pruebas utilizado, procesador, velocidad, etc, como la lista de pistas de audio utilizadas y porqué. También se enumeran los distintos codecs empleados así como algunas precauciones que se tomaron para garantizar las mismas condiciones durante todo el test y el procedimiento seguido durante todo el proceso.

3,1 – Banco de pruebas.

3,1,1 – Hardware

Se utilizó un ordenador tipo PC genérico para las pruebas cuyas características eran las siguientes:

- Procesador: Intel Core 2 Quad Q8300 @ 2,50GHz
- Memoria RAM: Mushkin 2048MB x4 @ 5-5-5-18
- Placa base: ASUS P5Q premium
- Gráfica: nVidia GeForce GTX 260 896MB Ram

3,1,1 – Software

El software mínimo utilizado en las pruebas consiste en el sistema operativo Windows 7 x64 Ultimate, un sistema operativo de 64 bit nos permitía utilizar toda la memoria física disponible, suponiendo que así el programa de codificación rendiría al máximo de su potencial.

Por otra parte, además de los drivers para el hardware para el funcionamiento óptimo del sistema, se precisaba de un programa de compresión. El programa de compresión utilizado fue Fre:ac (20130430-x64 snapshot) por utilizar el 100% del procesador, los 4 núcleos, mientras comprimía y también por permitirnos la elección de el codec NeroAAC. [7]

3,1,2 – Códecs

Los códecs de audio empleados en éste estudio fueron los de mayor implantación, es decir aquellos que los usuarios utilizan de manera más común.

3,1,2,1 – Lame Mp3

LAME es un códec MPEG Audio Layer III (MP3) de alta calidad licenciado bajo LGPL. Éste codec empezó a desarrollarse en 1998 y a día de hoy es considerado como el mejor códec MP3 a bitrates medios-altos y en bitrate variable (VBR), sobretodo gracias al trabajo de los desarrolladores y a su modelo de licencia open source [8].

3,1,2,1 – Nero AAC

Implementación comercial de ambos LC AAC y HE AAC, el códec Nero AAC fue creado por Nero AG como parte de su línea de productos Nero Digital. Por lo general es conocido por ser la mejor implementación de LC AAC en VBR. El también puede crear streams HEv1/v2 AAC con muy bajos bitrates y soporta la codificación surround multicanal [9].

3,1,2,1 – OGG Vorbis

Ogg vorbis es totalmente libre, sin patentes ni royalties. Se distribuye bajo una licencia BSD lo que implica que cualquiera puede implementarlo en aplicaciones propietarias, como por ejemplo Spotify. Vorbis está pensado para ser utilizado por un amplio espectro de aplicaciones, desde voz, 8kHz, a audio de alta calidad, 192kHz, bien sea en sonido monoaural como en sonido multicanal 5.1. [10]

3,1,2,1 – WMA pro v.10

Windows Media Audio es un tecnología propietaria de Microsoft Corporation. Microsoft afirma que WMA es capaz de proporcionar mayor calidad de sonido que Mp3 al mismo bitrate. Por su parte WMA pro v10 es un códec moderno y avanzado que permite opciones multicanal y de alta resolución. [11]

3,2 – Lista de canciones

A continuación se detallan todas las canciones utilizadas el género musical al que pertenecen, artista, título, duración en segundos y el tamaño que ocupan en KB.

Género	Artista	Canción	(s)	(KB)	Artista	Canción	(s)	(KB)
Blues	John Lee Hooker	Boom boom	259	44,732	Robert Johnson	Sweet home chicago	182	31,433
Chiptune	Anamaguchi	Endless Fantasy	352	60,769	Disasterpeace	Adventure	198	34,202
Clásica	Vivaldi	4 Station, Spring	215	37,107	Mozard	Sinfonía nº40 part II	197	33,953
Country	Hank Williams	Hey, Good Lookin'	175	30,285	Johnny Cash	At Folsom Prison	163	28,122
Dance	Cascada	Every Time We Touch	197	33,971	Milk Inc.	Never Again	197	34,107
Disco	Sylvester	You make me feel	395	68,190	Village People	Macho Man	314	54,092
DubStep	Knife Party	Bonfire	272	46,870	Skrillex	Scary Monsters	274	42,651
Fado	Amalia Rodrigues	Todo Eto e Fado	268	46,246	Maria da Fe	Até Que Vos Ma Doa	234	40,375
Funk	James Brown	Get Up	315	54,409	The Temptations	Papa Was a rolling stone	706	121,620
Gospel	Aaron Neville	Change is gonna come	280	48,352	Queen Latifah	Oh! Happy day	300	51,751
Grunge	Alice in Chains	Man in the Box	286	49,365	Nirvana	Smells like a teen spirit	302	52,052
Hard Rock	AC/DC	Back in Black	255	43,979	Guns n' Roses	Wellcome to the jungle	273	47,130
Hip-Hop	Sugar Hill Gang	Hot hot summer day	432	74,433	Sugar Hill Gang	Reaper's delight	878	141,369
Jazz	John Coltraine	My favourite Things	825	121,369	Miles Davis	So What	545	93,957
New Age	Enya	Orinoco	268	46,207	Mike Oldfield	Tubular bells	1,536	264,727
New Wave	Eurithmics	Sweet Dreams	217	37,536	Simple Minds	Don't You	263	45,357
Pop	John Lennon	Imagine	185	31,318	Beatles	Yesterday	125	21,649
Punk	The Clash	Londong Calling	199	34,424	GreenDay	Long View	239	41,188
R&B	Bell Biv Davoe	Poison	261	45,111	Janet Jackson	That the way love goes	256	45,814
Rock	Neil Young	Rocking in free world	284	48,986	Rolling Stones	It's only rock n' roll	307	52,886
Salsa	Lalo Rodriguez	Bailando	297	50,667	Frankie Ruiz	Devorame otra vez	311	53,669
Soul	Aretha Franklin	Respect!	148	25,592	Chuck Berry	Maibelline	142	24,559
Tango	Carlos Gardel	Mi buenos aires..	171	29,481	Francis Canaro	La cumparsita	195	33,631

Como se puede observar se analizaron 50 canciones de 25 géneros musicales distintos. Se

eligieron dos canciones por género para así evitar, en la medida de lo posible, interferencias intra-género como podría ser el Pop-Rock, y así analizar de la mejor manera posible cada género por separado.

Obviamente unos géneros nacen de otros y tienen, a su vez, influencias de un tercero. Es decir, los géneros musicales no son independientes entre sí. Hay géneros "independientes" de otros como el Tango del New Wave, pero hay otros como el HardRock viene del Rock que a su vez viene del Blues. Por éste motivo se eligieron las canciones más "representativas", o con más elementos característicos de cada género para así promediar y minimizar los elementos "ajenos" al género.

Cincuenta canciones de veinticinco géneros. No se utilizaron más por una cuestión de tiempo y se eligieron los géneros de los cuales nos resultó más sencillo encontrar el archivo origen sin pérdidas, ya sea en CD, Wav o Flac. Estos dos últimos asegurándose que proviniesen de CD o Vinilo.

3,3 – Precauciones tomadas

Se tomaron una serie de precauciones para intentar garantizar que todas las medidas se tomaran en las mismas condiciones y que de esta manera no fuesen desvirtuadas por procesos ejecutándose en segundo plano.

Las medidas tomadas fueron desactivar el antivirus, firewall, actualizaciones automáticas y cerrar todo aquel programa que no fuese necesario, con esto garantizamos tener todos los recursos disponibles y no tener ninguna bajada de rendimiento repentino debido a que algún programa se actualizase en segundo plano.

Por otro lado se creó un archivo "AAStart.wav" que era utilizado como baliza de salida y conseguir así tiempo para poner el cronómetro en marcha. De esta manera se conseguía que las medidas de tiempo de la primera canción empezasen en el momento adecuado.

3,4 – Procedimiento utilizado.

En primer lugar se convirtieron todas las canciones en formato Wav 44,1 kHz 16-bit estéreo, se renombraron todas las canciones para que el orden de compresión fuese al mostrado en la lista agilizando así poder apuntar los tiempos de manera más directa y sencilla.

Una vez hecho esto se configuró el programa para que añadiese al nombre del fichero las especificaciones de la comprensión, por ejemplo:

Blues_John_Lee_Hooker_Boom_Boom_MP3_32kbps.mp3

Con esto conseguíamos que las canciones apareciesen organizadas por bitrate, esto ahorra mucho tiempo a la hora de introducir en la hoja de datos todos los tamaños de salida resultante, algo que ya de por sí mismo ya llevaba bastante tiempo.

Este proceso se repitió para todos los codecs a los siguientes bitrates: 32, 40, 48, 56, 64, 80, 96, 112, 128, 160, 192, 224, 256 y 320kbps excepto para WMA que en las opciones sólo disponía de los siguientes bitrates para la configuración estéreo: 32, 48, 64, 80, 96, 128, 160, 192 y 256kbps. Siempre que se eligió la configuración ABR (average bitrate) ya sea bien eligiendo el perfil de la lista disponible o bien poniendo el bitrate objetivo en el cuadro de dialogo

Además como ya se ha comentado en el apartado anterior se midió el tiempo que tardaban los diferentes codecs a los siguientes bitrates: 80, 128, 192 y 320 kbps. Como 320kbps no estaba disponible para WMA se optó por tomar la medida a 256kbps, el máximo disponible para estéreo. En los bitrates seleccionados para la medida de tiempo se utilizó un programa de cronómetro que permitía almacenar los tiempos en un lista para después ser copiados a la hoja de cálculo.

Se anotaron todos los tamaños de salida y los tiempos requeridos para la comprensión en la hoja de cálculo y se realizaron tablas y gráficas, los resultados pueden verse en el siguiente apartado.

4 – Resultados.

A continuación se detallarán los resultados obtenidos tanto mediante gráficas como por tablas y en el punto 4,3 se evaluarán estos en busca de resultados destacables.

4,1 – Resultados por Códec

A continuación se detallan los resultados por códec, quedando en comparación directa los diferentes niveles de compresión alcanzados por el códec en los diferentes géneros. Los resultados son el promedio de los porcentajes de compresión obtenidos en las dos canciones representativas de cada género, tal y como se ha dicho anteriormente en el texto para minimizar las influencias que otro género pudiese haber tenido.

4,1,1 – MP3

En la siguiente tabla [Tabla 1] tenemos los porcentajes de compresión alcanzados por el códec Mp3 en los diferentes bitrates y géneros.

Lame Mp3	32	40	48	56	64	80	96	112	128	160	192	224	256	320
Blues	97,58%	97,05%	96,42%	95,87%	95,33%	94,27%	93,33%	92,14%	91,27%	89,44%	87,32%	85,14%	82,87%	79,13%
Chiptune	97,66%	97,14%	96,51%	95,96%	95,40%	94,33%	93,27%	92,12%	91,12%	89,10%	86,93%	84,66%	82,32%	78,41%
Clásica	97,82%	97,31%	96,76%	96,24%	95,68%	94,65%	93,51%	92,51%	91,39%	89,17%	86,98%	84,70%	82,44%	78,91%
Country	97,63%	97,11%	96,51%	95,96%	95,40%	94,36%	93,31%	92,22%	91,19%	89,08%	86,94%	84,78%	82,58%	78,89%
Dance / Tranci	97,66%	97,12%	96,55%	96,01%	95,43%	94,44%	93,30%	92,25%	91,17%	88,99%	86,72%	84,44%	82,13%	78,23%
Disco	97,63%	97,08%	96,51%	95,96%	95,38%	94,41%	93,26%	92,18%	91,07%	88,88%	86,61%	84,32%	82,04%	78,20%
Dub Step	97,71%	97,18%	96,59%	96,06%	95,48%	94,46%	93,30%	92,24%	91,10%	88,87%	86,63%	84,35%	82,11%	78,38%
Fado	97,58%	97,07%	96,45%	95,90%	95,37%	94,27%	93,30%	92,18%	91,25%	89,36%	87,42%	85,29%	83,00%	79,07%
Folk	97,93%	97,56%	97,04%	96,63%	96,28%	94,84%	94,03%	92,95%	92,18%	90,30%	88,31%	86,33%	84,44%	81,66%
Funk	97,54%	96,99%	96,42%	95,84%	95,29%	94,25%	93,20%	92,10%	91,07%	88,97%	86,74%	84,48%	82,19%	78,40%
Gospel	97,66%	97,12%	96,54%	96,00%	95,42%	94,43%	93,30%	92,23%	91,15%	88,97%	86,72%	84,44%	82,20%	78,57%
Grunge	97,76%	97,23%	96,64%	96,11%	95,53%	94,54%	93,38%	92,36%	91,26%	89,09%	86,84%	84,56%	82,27%	78,54%
Hard Rock	97,75%	97,23%	96,63%	96,08%	95,51%	94,54%	93,39%	92,36%	91,26%	89,09%	86,87%	84,60%	82,35%	78,67%
Hip Hop	97,59%	97,03%	96,44%	95,87%	95,27%	94,30%	93,14%	92,05%	90,94%	88,74%	86,48%	84,22%	82,00%	78,49%
Jazz	97,69%	97,15%	96,54%	95,98%	95,38%	94,41%	93,18%	92,23%	91,02%	88,66%	86,38%	84,23%	82,07%	78,57%
NewAge	97,78%	97,26%	96,69%	96,16%	95,59%	94,60%	93,43%	92,49%	91,34%	89,10%	86,87%	84,64%	82,44%	78,89%
NewWave	97,65%	97,11%	96,52%	95,96%	95,37%	94,40%	93,24%	92,21%	91,13%	88,95%	86,71%	84,44%	82,20%	78,56%
Pop	97,74%	97,21%	96,65%	96,11%	95,53%	94,57%	93,42%	92,40%	91,30%	89,08%	86,89%	84,64%	82,39%	78,94%
Punk	97,67%	97,13%	96,53%	95,96%	95,38%	94,44%	93,27%	92,23%	91,16%	89,00%	86,77%	84,52%	82,24%	78,58%
R&B	97,63%	97,08%	96,48%	95,93%	95,35%	94,35%	93,25%	92,15%	91,12%	89,02%	86,78%	84,48%	82,18%	78,39%
Reggae	97,68%	97,15%	96,55%	95,99%	95,41%	94,48%	93,31%	92,22%	91,12%	88,95%	86,70%	84,43%	82,18%	78,59%
Rock n' Roll	97,68%	97,15%	96,55%	96,01%	95,44%	94,39%	93,35%	92,19%	91,18%	89,12%	86,93%	84,66%	82,33%	78,50%
Salsa	97,64%	97,09%	96,49%	95,92%	95,32%	94,31%	93,11%	92,04%	90,86%	88,61%	86,37%	84,12%	81,92%	78,45%
Soul	97,58%	97,05%	96,46%	95,89%	95,36%	94,33%	93,34%	92,17%	91,13%	89,02%	86,77%	84,50%	82,26%	78,60%
Tango	97,63%	97,12%	96,54%	96,00%	95,47%	94,42%	93,40%	92,34%	91,32%	89,25%	87,14%	84,92%	82,59%	78,63%
Media	97,67%	97,15%	96,56%	96,01%	95,45%	94,43%	93,33%	92,26%	91,20%	89,07%	86,87%	84,64%	82,39%	78,73%

Tabla 1: Porcentajes de compresión Lame Mp3

En las siguientes figuras, [Figura 2 – Figura 5] se muestran algunos resultados de la Tabla 1 de manera visual para una mejor y más rápida comprensión e interpretación de los datos obtenidos.

En estas gráficas se puede apreciar como los resultados varían dependiendo del género con algún valor que destaca sobre el resto y sin nos fijamos en los datos los niveles de compresión alcanzados por esos picos se podrían equiparar a los los del bitrate anterior. Los posibles motivos se revisarán en el siguiente apartado, el 5°.

[Figura 2]: En esta figura se muestran los diferentes porcentajes de compresión obtenidos en los bitrates de 32kbps y 40kbps. El porcentaje medio de compresión es de 97,67% y 97,15% respectivamente. En ambos casos el género que muestra una peor compresión es el Funk con un 97,54% y un 96,99% respectivamente. Por el otro lado tenemos al Folk como género que mejores resultados obtiene con un 97,93% y un 97,56% con una desviación sobre la media del orden de un 0,26% y un 0,41% respectivamente.

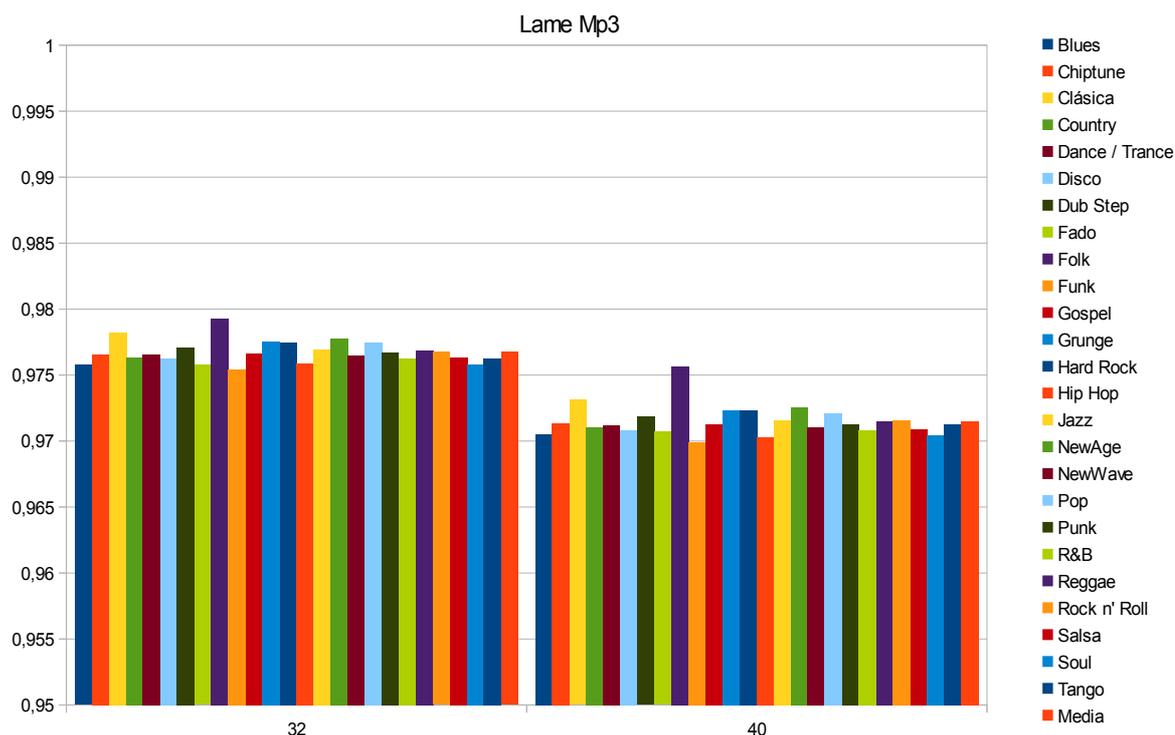


Figura 2: Ratio compresión Mp3 a 32 y 40kbps

[Figura 3]: Al igual que en la figura anterior, la siguiente tabla muestra los resultados obtenidos a 64kbps y 80kbps. Se puede apreciar como el Folk sigue destacando por encima de el resto de géneros con una mejora del 0,83% y 0,41% sobre una media del 95,45% y 94,43% respectivamente.

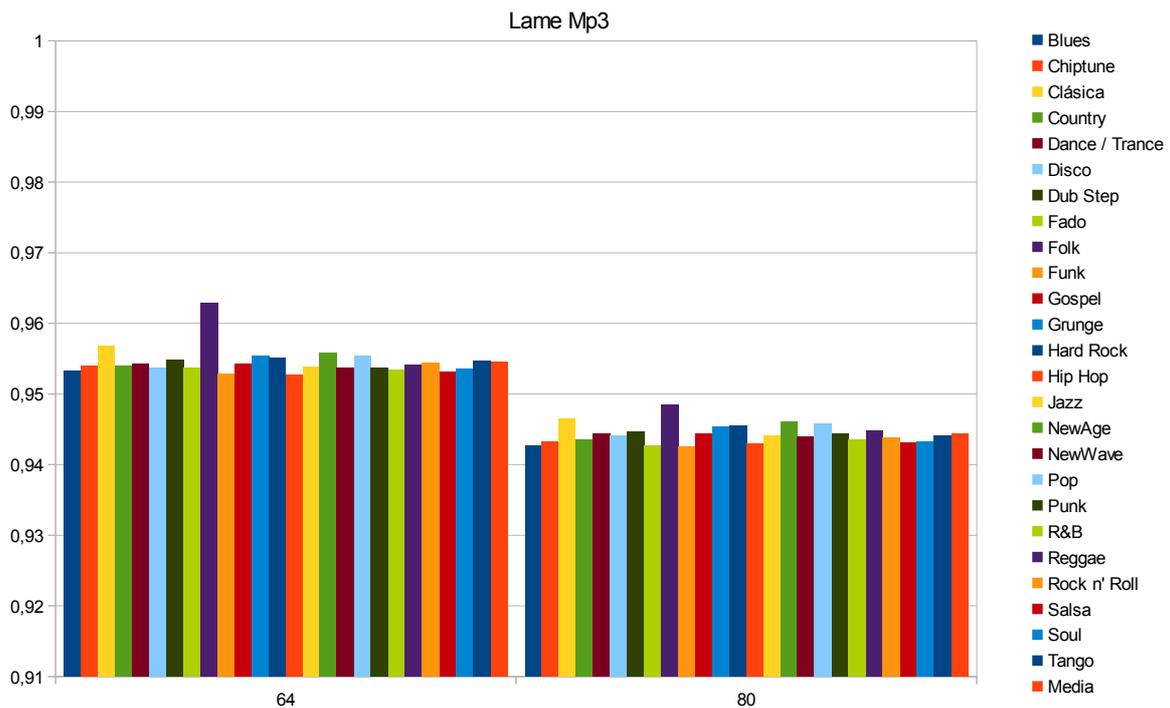


Figura 3: Ratio compresión Mp3 64 y 80kbps

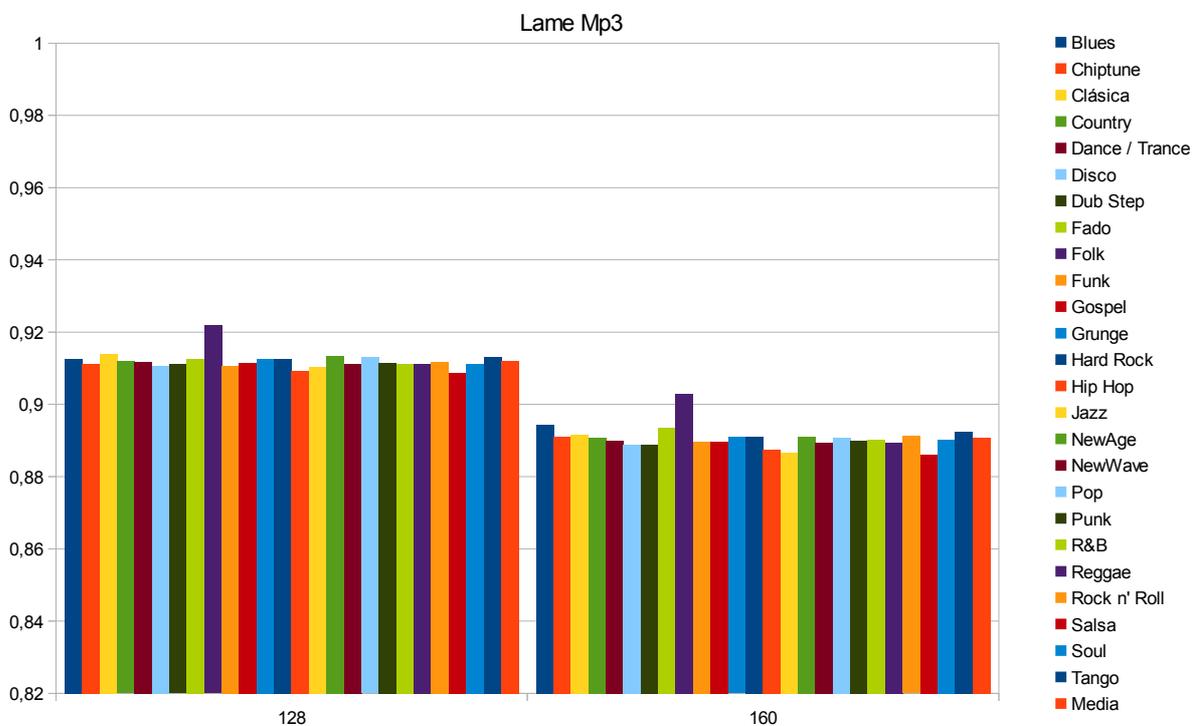


Figura 4: Ratio compresión Mp3 a 128 y 160 kbps

[Figura 4]: En la anterior figura vemos que los resultados se mantienen, pero esta vez para

los bitrates de 128kbps y 160kbps. El Folk sigue destacando sobre los demás obteniendo diferencias de varias décimas.

[Figura 5]: En la siguiente figura, la correspondiente a los bitrates de 256kbps y 320kbps, se aprecia como el Folk sigue destacando de manera notable. Por su parte, el resto de géneros no se alejan en demasía de la media conseguida en los distintos bitrates.

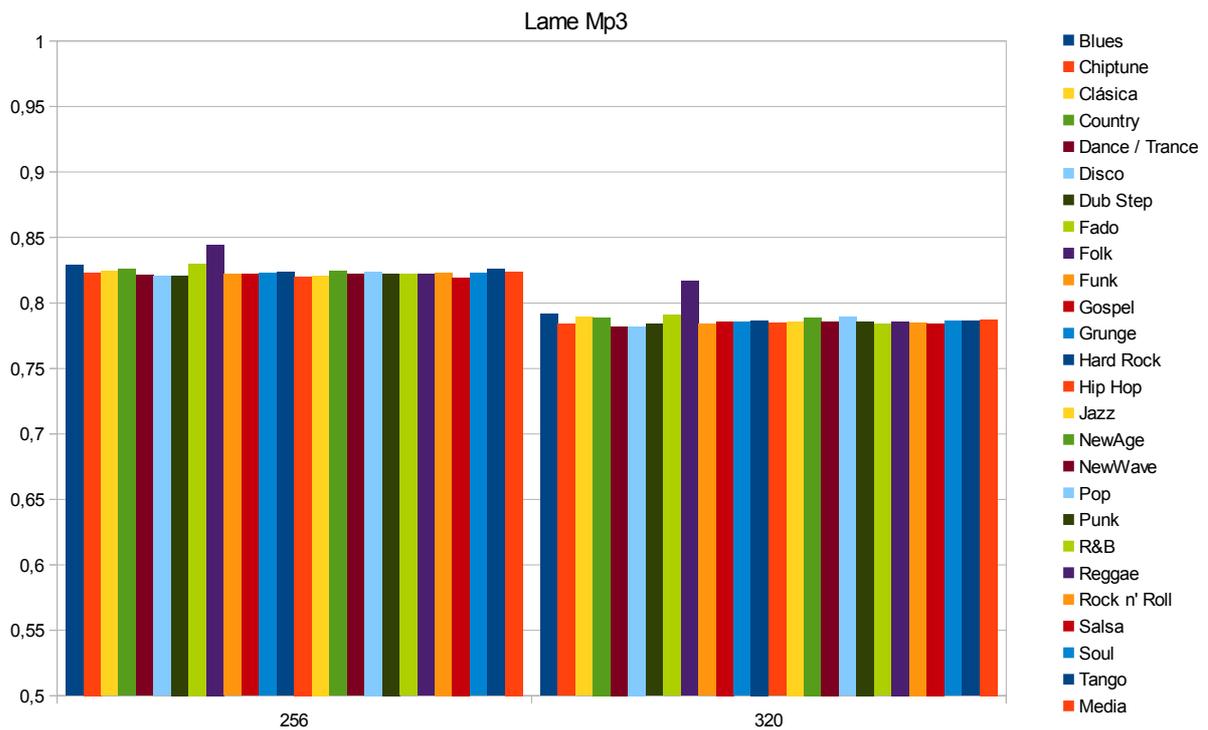


Figura 5: Ratio compresión Mp3 a 256 y 320kbps

4,1,2 – Nero AAC

De la misma manera que con el Mp3 primero se mostrarán los porcentajes de compresión en una tabla [Tabla 2] para después ilustrarlos mediante figuras [Figura 6 – Figura 9].

A diferencia de los resultados mostrados por el códec Mp3 los resultados obtenidos por el códec Nero AAC son bastante “planos”, es decir parece ser que los niveles de compresión son prácticamente independientes del género musical a tratar, salvo en altos bitrates donde se vuelve a apreciar como hay un estilo musical que consigue alcanzar niveles de compresión superiores al resto. Como anteriormente se ha comentado los posibles motivos serán analizados posteriormente.

Nero AAC	32	40	48	56	64	80	96	112	128	160	192	224	256	320
Blues	97,63%	97,07%	96,52%	95,96%	95,40%	94,27%	93,16%	92,02%	90,91%	88,68%	86,42%	84,22%	82,09%	79,19%
Chiptune	97,61%	97,05%	96,48%	95,92%	95,35%	94,22%	93,08%	91,94%	90,82%	88,54%	86,24%	83,94%	81,70%	77,19%
Clásica	97,59%	97,04%	96,48%	95,92%	95,36%	94,24%	93,11%	91,98%	90,85%	88,57%	86,28%	83,96%	81,76%	77,26%
Country	97,61%	97,05%	96,49%	95,93%	95,37%	94,24%	93,12%	91,99%	90,86%	88,60%	86,32%	84,04%	81,80%	77,30%
Dance / Tranc	97,60%	97,04%	96,50%	95,93%	95,36%	94,23%	93,10%	91,96%	90,83%	88,54%	86,26%	83,97%	81,74%	77,27%
Disco	97,61%	97,05%	96,49%	95,92%	95,35%	94,22%	93,08%	91,95%	90,82%	88,54%	86,26%	83,97%	81,72%	77,23%
Dub Step	97,60%	97,04%	96,49%	95,92%	95,35%	94,23%	93,10%	91,96%	90,83%	88,54%	86,27%	83,97%	81,74%	77,24%
Fado	97,64%	97,08%	96,52%	95,97%	95,41%	94,29%	93,17%	92,05%	90,94%	88,70%	86,44%	84,16%	81,98%	77,55%
Folk	97,60%	97,05%	96,50%	95,96%	95,41%	94,29%	93,19%	92,08%	90,94%	88,76%	86,41%	84,75%	83,18%	80,85%
Funk	97,62%	97,06%	96,49%	95,93%	95,36%	94,23%	93,09%	91,96%	90,83%	88,57%	86,29%	84,01%	81,74%	77,81%
Gospel	97,61%	97,05%	96,49%	95,93%	95,34%	94,21%	93,10%	91,97%	90,81%	88,54%	86,27%	83,99%	81,76%	77,26%
Grunge	97,61%	97,05%	96,49%	95,92%	95,36%	94,24%	93,11%	91,98%	90,85%	88,58%	86,31%	84,05%	81,82%	77,32%
Hard Rock	97,61%	97,05%	96,49%	95,93%	95,36%	94,24%	93,12%	92,00%	90,87%	88,59%	86,33%	84,07%	81,85%	77,39%
Hip Hop	97,62%	97,06%	96,50%	95,94%	95,38%	94,25%	93,13%	92,01%	90,87%	88,62%	86,33%	84,05%	81,83%	77,38%
Jazz	97,62%	97,06%	96,50%	95,93%	95,37%	94,24%	93,12%	91,99%	90,85%	88,60%	86,31%	84,04%	81,80%	77,30%
NewAge	97,62%	97,06%	96,50%	95,94%	95,37%	94,25%	93,12%	91,99%	90,87%	88,61%	86,34%	84,07%	81,84%	77,34%
NewWave	97,62%	97,06%	96,50%	95,94%	95,39%	94,25%	93,12%	91,99%	90,87%	88,60%	86,33%	84,04%	81,82%	77,34%
Pop	97,58%	97,01%	96,44%	95,88%	95,31%	94,19%	93,06%	91,93%	90,83%	88,49%	86,24%	83,94%	81,75%	77,24%
Punk	97,60%	97,04%	96,48%	95,92%	95,35%	94,22%	93,10%	91,97%	90,84%	88,56%	86,28%	83,99%	81,77%	77,28%
R&B	97,61%	97,05%	96,49%	95,92%	95,36%	94,22%	93,09%	91,96%	90,83%	88,56%	86,26%	83,97%	81,74%	77,22%
Reggae	97,60%	97,04%	96,49%	95,92%	95,35%	94,23%	93,10%	91,97%	90,84%	88,56%	86,29%	83,99%	81,77%	77,28%
Rock n' Roll	97,61%	97,04%	96,47%	95,90%	95,34%	94,21%	93,08%	91,95%	90,81%	88,55%	86,27%	84,01%	81,79%	77,33%
Salsa	97,62%	97,06%	96,51%	95,95%	95,39%	94,26%	93,14%	92,01%	90,88%	88,63%	86,34%	84,04%	81,81%	77,34%
Soul	97,61%	97,06%	96,51%	95,96%	95,40%	94,29%	93,18%	92,06%	90,94%	88,71%	86,45%	84,20%	81,99%	77,55%
Tango	97,61%	97,06%	96,50%	95,94%	95,33%	94,20%	93,10%	91,97%	90,80%	88,54%	86,25%	83,96%	81,77%	77,33%
Media	97,61%	97,05%	96,49%	95,93%	95,36%	94,24%	93,11%	91,99%	90,86%	88,59%	86,31%	84,06%	81,86%	77,55%

Tabla 2: Porcentajes de compresión obtenidos mediante el códec NeroAAC

[Tabla2]: Resultados obtenidos utilizando el códec NeroAAC. No se aprecian grandes diferencias las diferencias máximas son del orden del 0,03% en los bitrates bajos. Ya en los bitrates altos las diferencias llegan a 3,3% en Géneros como el Folk.

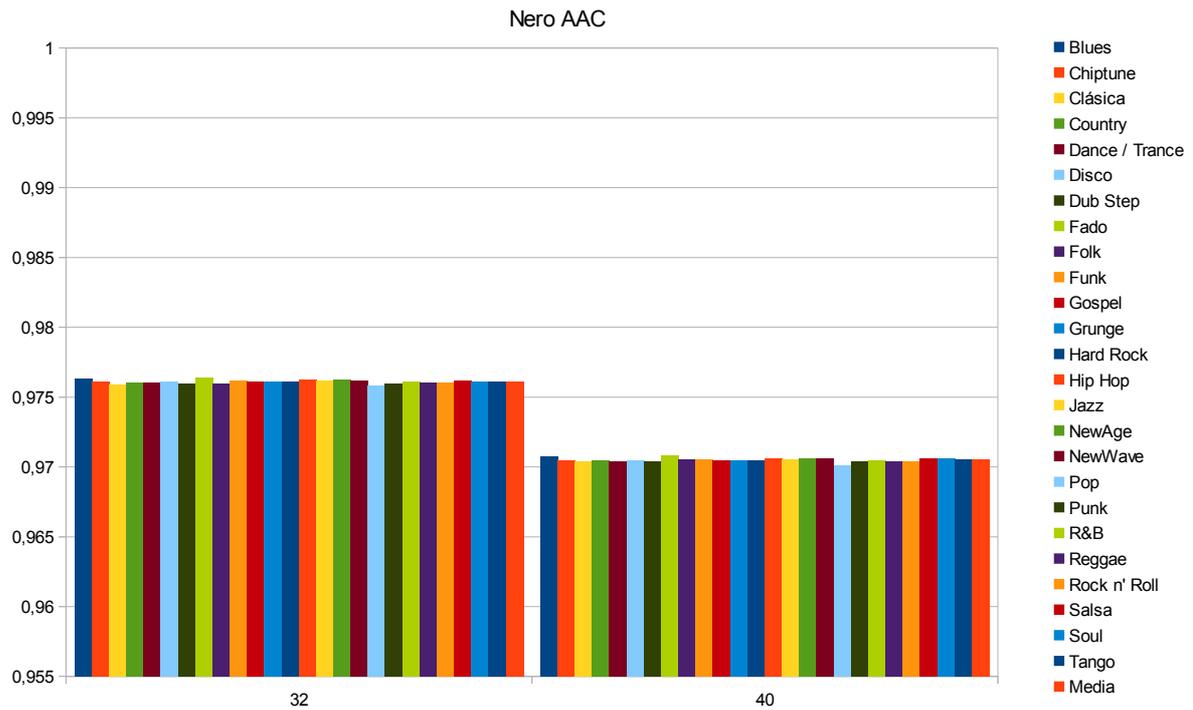


Figura 6: Ratio compresión AAC a 32 y 40kbps

[Figura 6]: La anterior figura muestra los resultados obtenidos para los bitrates de 32 y 40kbps, donde las diferencias son mínimas y podría decirse que nada relevantes.

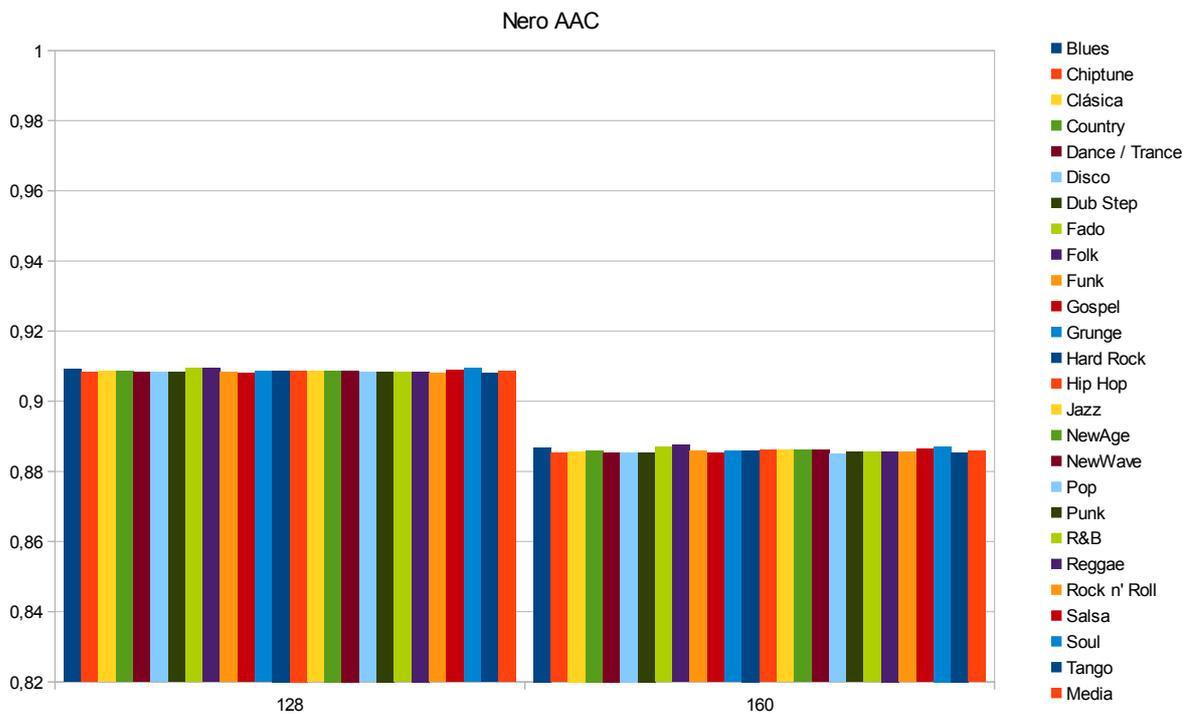


Figura 7: Ratio compresión AAC a 128 y 160kbps

[Figura 7]: Al igual que en la figura 6 no se aprecia nada destacable a los 128 y 160kbps.

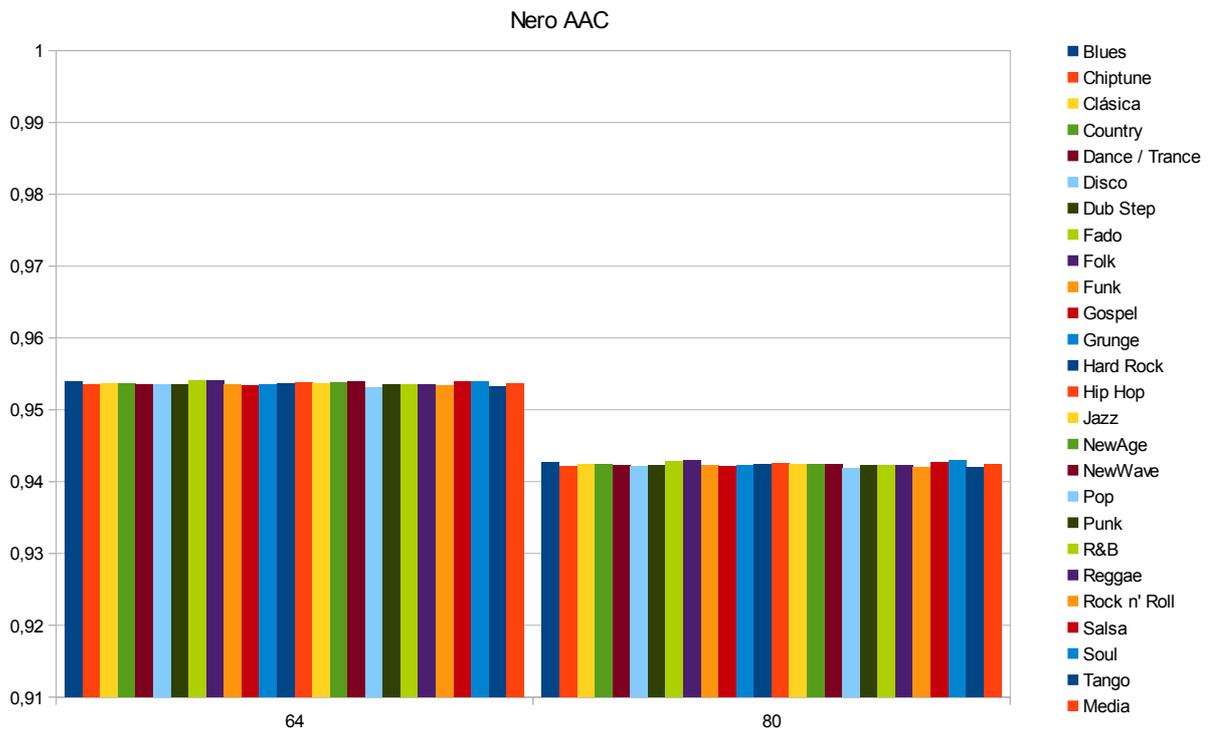


Figura 8: Ratio compresión AAC a 64 y 80 kbps

[Figura 8]: En la anterior figura no se aprecia ninguna gran diferencia en los resultados a los bitrates de 64 y 80 kbps, diferencias mínimas del orden del 0,01% entre los diferentes géneros.

[Figura 9]: En la siguiente figura si que se aprecian diferencias notables, a los 256kbps el Folk destaca sobre la média con un 1,32% de mejora. Ya en los 320kbps se puede obserbar como el Folk destaca con un 3,3% de mejora sobre la media, también se puede ver como el Blues tambien mejora sobre la media algo menos que la mitad que el Folk con un 1,64%.

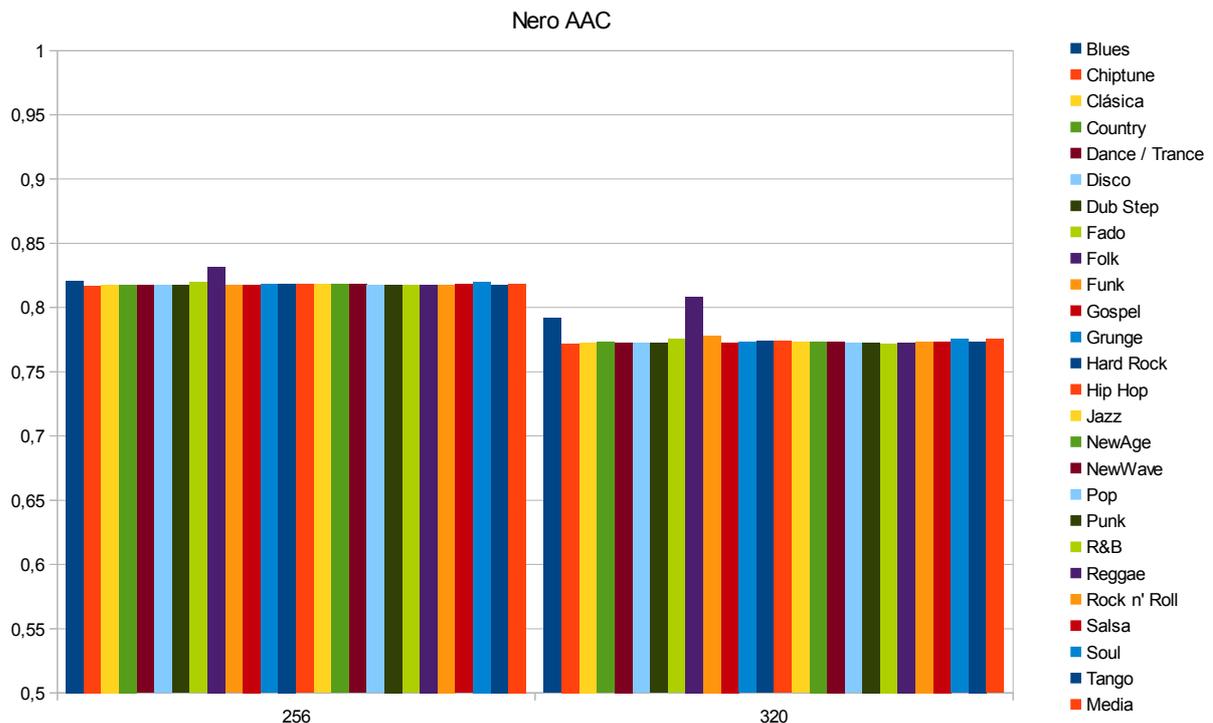


Figura 9: Ratio compresión AAC a 256 y 320 kbps

4,1,3 – Ogg Vorbis

Resultados obtenidos con el códec OGG, los resultados se organizan por bitrate de salida y primero mediante una tabla [Tabla 3] para después ilustrarlos mediante varias figuras [Figura 10 – Figura 13].

OGG Vorbis obtiene unos resultados de compresión también planos, independientemente del género musical a tratar, sin embargo en este códec no se obtienen los picos obtenidos por otros codecs en según que géneros, creemos que es debido a cómo Ogg maneja los datos a la hora de comprimirlos.

[Tabla 3]: Al igual que en las anteriores subsecciones la siguiente tabla muestra los porcentajes de compresión obtenidos para el códec OGG, y lo hace de una manera numérica. Como se puede apreciar las diferencias obtenidas con éste códec son mínimas, del orden de un 0,01% a un 0,2%, como ya hemos comentado diferencias mínimas.

OGG Vorbis	32	40	48	56	64	80	96	112	128	160	192	224	256	320
Blues	97,71%	97,15%	96,57%	96,01%	95,43%	94,31%	93,15%	92,02%	90,87%	88,63%	86,36%	84,08%	81,82%	77,28%
Chiptune	97,69%	97,12%	96,54%	95,97%	95,39%	94,26%	93,09%	91,95%	90,79%	88,53%	86,23%	83,93%	81,65%	77,06%
Clásica	97,70%	97,13%	96,55%	95,98%	95,41%	94,28%	93,11%	91,98%	90,83%	88,57%	86,27%	84,00%	81,71%	77,14%
Country	97,70%	97,13%	96,55%	95,98%	95,40%	94,28%	93,11%	91,97%	90,82%	88,57%	86,28%	83,99%	81,71%	77,13%
Dance / Trance	97,68%	97,12%	96,53%	95,97%	95,38%	94,25%	93,09%	91,94%	90,79%	88,53%	86,23%	83,93%	81,64%	77,06%
Disco	97,69%	97,12%	96,54%	95,97%	95,38%	94,25%	93,09%	91,94%	90,79%	88,52%	86,23%	83,93%	81,64%	77,05%
Dub Step	97,69%	97,12%	96,54%	95,97%	95,39%	94,25%	93,09%	91,95%	90,79%	88,53%	86,23%	83,93%	81,64%	77,05%
Fado	97,72%	97,16%	96,59%	96,03%	95,45%	94,34%	93,18%	92,06%	90,92%	88,69%	86,42%	84,16%	81,90%	77,38%
Folk	97,68%	97,12%	96,53%	95,96%	95,38%	94,25%	93,07%	91,94%	90,78%	88,52%	86,22%	83,93%	81,63%	77,10%
Funk	97,69%	97,12%	96,54%	95,97%	95,39%	94,26%	93,09%	91,95%	90,80%	88,53%	86,24%	83,94%	81,66%	77,07%
Gospel	97,69%	97,12%	96,54%	95,97%	95,38%	94,25%	93,09%	91,95%	90,79%	88,53%	86,23%	83,93%	81,64%	76,96%
Grunge	97,70%	97,13%	96,55%	95,98%	95,40%	94,27%	93,11%	91,97%	90,81%	88,56%	86,27%	83,98%	81,69%	77,12%
Hard Rock	97,70%	97,14%	96,56%	95,99%	95,41%	94,28%	93,12%	91,99%	90,84%	88,59%	86,30%	84,02%	81,74%	77,18%
Hip Hop	97,70%	97,13%	96,55%	95,98%	95,39%	94,26%	93,11%	91,97%	90,81%	88,55%	86,27%	83,97%	81,69%	77,12%
Jazz	97,70%	97,13%	96,55%	95,98%	95,40%	94,28%	93,11%	91,97%	90,82%	88,57%	86,27%	83,98%	81,70%	77,13%
NewAge	97,71%	97,14%	96,56%	95,99%	95,42%	94,29%	93,13%	91,99%	90,84%	88,59%	86,30%	84,02%	81,75%	77,18%
NewWave	97,70%	97,13%	96,55%	95,99%	95,40%	94,28%	93,12%	91,98%	90,83%	88,58%	86,29%	84,00%	81,73%	77,16%
Pop	97,68%	97,12%	96,53%	95,97%	95,38%	94,25%	93,08%	91,94%	90,79%	88,53%	86,22%	83,93%	81,64%	77,05%
Punk	97,69%	97,13%	96,54%	95,98%	95,39%	94,26%	93,10%	91,96%	90,81%	88,55%	86,26%	83,97%	81,68%	77,11%
R&B	97,69%	97,12%	96,54%	95,97%	95,38%	94,25%	93,09%	91,94%	90,79%	88,52%	86,23%	83,92%	81,63%	77,05%
Reggae	97,69%	97,13%	96,54%	95,97%	95,39%	94,26%	93,10%	91,96%	90,80%	88,55%	86,26%	83,96%	81,67%	77,10%
Rock n' Roll	97,69%	97,13%	96,54%	95,97%	95,39%	94,25%	93,09%	91,94%	90,79%	88,53%	86,23%	83,93%	81,64%	77,06%
Salsa	97,70%	97,14%	96,56%	95,99%	95,40%	94,28%	93,12%	91,99%	90,83%	88,59%	86,31%	84,02%	81,75%	77,19%
Soul	97,71%	97,15%	96,58%	96,02%	95,44%	94,33%	93,18%	92,05%	90,91%	88,68%	86,42%	84,15%	81,89%	77,37%
Tango	97,69%	97,12%	96,53%	95,97%	95,39%	94,26%	93,08%	91,95%	90,79%	88,53%	86,23%	83,94%	81,65%	77,06%
Media	97,70%	97,13%	96,55%	95,98%	95,40%	94,27%	93,11%	91,97%	90,82%	88,56%	86,27%	83,98%	81,70%	77,13%

Tabla 3: Porcentajes de compresión obtenidos mediante el códec OGG Vorbis

[Figura 10]: En la siguiente tabla tenemos, esta vez de manera gráfica, los porcentajes pertenecientes a los bitrates de 32kbps y 40kbps. Diferencias mínimas donde ningún género destaca mucho mas que otro.

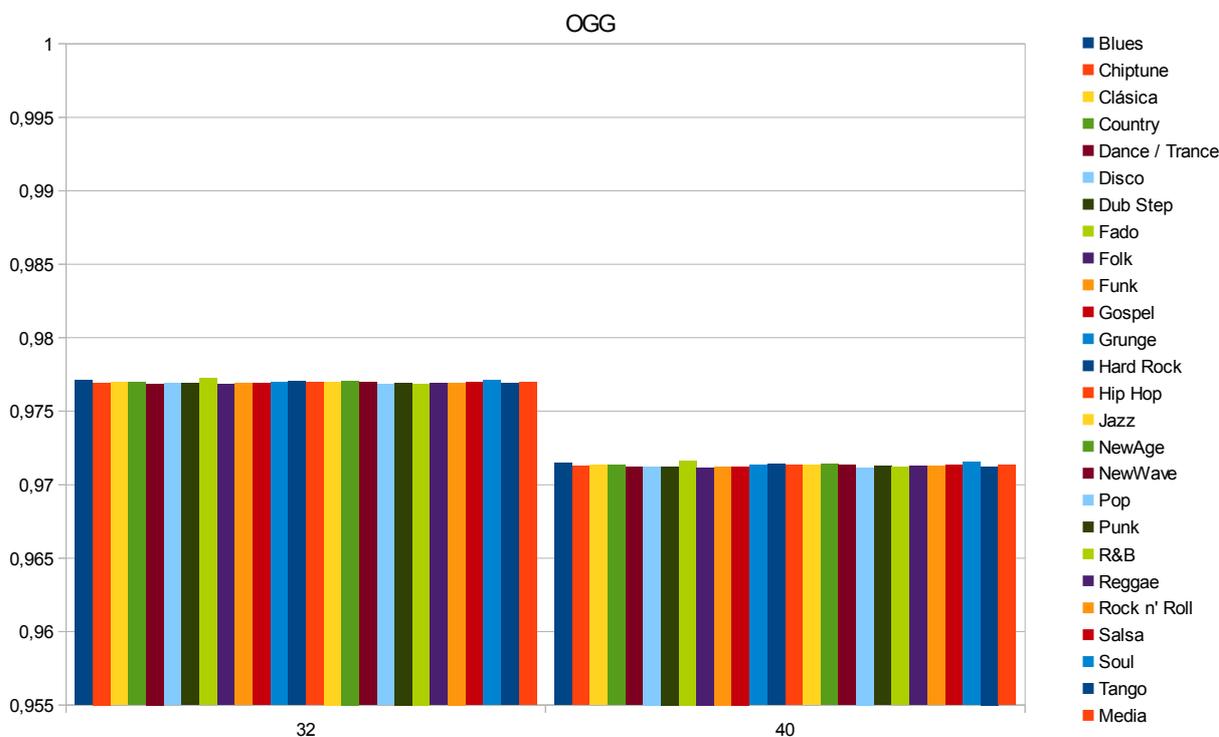


Figura 10: Ratio compresión OGG a 32 y 40kbps

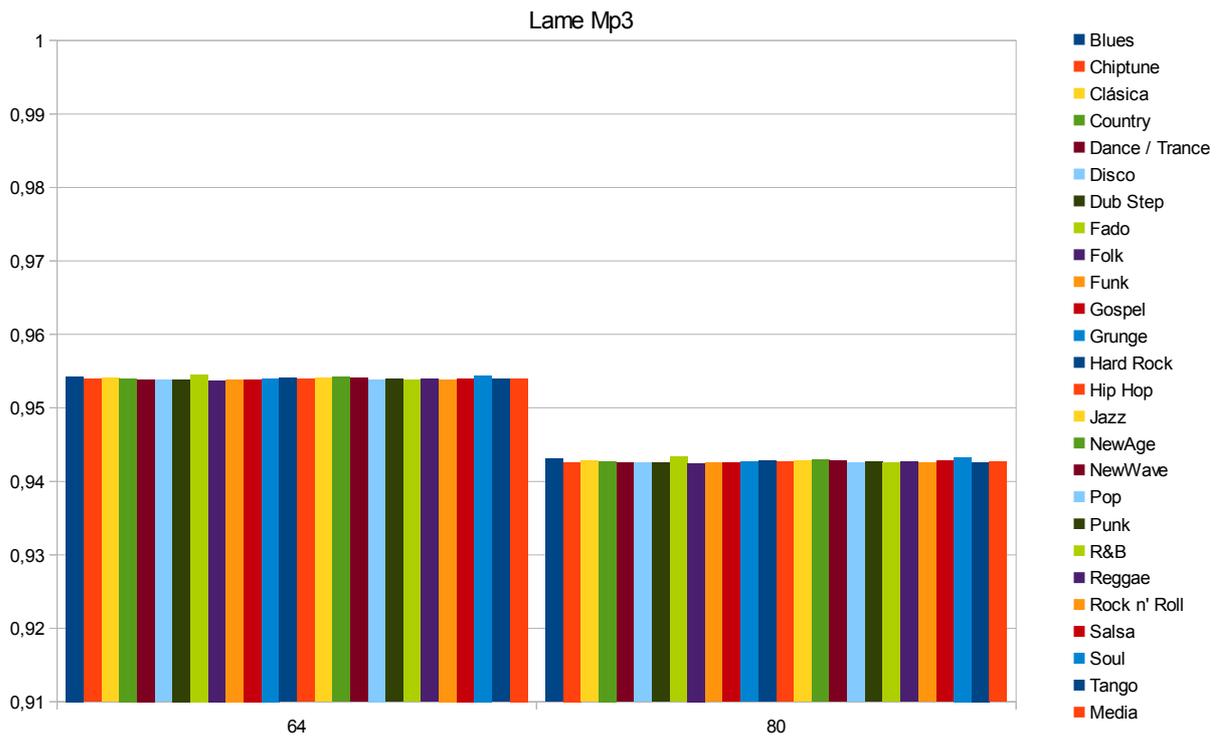


Figura 11: Ratio compresión OGG a 64 y 80kbps

[Figura 11]: La anterior figura muestra los resultados de 64 y 80kbps. Nada destacable.

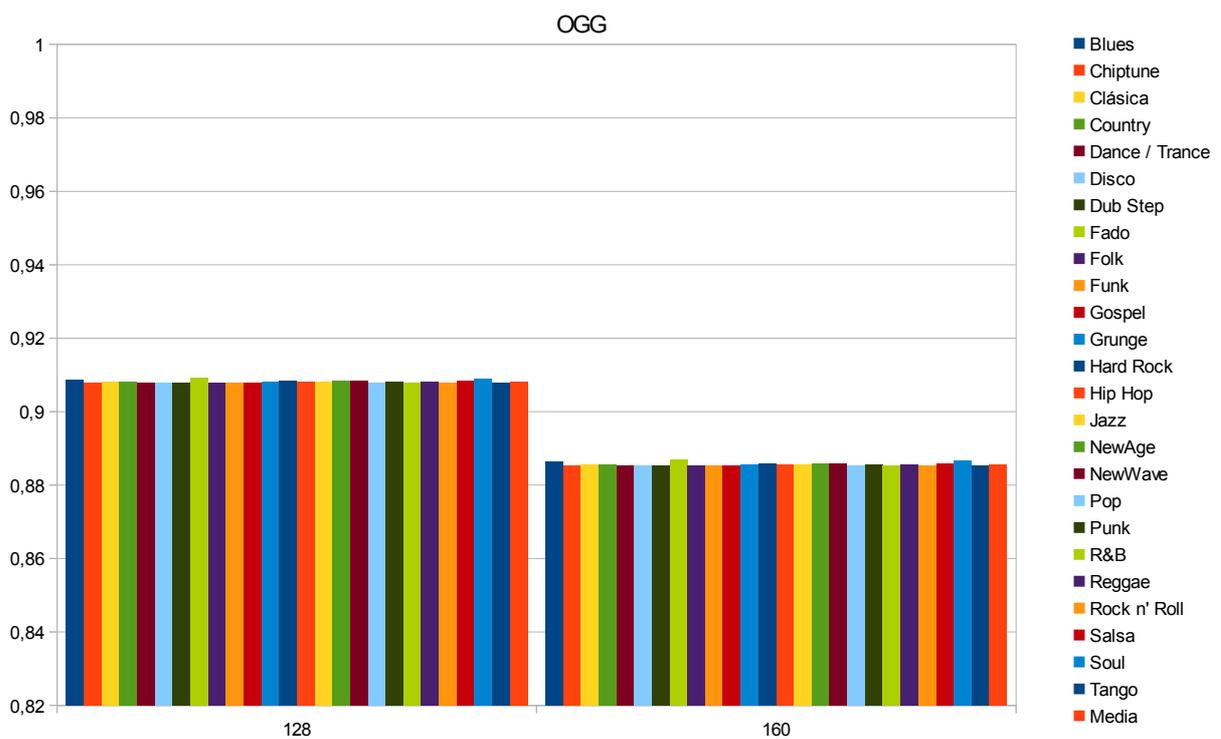


Figura 12: Ratio compresión OGG a 128 y 160kbps

[Figura 12]: La anterior figura, al igual que las precedentes muestra los resultados de manera gráfica, esta vez para los bit rates de 128kbps y 160kbps. Otra vez diferencias mínimas con una desviación típica sobre la media del orden de unas pocas centésimas porcentuales.

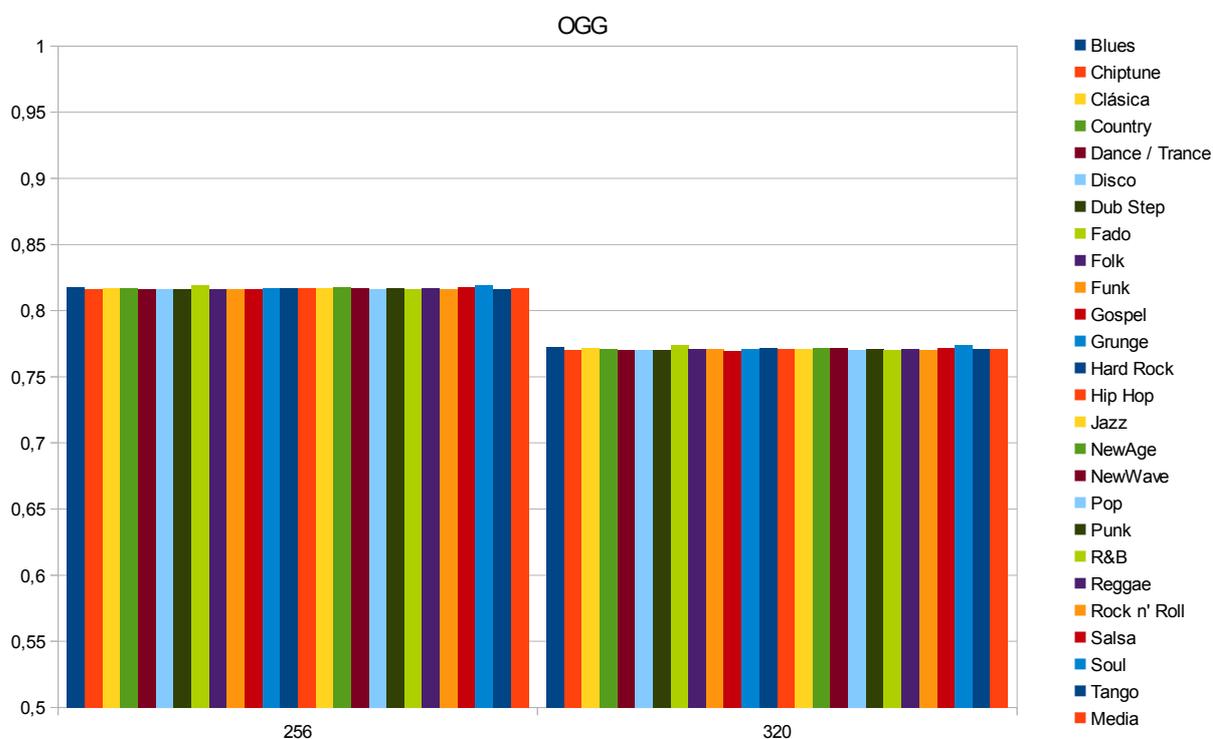


Figura 13: Ratio compresión OGG a 256 y 320kbps

[Figura 13]: Para finalizar la subsección la figura anterior ilustra los resultados obtenidos para los bitrates más altos el de 256kbps y el de 320kbps. Una vez más el códec nos devuelve unos resultados sin nada más destacable que su estabilidad con una medias porcentuales de compresión de un 81,70% para 256kbps y de un 73,13% para el bitrate de 320kbps.

4,1,4 – WMA pro v.10

Y para finalizar tenemos los resultados obtenidos con el códec WMA en su versión profesional número 10, como ha sido habitual a lo largo del texto continuaremos con una tabla [Tabla 4] en la que se muestran los porcentajes para, acto seguido, ilustrarlos con varias figuras [Figura 14 – Figura 17].

Se puede apreciar como WMA obtiene unos ratios de compresión absolutamente independientes del género musical produciéndose como máximo unas ligeras variaciones del 0,02%.

[Tabla 4]: La siguiente tabla muestra los porcentajes de compresión obtenidos para los bitrates disponibles. Se puede apreciar como los porcentajes se mueven siempre dentro de la misma décima.

WMA Pro 10	32	48	64	80	96	128	160	192	256
Blues	97,67%	96,53%	95,39%	94,26%	93,12%	90,85%	88,58%	86,31%	81,77%
Chiptune	97,66%	96,51%	95,38%	94,24%	93,10%	90,83%	88,55%	86,27%	81,71%
Clásica	97,66%	96,52%	95,38%	94,25%	93,11%	90,83%	88,56%	86,28%	81,73%
Country	97,66%	96,54%	95,37%	94,24%	93,10%	90,82%	88,54%	86,26%	81,70%
Dance / Trance	97,66%	96,52%	95,39%	94,26%	93,11%	90,85%	88,56%	86,29%	81,74%
Disco	97,67%	96,53%	95,40%	94,27%	93,13%	90,87%	88,60%	86,33%	81,79%
Dub Step	97,67%	96,52%	95,39%	94,25%	93,12%	90,84%	88,57%	86,29%	81,76%
Fado	97,67%	96,54%	95,40%	94,27%	93,13%	90,86%	88,60%	86,33%	81,79%
Folk	97,66%	96,52%	95,39%	94,25%	93,11%	90,84%	88,57%	86,29%	81,75%
Funk	97,68%	96,54%	95,40%	94,27%	93,13%	90,86%	88,59%	86,32%	81,77%
Gospel	97,67%	96,53%	95,40%	94,26%	93,13%	90,86%	88,59%	86,32%	81,78%
Grunge	97,67%	96,53%	95,40%	94,27%	93,13%	90,86%	88,59%	86,32%	81,79%
Hard Rock	97,67%	96,53%	95,40%	94,26%	93,12%	90,85%	88,57%	86,30%	81,75%
Hip Hop	97,68%	96,54%	95,41%	94,27%	93,14%	90,87%	88,60%	86,33%	81,80%
Jazz	97,68%	96,54%	95,41%	94,28%	93,14%	90,87%	88,60%	86,34%	81,80%
NewAge	97,68%	96,54%	95,41%	94,27%	93,14%	90,87%	88,60%	86,33%	81,79%
NewWave	97,67%	96,53%	95,40%	94,26%	93,13%	90,85%	88,58%	86,31%	81,77%
Pop	97,66%	96,50%	95,37%	94,24%	93,10%	90,82%	88,53%	86,27%	81,69%
Punk	97,67%	96,53%	95,40%	94,27%	93,13%	90,86%	88,60%	86,33%	81,79%
R&B	97,66%	96,52%	95,40%	94,25%	93,12%	90,85%	88,57%	86,29%	81,75%
Reggae	97,67%	96,53%	95,40%	94,26%	93,13%	90,86%	88,59%	86,32%	81,78%
Rock n' Roll	97,67%	96,53%	95,39%	94,26%	93,12%	90,85%	88,57%	86,30%	81,75%
Salsa	97,67%	96,54%	95,40%	94,27%	93,14%	90,87%	88,60%	86,33%	81,79%
Soul	97,66%	96,53%	95,39%	94,26%	93,12%	90,86%	88,59%	86,32%	81,78%
Tango	97,66%	96,52%	95,39%	94,26%	93,12%	90,86%	88,59%	86,32%	81,78%
Media	97,67%	96,53%	95,39%	94,26%	93,12%	90,85%	88,58%	86,31%	81,76%

Tabla 4: Porcentajes de compresión resultantes de utilizar el códec WMA pro 10

[Figura 14]: La siguiente figura muestra los resultados obtenidos para los bitrates de 32kbps y 48kbps de una manera gráfica, como se puede apreciar los resultados obtenidos son prácticamente idénticos entre los diferentes géneros musicales. La media se sitúa en el 97,67% y el 96,53% respectivamente.

[Figura 15]: En la segunda figura de la siguiente página se puede comprobar la regularidad que el códec WMA da como resultado a los 96kbps y los 128kbps. La media está en un 95,39% y un 94,26% para los bitrates de 96 y 128kbps.

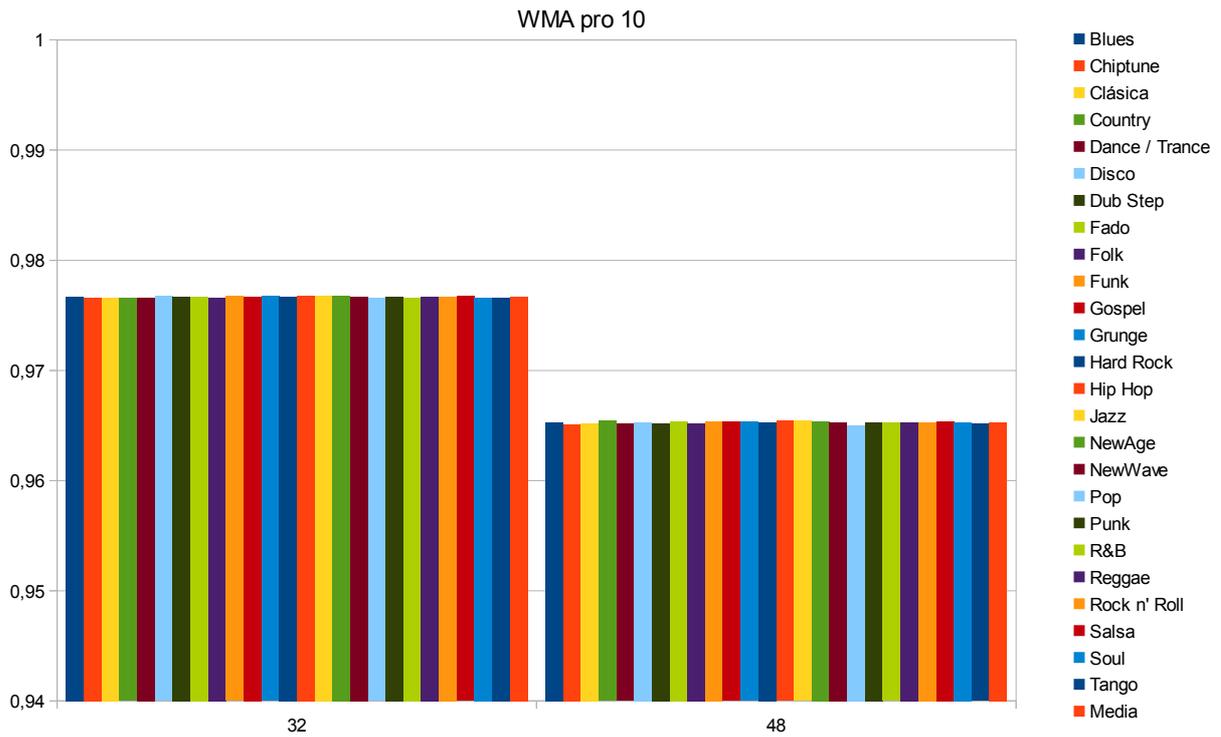


Figura 14: Ratio compresión WMA pro v10 a 32 y 48kbps

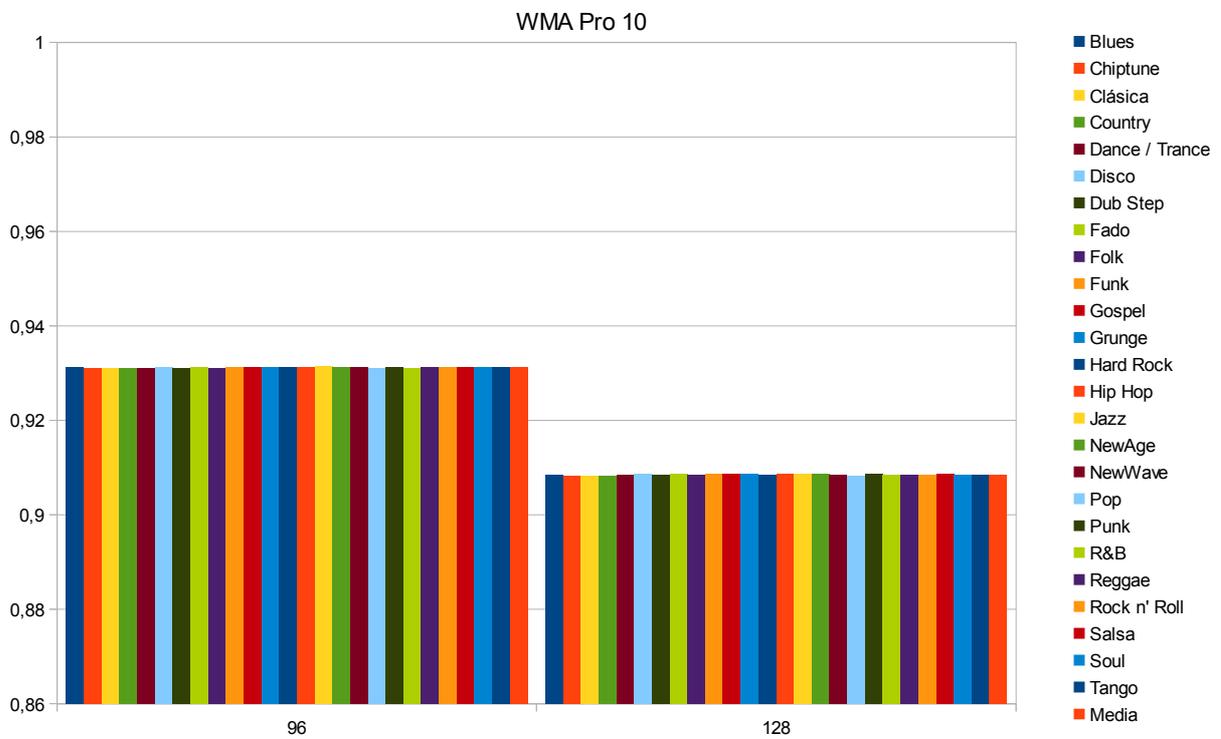


Figura 15: Ratio compresión WMA pro v10 a 96 y 128kbps

[Figura 16]: La siguiente figura ilustra los resultados planos típicos de este códec a 160 y 192kbps.

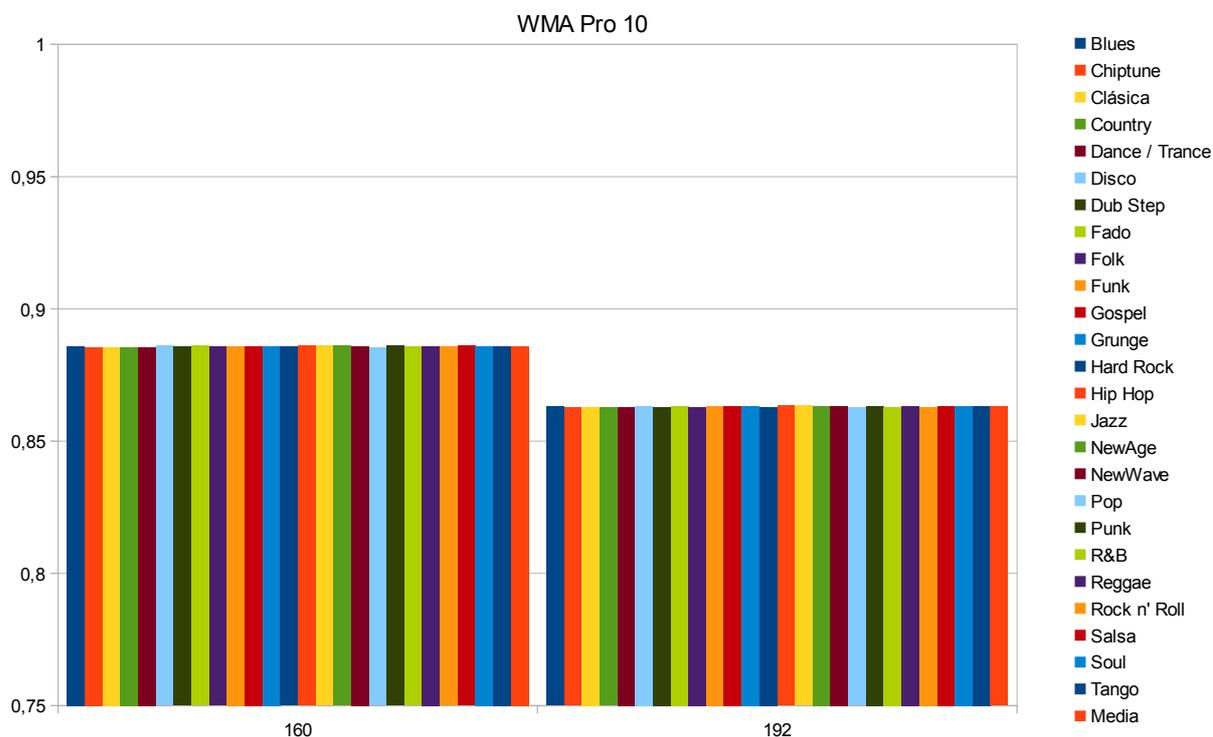


Figura 16: Ratio compresión WMA pro v10 a 160 y 192kbps

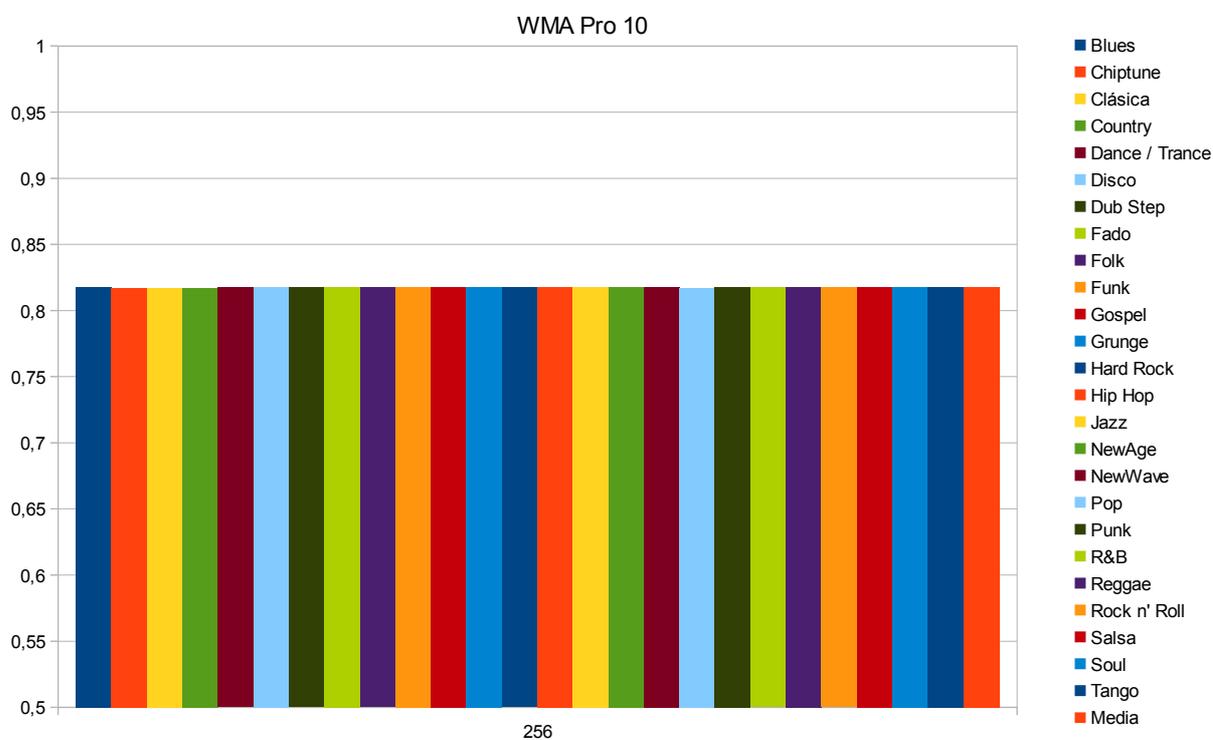


Figura 17: Ratio compresión WMA pro v10 a 256kbps

[Figura 17]: La anterior figura mostraba los resultados obtenidos para el bitrate más alto disponible 256kbps, a diferencia del códec NeroAAC, WMA pro v.10 no muestra ninguna diferencia entre género a bitrates altos.

4,2 – Tiempos de compresión por género musical.

En este apartado se mostrarán los tiempos de compresión medios que cada códec ha precisado para comprimir un segundo de cada género musical, estos tiempos han sido promediados entre las dos canciones elegidas de cada género. A primera vista, [Figura 18], puede parecer que las medidas de tiempo sobre el códec WMA pueden ser equivocadas pero han sido repetidas en varias ocasiones para cerciorarse de que son correctas.

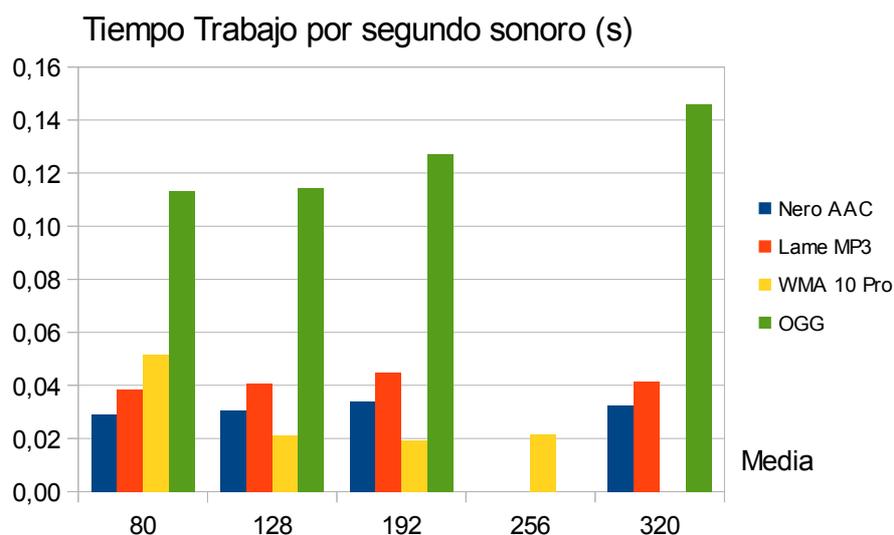


Figura 18: Tiempos de compresión medios (s)

Tras calcular el tiempo que se tardaba en comprimir un segundo de cada canción se calculó la media de cada género y posteriormente se realizó un promedio de lo que se tardaría en comprimir un segundo de cualquier canción. Éste resultado es el que se muestra en la Figura 18.

Se puede apreciar la gran diferencia existente entre OGG y el resto de codecs, resultando bastante más lento a las hora de comprimir, esta lentitud es bastante más notable cuando se comprime por lotes. Suponemos que la falta de optimización se debe a que es un codec open source

100% y tampoco ha tenido el éxito a la hora de implantarse que tuvo en su día el Mp3, algo que hubiese producido que la comunidad open source se hubiese volcado más a fondo en el trabajo de este códec optimizando así todos los algoritmos de compresión.

Nero AAC y Lame MP3 se muestran estables a través de todos los bitrates mostrando un tiempo de compresión por segundo bastante regular. En caso del Lame Mp3 ha tenido mucho tiempo para ser optimizado y, de hecho, es el único códec de MP3 que se sigue implementando actualmente. Con el Nero AAC tenemos una empresa grande detrás que lucha por hacer su códec AAC mejor que el de la competencia.

Por último tenemos WMA en su versión profesional número 10. Tenemos un caso extraño, pues éste códec se toma casi el doble de tiempo para comprimir a bajos bitrates que a altos, creemos que esto es debido a los algoritmos de compresión, que a bajos bitrates son más exigentes para intentar sacar el mejor sonido posible. El codec, propiedad de microsoft, es el que se muestra más veloz a altos bitrates, tardando un 50% menos tiempo que el códec Nero AAC y siendo siete veces más rapido que el códec libre OGG. Se puede decir que los ingenieros de Microsoft han realizado un muy buen trabajo de optimización.

4,3 – Casos de interés

Como ha quedado ilustrado tanto en tablas como en las figuras anteriores relativas al porcentaje de compresión ha destacado tanto en AAC cómo en MP3 unos picos de compresión superiores a la media por bastantes puntos incluso en algunas ocasiones llegar a alcanzar los valores de compresión media del bitrate inferior.

Por ejemplo, Blues Mp3 a 224kbps tiene una compresión del 86,33% siendo la media de ese bitrate 84,64 consiguiendo una mejora de 1,69%. Siguiendo en Mp3 Tenemos que Blues a 64kbps consigue una compresión del 96,28% siendo superior a la media del bitrate anterior 56kbps Avg. 96,01%.

Se trata de los géneros Folk y Blues. En primer lugar hay que analizar la composición sonora de estos géneros, o al menos de estas canciones en particular.

En el Folk nos encontramos al cantante, una guitarra solista y, a veces, una armónica, esto reduce tanto en ancho de banda como el margen dinámico. Como sabemos la voz humana se centra en las frecuencias medias y tanto la guitarra como la armónica presentan un ancho de banda más

limitado que por ejemplo, un piano o la música electrónica. Además este estilo de música presenta pocos transitorios abruptos y es todo muy homogéneo y fluido.

Por otro lado en el Blues, tenemos unas características parecidas, sobre todo en el periodo clásico, más adelante se incorporaría el piano y la percusión, siempre dependiendo del artista en cuestión y sus preferencias musicales.

Éstas características permiten a los codecs de audio aunar las bandas de frecuencia para una mejor utilización del bit rate disponible, y reducir el ancho de banda a filtrar consiguiendo ahorrar bastante espacio, pues comprimir frecuencias que no se van a utilizar sería redundante.

5 – Conclusiones.

Dado que el objetivo de este estudio era analizar el rendimiento de los principales codecs de audio utilizados actualmente y no la calidad sonora aparente resultante para posteriormente y junto a otros estudios comprobar si existía diferencia a la hora de comprimir los diferentes géneros musicales y además ayudar a implementar un algoritmo que permitiese la elección óptima de un códec o otro a la hora de establecer una emisora on-line o comprimir nuestra colección de discos ahorrando el máximo de tiempo y espacio/ancho de banda posible.

Para comenzar tenemos las dos siguientes figuras, [Figura 19 y Figura 20], que muestran una comparación directa del porcentaje de compresión por género y códec.

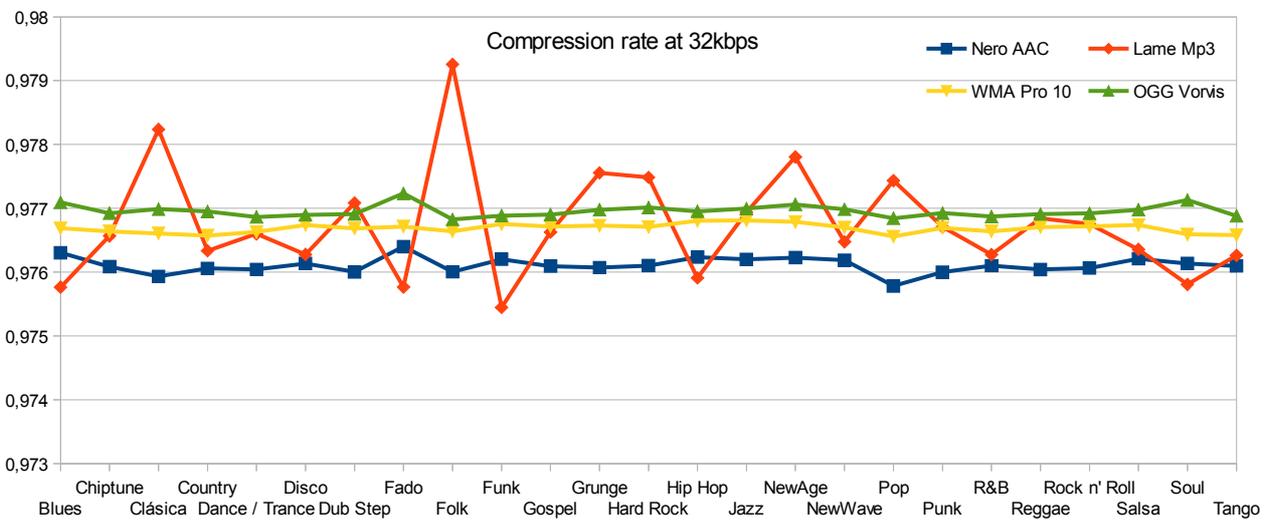


Figura 19: Factor de compresión a 32kbps

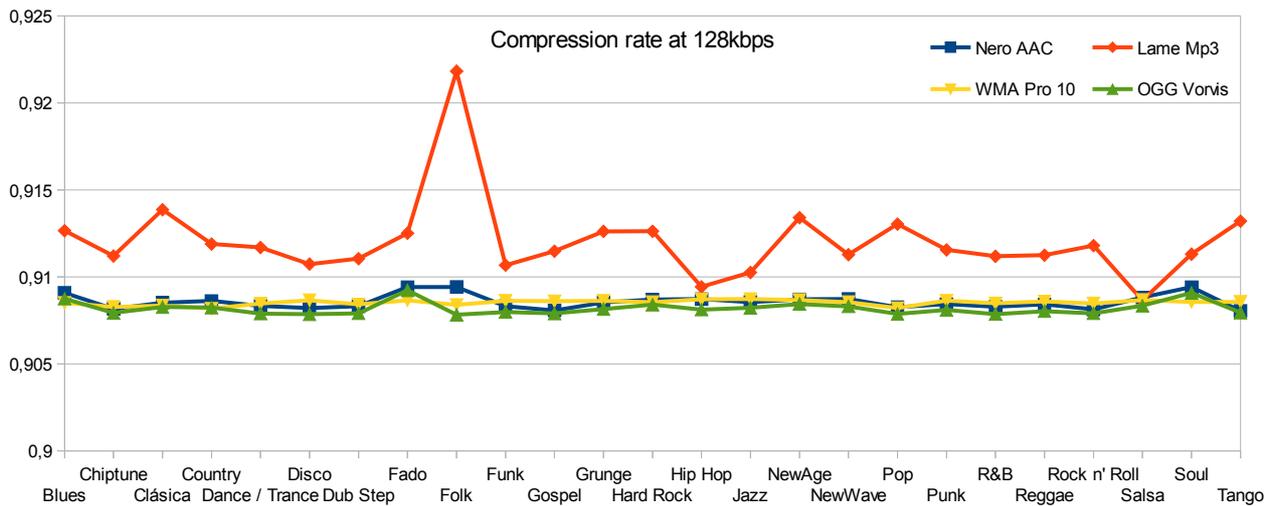


Figura 20: Factor de compresión a 128kbps

A la vista de las gráficas podemos concluir que Mp3 consigue una mayor compresión a medio- altos bitrates y que con un bitrate muy restringido igual es una mejor opción elegir otro códec dado que nos permitirá arañar un poco de ancho de banda, esto es muy importante para difusiones por la red. Aunque la diferencia sea mínima sigue siendo importante tanto si tenemos miles de clientes como en el cómputo mensual de tráfico de red, estas diferencias pueden ser importantes de cara al coste aplicable por tráfico de red.

Por supuesto a la vista de los tiempos de compresión podemos decir que todos los códecs son válidos para la compresión en tiempo real, puede que el códec OGG no sea válido para éste uso si la capacidad de procesamiento de la máquina a utilizar esté limitada, como podría ser un dispositivo móvil tipo arduino o raspberry pi.

Mirando ya hacia el entorno a nivel personal y no ya el empresarial tenemos ya un ranking de velocidad a la hora de comprimir, teniendo en cuenta que queremos una calidad aceptable de unos 128-192kbps, es el siguiente:

- 1º – WMA pro v.10
- 2º – Nero AAC
- 3º – Mp3 Lame
- 4º – OGG Vorbis

Respecto a los factores de compresión medios tenemos:

1º – Mp3 Lame (+0,5%)

2º – Nero AAC, WMA pro 10

3º – OGG (-0,1%)

Para decir cual de estos códecs sería el adecuado a nivel personal tendríamos que valorar el espacio disponible en disco, el tiempo que estaría dispuesto a utilizar comprimiendo y la compatibilidad de los reproductores portátiles.

Puesto que la diferencia de espacio no es tan relevante en el entorno personal, cada día disponemos de discos cada vez más grandes tanto en dispositivos móviles como en fijos, creemos que aquí la velocidad de compresión resulta mucho más relevante.

Tendiendo estos criterios en mente, nuestra elección sería, en primer lugar WMA siempre que dispongamos de el por usar Windows 7 y no un Linux o Mac. Si resulta no ser una opción por compatibilidad o falta de disposición recomendaríamos el uso de Nero AAC también muy rápido y con unos tamaños de salida muy comedidos.

Añadir que este estudio no es definitivo y debería ser completado por otros para así tener una caracterización mucho más precisa de cada uno de los codecs. Tal y como veremos en el siguiente apartado.

Para finalizar decir que los resultados no han sido los esperados, a nivel personal esperabamos que el estilo de música llamado chiptune, música de baja definición o de 8.bit como la que generaban las antiguas consolas, (Nes, MegaDrive, etc..) tuviese un factor de compresión mayor, pero no ha sido así. También esperamos que estas músicas se puedan comprimir a menor bitrate perdiendo menos calidad que otros tipos de música. También resultó ser una sorpresa ver esos picos de compresión en el Folk y Blues. Añadir que al final ha resultado ser un estudio que ha aportado datos interesantes y ha estado bien haber participado en el.

6 – Línea de estudio posterior.

Existe la necesidad de complementar éste estudio con un estudio de percepción sonora para evaluar el comportamiento de estos codecs en cada bitrate, porque existe la posibilidad que un cierto códec a 128kbps saque la calidad de otro a 160kbps.

Con ésta futura línea de investigación podremos caracterizar de una manera más precisa cada uno de los codecs más utilizados para, así, en un tercer proyecto implementar el algoritmo que permita una óptima selección del códec según el género.

Creemos que la investigación según el género musical es una vía interesante y que no ha sido estudiada en profundidad puesto que ,como hemos podido comprobar, los codecs no están optimizados por géneros. Y dado que existe una diferencia auditiva considerable, canciones llenas de transitorios otras mas fluidas, es posible que se puedan conseguir avances en esta dirección.

7 – Referencias / Bibliografía.

REFERENCIAS

- [1] S. Lei, L. Changbin, L. Bin. "Network utility maximization for triple-play services", Computer Communications, Volume 31, Issue 10, 25 June 2008, pp. 2257 - 2269
- [2] J. Mueller, T. Magedanz, J. Fiedler, 'NNodeTree: a scalable peer-to-peer live streaming overlay architecture for next-generation-networks', Network Protocols and Algorithms, (2009), Vol. 1, No. 2, pp.61–84.
- [3] Miguel Edo, Miguel Garcia, Carlos Turro and Jaime Lloret, IP Telephony development and performance over IEEE 802.11g WLAN, The Fifth International Conference on Networking and Services (ICNS 2009), Valencia (Spain), April 20-25, 2009.
- [4] Miguel Garcia, Diana Bri, Carlos Turró, Jaime Lloret, A User-Balanced System for IP Telephony in WLAN, The Second International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies (UBICOMM 2008), Valencia (Spain), September 29 - October 4, 2008
- [5] Ali Safa Sadiq, Kamalrulnizam Abu Bakar, Kayhan Zrar Ghafoor, Jaime Lloret, An Intelligent Vertical Handover Scheme for Audio and Video Streaming in Heterogeneous Vehicular Networks, Mobile Networks and Applications, 2014. DOI: 10.1007/s11036-013-0465-8
- [6] Denis Rosário, Rodrigo Costa, Helder Paraense, Kássio Machado, Eduardo Cerqueira, Torsten Braun, Zhongliang Zhao, A Hierarchical Multi
- [7] Free Audio Converter Website. At <http://www.freac.org/>
- [8] Lame Mp3. Available at: <http://lame.sourceforge.net>
- [9] Nero AAC. Available at: <http://www.nero.com/esp/company/about-nero/nero-aac-codec.php>
- [10] OGG Vorbis. Available at: <http://www.vorbis.com/>
- [11] WMA pro v10. Available at:
<http://msdn.microsoft.com/enus/library/windows/desktop/gg153556%28v=vs.85%29.aspx>

Bibliografía consultada.

Introducción:

http://es.wikipedia.org/wiki/Historia_de_la_m%C3%BAsica

http://es.wikipedia.org/wiki/Historia_del_libro

Enmascaramiento:

http://es.wikipedia.org/wiki/Umbral_de_enmascaramiento

http://www.lpi.tel.uva.es/~nacho/docencia/ing_ond_1/trabajos_06_07/io1/public_html/enascaramiento.htm

<http://www.eumus.edu.uy/docentes/maggiolo/acuapu/enm.html>

Umbral auditivo:

<http://www.elruido.com/portal/web/guest/umbrales-auditivos>

http://es.wikipedia.org/wiki/Umbral_de_audici%C3%B3n

<http://neurofisiologiagranada.com/peatc/peatc-umbralesubjetivo.htm>

Compresión con pérdidas:

http://es.wikipedia.org/wiki/Algoritmo_de_compresi%C3%B3n_con_p%C3%A9rdida

<http://computer.howstuffworks.com/mp3.htm>

Códecs:

AAC

http://wiki.hydrogenaudio.org/index.php?title=Nero_AAC

MP3

<http://lame.sourceforge.net/>

OGG

<http://es.wikipedia.org/wiki/Vorbis>

WMA

http://msdn.microsoft.com/en-us/library/windows/desktop/gg153556%28v=vs.85%29.aspx#windows_media_audio_10_professional