

DESEMPEÑO DE CODIFICADORES DE AUDIO SIN PÉRDIDAS Y UNA HERRAMIENTA QUE SIMPLIFICA SU ANÁLISIS

Marengo Rodriguez, F. A.; Roveri, E. A.; Rodríguez Guerrero, J. M.;
Treffiló, M.; Miyara, F.

Laboratorio de Acústica y Electroacústica, Escuela de Ingeniería Electrónica, Facultad de Ciencias Exactas, Ingeniería y Agrimensura, Universidad Nacional de Rosario, Rosario.
Correo electrónico: fmarengorodriguez@yahoo.com.ar

Resumen – A raíz de la mayor capacidad de transmisión y recursos de almacenamiento disponibles, muchos usuarios de equipos reproductores de audio se han volcado por la codificación de audio sin pérdidas en lugar de la codificación perceptual (es decir, que producen pérdida de información no percibidas por el oído común como MP3). Hasta el momento, se hicieron muchas comparaciones entre los codec sin pérdidas, pero ninguna de ellas nos da noción de cuál codec es el mejor para comprimir un archivo de audio dado en formato PCM estéreo. Este trabajo pretende cubrir esa necesidad, haciendo una comparación exhaustiva entre varios codec populares en términos del factor de compresión y de la velocidad de codificación para piezas musicales de distintos géneros provenientes de CD comerciales. Los géneros analizados son folklore, tango, música clásica, rock y jazz. También se incluyó como categoría adicional recitales de rock en vivo, a fin de evaluar la influencia del ruido ambiente sobre el desempeño de la codificación. En este trabajo se demuestra que la compresión alcanzada depende esencialmente del género musical, y que cada codificador se caracteriza por su velocidad. Esto se debe a una base tecnológica común a todos los codec analizados, compuesta por predictores y bloques codificadores por entropía. También se muestra que cada codec se caracteriza por una zona operativa en el dominio velocidad-compresión. Finalmente, se plantea el diseño de una herramienta gráfica a fin de exponer resumidamente los mejores codec en términos de velocidad y compresión.

1. INTRODUCCIÓN

La compresión de información ha jugado un papel fundamental en la tecnológica digital. Para comprimir audio en particular, se diseñaron los codificadores de audio o codec, los cuales explotan la relativa predecibilidad a corto plazo de la señal tratada (*codecs sin pérdidas* o CSP) e incluso pueden suprimir parte de la misma de forma que el resultado no varía apreciablemente para el oyente (codecs perceptuales como MP3).

A pesar que los CSP logran comprimir menos que los codecs perceptuales (1,5 a 6 veces en lugar de 8 a 13 veces), su uso se difunde cada vez más por varias razones. Por ejemplo, el avance de la tecnología permite acceder más fácilmente a recursos de almacenamiento con mayor capacidad de memoria y a canales de transmisión de datos más potentes que achican la brecha entre los CSP y los codec perceptuales. Asimismo, la distorsión introducida por éstos se evidencia en los equipos de reproducción de sonido de alta fidelidad

que bajan de costo continuamente. Además, no se permite perder información cuando se transmite datos entre equipos de masterización en estudios de sonido.

Hasta el momento, no hay ninguna guía que oriente al usuario común sobre las ventajas de cada codec, en términos de velocidad y compresión para distintos géneros musicales. En este trabajo se analiza un conjunto de CSP con diversas piezas musicales completas provenientes de CD comerciales, donde las muestras digitales de 16 bits fueron extraídas a 44 100 Hz.

2. FUNCIONAMIENTO DE LOS CODEC SIN PÉRDIDAS

Básicamente, los codificadores sin pérdidas procesan la señal en cuadros de longitud determinada, y en cada uno de ellos se realiza la predicción de las muestras de entrada, obteniéndose un conjunto de datos pequeños llamado residuo, que posteriormente se comprime con un codificador de entropía [1]. Éstos se multiplexan con los parámetros del predictor para conformar la trama de salida del codificador.

Para la predicción que realizan los CSP aquí analizados, cada muestra de entrada se expresa como combinación lineal de entradas y salidas pasadas. La mayoría de los CSP ignoran estas últimas.

Utilizamos los codecs enumerados en [2] seleccionando los modos por defecto (normal) y los que maximizan su velocidad (rápido) y compresión (óptimo).

3. ANÁLISIS PROPUESTO

Se eligió más de una docena de piezas musicales de: folklore, tango, jazz, música clásica, pop, rock y recitales de rock en vivo. Esta última opción se tuvo en cuenta para analizar la influencia del ruido ambiente. En cada género, se contempló un gran abanico de artistas y ritmos, y se calculó el valor promedio del factor de compresión FC así como del rate R . El primero se define como el cociente entre los tamaños del archivo original y codificado. El rate es el cociente entre el tiempo de reproducción del archivo original y el tiempo de codificación.

La ejecución de los CSP se llevó a cabo empleando una herramienta gráfica propia que se adapta a nuestros requerimientos de ensayo [3], ya que podemos conocer su retardo asociado (0,38 s), pausar la ejecución del proceso para detectar la presencia de errores y presentar los resultados de interés en forma tabulada.

4. RESULTADOS

Se estudió el desempeño de los CSP en una PC portátil equipada con placa madre HP Pavilion dv7, procesador AMD Turion II Ultra Dual-Core de 2,4GHz y memoria SDRAM DDR2-800 de 2 x 2GB. Funcionó con sistema operativo Windows Vista 6.1.7600.

4.1 Codificación en modo normal

El consumo de CPU estuvo comprendido entre el 24 % y el 28 %. Al agrupar la información por género, se obtuvieron los valores de FC expuestos en la Fig. 1 (a). En la parte inferior de este panel, se agregó el promedio de FC para cada género. Cabe aclarar que se ordenó la información en orden decreciente de compresión, siendo más favorable para el jazz y menos para el pop.

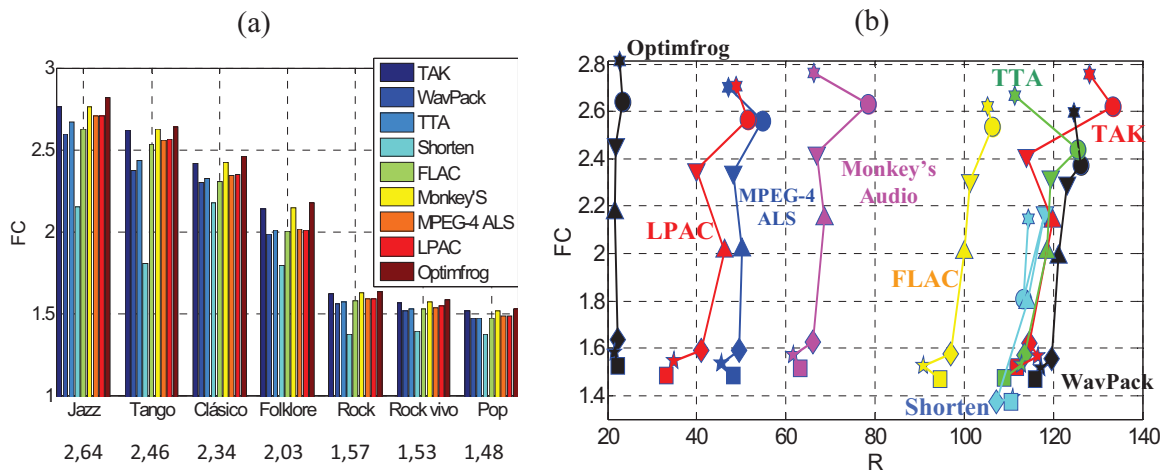


Figura 1: (a) Compresión obtenida con los codificadores analizados en cada género musical. (b) Zona operativa de cada codec para jazz (★), tango (●), música clásica (▼), folklore (▲), rock (◆), rock en vivo (★) y pop (■).

La Fig. 1 (b) ilustra las zonas operativas de cada codec en el dominio $R-FC$. Por un lado, se puede caracterizar cada CSP por su velocidad, siendo TAK, WavPack, TTA y Shorten los más veloces y Optimfrog el más lento. Por otro lado, la velocidad en cada codec se ve afectada por el género que procesa. Por ejemplo, lleva menos tiempo comprimir tango y más tiempo comprimir música clásica, rock en vivo y pop.

5. MÉTODO DE FRONTERA

Dada la gran cantidad de configuraciones codecs-modos ó C-M, el análisis anterior se vuelve muy engorroso. Con fines de simplificación, se propone una herramienta gráfica orientada al usuario, simple, aproximada y que elimine los C-M más lentos y de menor compresión. El resultado es un conjunto C-M óptimos en el dominio $R-FC$.

La Fig. 2 plantea dos casos posibles de C-M llamados “A” y “B”. En el caso (i), el codec

A es más lento pero comprime más que el codec B: ambos codec representan opciones óptimas. Por otro lado, en (ii) se descarta el codec A por ser más lento y de menor compresión que B. Al contemplar todos los C-M ordenados según valores decrecientes de R , se obtiene el conjunto completo de C-M óptimos que describen una curva monótona decreciente y constituyen la *frontera* del máximo rendimiento.

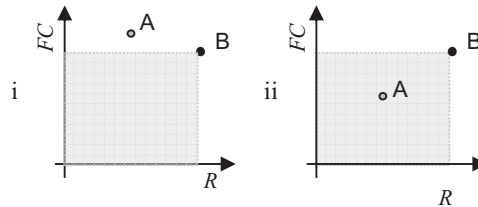


Figura 2: Dos situaciones posibles de C-M: en (i) “A” y “B” son óptimos, en (ii) se descarta “A”.

Cabe agregar que en este proceso conviene trabajar con valores normalizados de rate y compresión, dado que sólo interesa la comparación mutua entre los datos originales. Dicha normalización se lleva a cabo dividiendo los valores $\{R, FC\}$ de entrada entre los correspondientes a Optimfrog-óptimo. Éste es un buen parámetro, no solo porque corresponde a la mejor compresión, sino también porque al ser su rate bajo, la dispersión provocada por el retardo del sistema operativo y la interfaz gráfica se minimizan.

5.1 Aplicación

Para obtener la frontera del conjunto de todos los géneros musicales, se utilizaron los datos provenientes de 2 archivos representativos en cada género, es decir con valores de R y de FC cercanos a los promedios respectivos. Luego se promediaron los resultados en cada C-M ponderando las piezas según su tiempo de reproducción (ver círculos amarillos en la Fig. 3 (a)), por ser éste el que inmuniza los efectos de dispersión del rate. Luego, se aplicó al algoritmo de selección de puntos óptimos a los mencionados centroides resultando la frontera ilustrada en la Fig. 3 (b).

6. CONCLUSIONES

Se observó que la compresión alcanzada es más dependiente del género musical tratado que del codec utilizado, debido a que éstos poseen la misma la estructura básica. Se mostró que las piezas de jazz y tango pueden reducir su espacio de memoria inicial en un porcentaje mayor a 60. Por el contrario, las piezas de pop y rock se reducen un poco más de 35 %. Se observó también que el ruido contenido en las grabaciones en vivo disminuyó la compresión en aproximadamente un 10 %. Postulamos que estos resultados se atribuyen a las

amplitudes presentes en la señal original, ya que las mismas suelen ser mayores en los géneros de rock y pop. En términos de velocidad, se pudo caracterizar cada codec por una zona operativa.

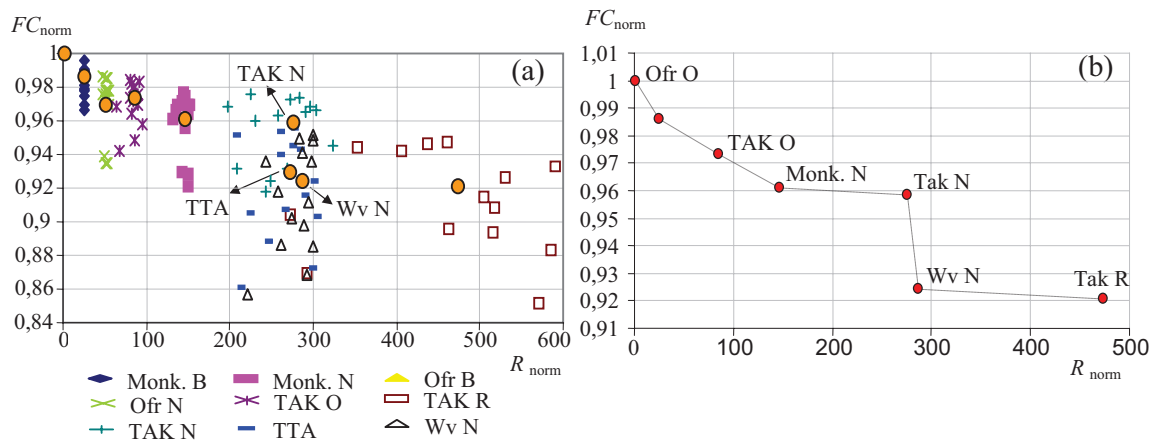


Figura 3: (a) Representación de valores de R y FC normalizados asociados a los 14 archivos analizados en los C-M seleccionados. (b) Frontera obtenida.

Se desarrolló para el usuario común una herramienta gráfica sencilla y de fácil utilización, diseñada con un algoritmo de selección de puntos C-M óptimos en el dominio velocidad-compresión. Esta herramienta sirve de guía para optar por el par C-M más conveniente según las necesidades de compresión y velocidad, y su uso se puede extender a otros codificadores y géneros musicales, adaptándose a los requerimientos de consumo del usuario.

7. REFERENCIAS

- [1] Salomon, D. “*Data Compression - The Complete Reference*”. 4ta ed. Springer-Verlag. Reino Unido. 2007.
- [2] Marengo Rodríguez, F.A.; Roveri, E.A.; Rodríguez Guerrero, J.M.; Treffiló, M.; Miyara, F. “*Análisis comparativo de codificadores de audio sin pérdidas*”, 2do Congr. Intern. de Acúst. UNTREF, Bs. As., Argentina. 2010.
- [3] Rodríguez Guerrero, J.M. “*Herramienta gráfica para ensayo de codificadores de audio sin pérdidas*”. Disp. en <http://www.fceia.unr.edu.ar/acustica/>
- [4] Robinson, T. “*Shorten: Simple lossless and near-lossless waveform compression*”. Rep. téc. CUED/F-INFENG/TR.156. Univ. Cambridge. Reino Unido. 1994.
- [5] Coalson, J. “*FLAC - Free lossless audio codec*” Codec libre de cód. abierto. <http://flac.sourceforge.net/>
- [6] Ashland, M.T. “*Monkey's Audio*”. Codec de audio. <http://www.monkeysaudio.com/index.html>
- [7] Liebchen, T. “*MPEG-4 ALS. The standard for lossless audio coding*”. Journ. Acoust. Soc. Korea, vol. 28, pp. 618–629. 2009.
- [8] Liebchen, T. “*Lossless predictive audio coder (LPAC)*”. Ejecutable en http://www.nue.tu-berlin.de/menue/mitarbeiter/tilman_liebchen/lpac_-_lossless_audio_codec_for_windows_and_linux/
- [9] Ghido, F. “*Ghido's Data Compression Page*”. Codec de audio. <http://www.losslessaudio.org/>
- [10] Becker, T. “*Tom's audio compressor (TAK)*”. <http://thbeck.de/Tak/Tak.html>
- [11] Djuric, A. “*TTA Lossless audio codec - True audio compressor algorithms*”. http://en.true-audio.com/TTA_Lossless_Audio_Codec_-_True_Audio_Compressor_Algorithms
- [12] Bryant, D. “*WavPack. Hybrid Lossless Audio Compression*” <http://www.wavpack.com/>