



*2° Congreso Internacional de Acústica UNTREF
Septiembre 2010, Argentina*

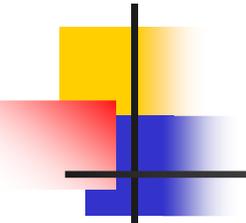
ANÁLISIS COMPARATIVO DE CODIFICADORES DE AUDIO SIN PÉRDIDAS

FERNANDO A. MARENGO RODRIGUEZ, ERIBERTO A. ROVERI,
JUAN MANUEL RODRÍGUEZ GUERRERO, MAURO TREFFILÓ,
FEDERICO MIYARA

Laboratorio de Acústica y Electroacústica
Escuela de Ingeniería Electrónica
Facultad de Ciencias Exactas, Ingeniería y Agrimensura
Universidad Nacional de Rosario
Rosario, Argentina

<http://www.fceia.unr.edu.ar/acustica/>





Motivación del trabajo

- Entre los audiófilos, los codificadores de audio sin pérdidas (CSP) han ganado popularidad frente a los perceptuales.
- No se registra antecedente sobre la comparación del desempeño de diferentes CSP para archivos de audio provenientes de CD.

Tipos de codec

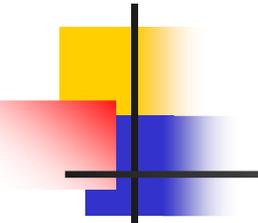
Perceptuales

Sin pérdidas (CSP)

Minimizan redundancia en la señal de entrada.

- Filtran irrelevancias psicoacústicas.
- Pros: gran compresión → 8 a 13 veces.
- Contra: Suprimen información.
- Ejemplos: MPEG-1 audio layer 3 (MP3) y Ogg Vorbis (formato libre y abierto).

- Pros: No introducen distorsión.
- Contra: menor compresión ~ 1,5 a 6 veces.
- Ejemplos: FLAC, WavPack, Monkey's Audio, OptimFrog, LPAC, MPEG-4 ALS, Shorten, TAK, TTA.



¿Por qué razones los CSP son cada vez más utilizados?

- **Tecnológicas:**

- Recursos de almacenamiento con mayor capacidad de memoria.
- La creciente velocidad de transmisión de datos vía Internet minimiza la ventaja de los codec perceptuales frente a los CSP.

- **Económicas:**

- En los equipos de reproducción de sonido de alta fidelidad se vuelve evidente la distorsión que introduce la codificación perceptual. Dichos equipos disminuyen de costo y su acceso es más masivo.

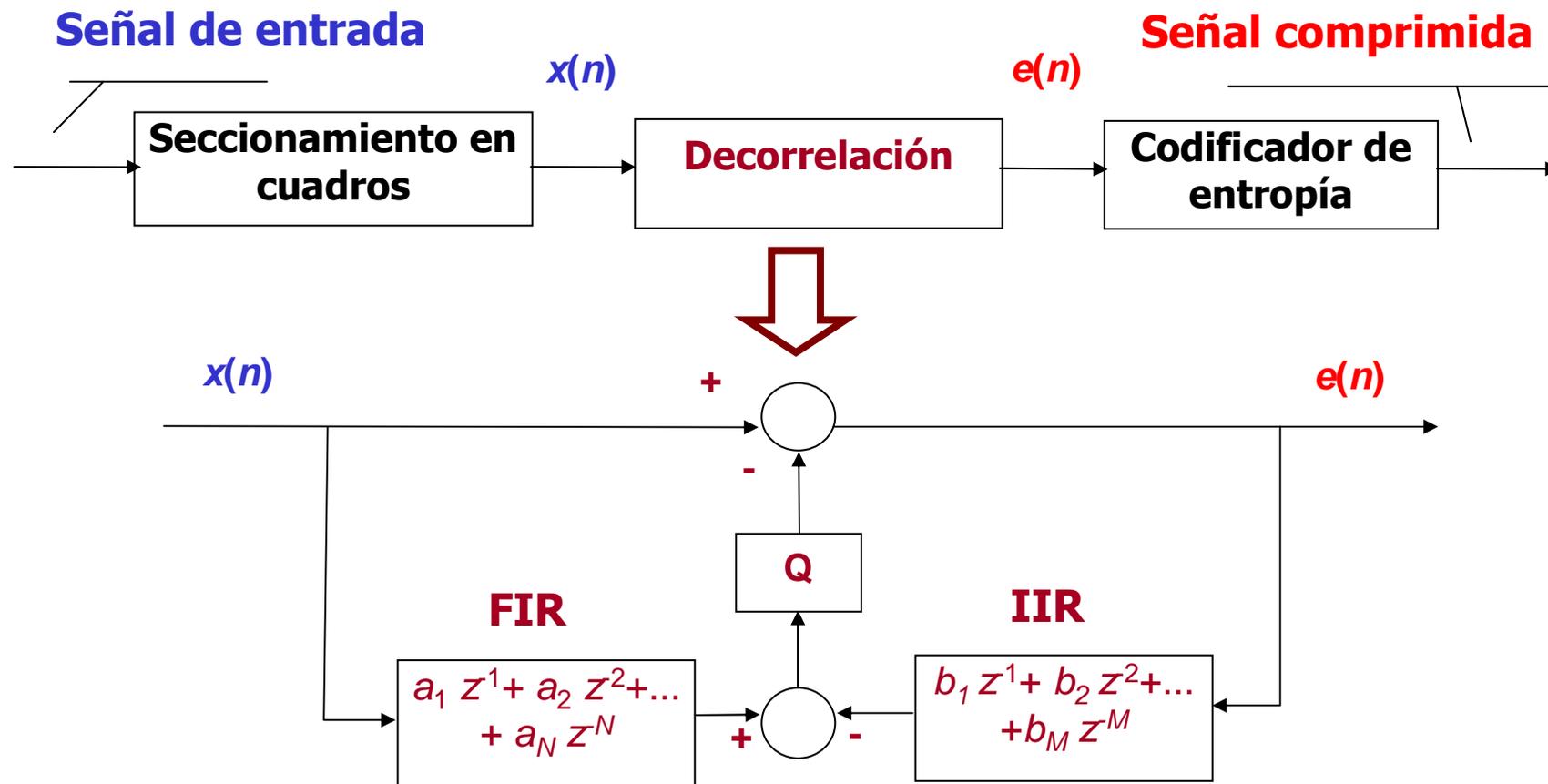
- **Masterización:**

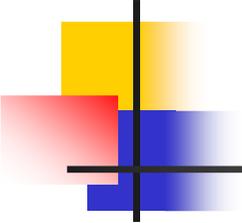
- Las grabaciones de estudio se pueden transmitir entre equipos remotos en formato comprimido sin distorsión: la única opción viable es usar CSP.

- **Medición:**

- Las grabaciones de audio que realizan los especialistas en acústica suelen ocupar espacio considerable en formato WAV. Se admite como alternativa el uso de los CSP ya que no distorsionan la información y economizan el espacio de memoria ocupado.

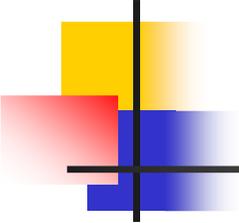
Funcionamiento de los CSP





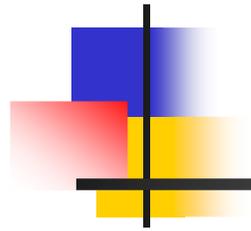
Análisis de rendimiento

- Se analizaron los CSP con estos géneros:
 - Folklore, Tango, Jazz, Música clásica, Pop, Rock
 - Recitales de rock en vivo (influencia del ruido ambiente).
 - Piezas musicales: más de 12 por género.
- Cada codec funcionó en los modos normal (por defecto), rápido y óptimo.
- Se ensayaron los codec en diferentes PC.
- Parámetros de rendimiento:
 - Factor de compresión $FC = \text{input_file_size} / \text{output_file_size}$
 - Rate $R = \text{file_playtime} / (\text{de})\text{coding_time}$.



Hardware ensayado

- PC portátil, AMD Turion II Ultra Dual-Core M600 de 2,4 GHz, 4GB de memoria DDR2 y Windows Vista.
- PC de escritorio, Pentium IV HT de 3GHz, 1GB de memoria DDR y Windows XP.
- PC de escritorio, AMD Athlon 64 x2 Dual Core 2,2 Ghz, 2GB de memoria DDR2 y Windows XP.



Interfaz gráfica



CompresorDescompresor



Dir. Origen: D:\Origen

Seleccionar

Compresor: D:\Codecs\mp4als.exe

Seleccionar

Dir. Salida: D:\Destino

Seleccionar

Formato de salida: als

Parametros: -v %infile %outfile

"%infile": Archivo de entrada "%outfile": Archivo de salida

Del dir. salida?

Comprime

Descomprime

Pausa?

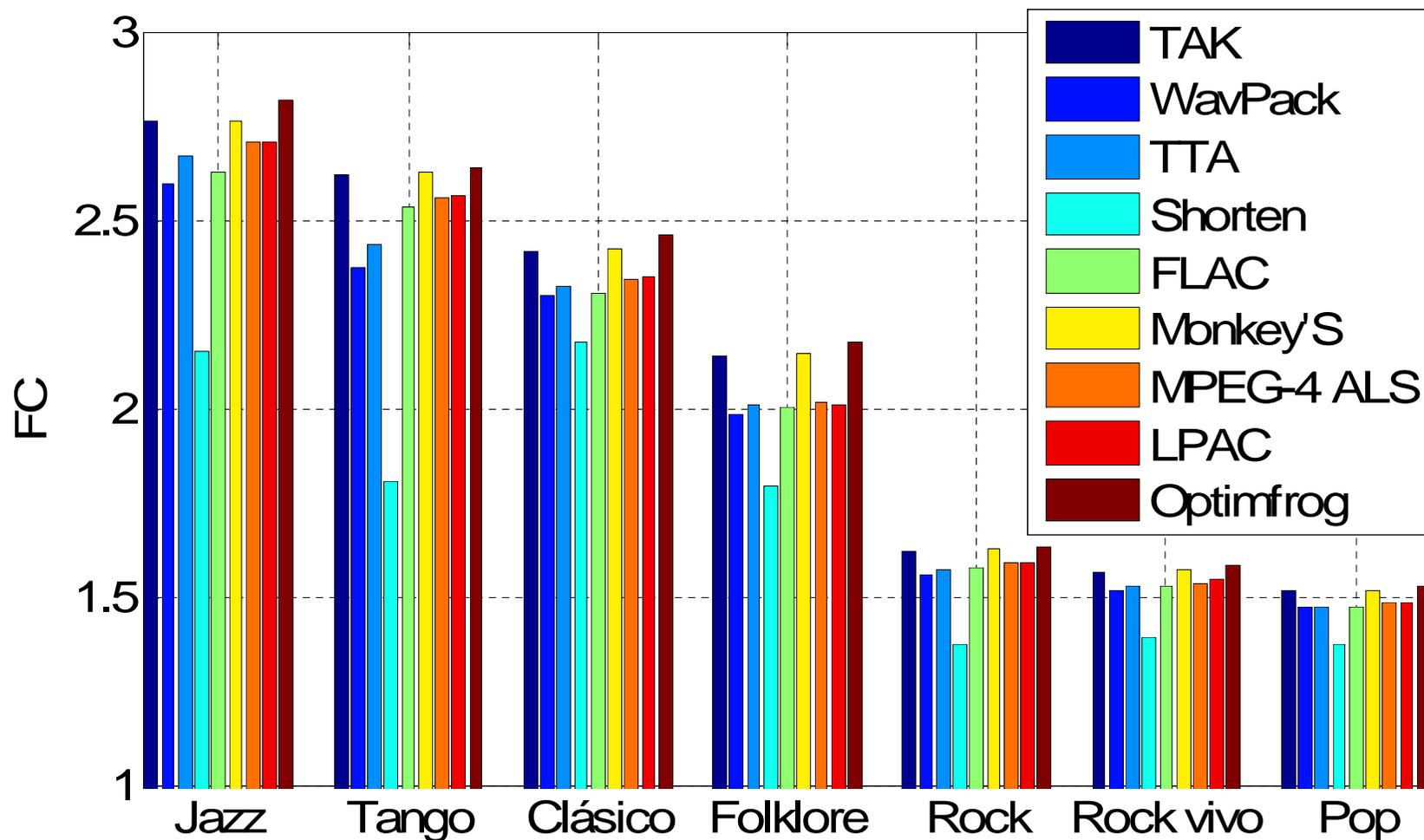
	Nombre	Tamaño - bytes	Extensión	Tiempo
▶	Archivo01.als	33544432	.als	10,294
	Archivo02.als	20539935	.als	8,115
	Archivo03.als	12954231	.als	5,537
	Archivo04.als	12362048	.als	6,182

Run!

Stop!

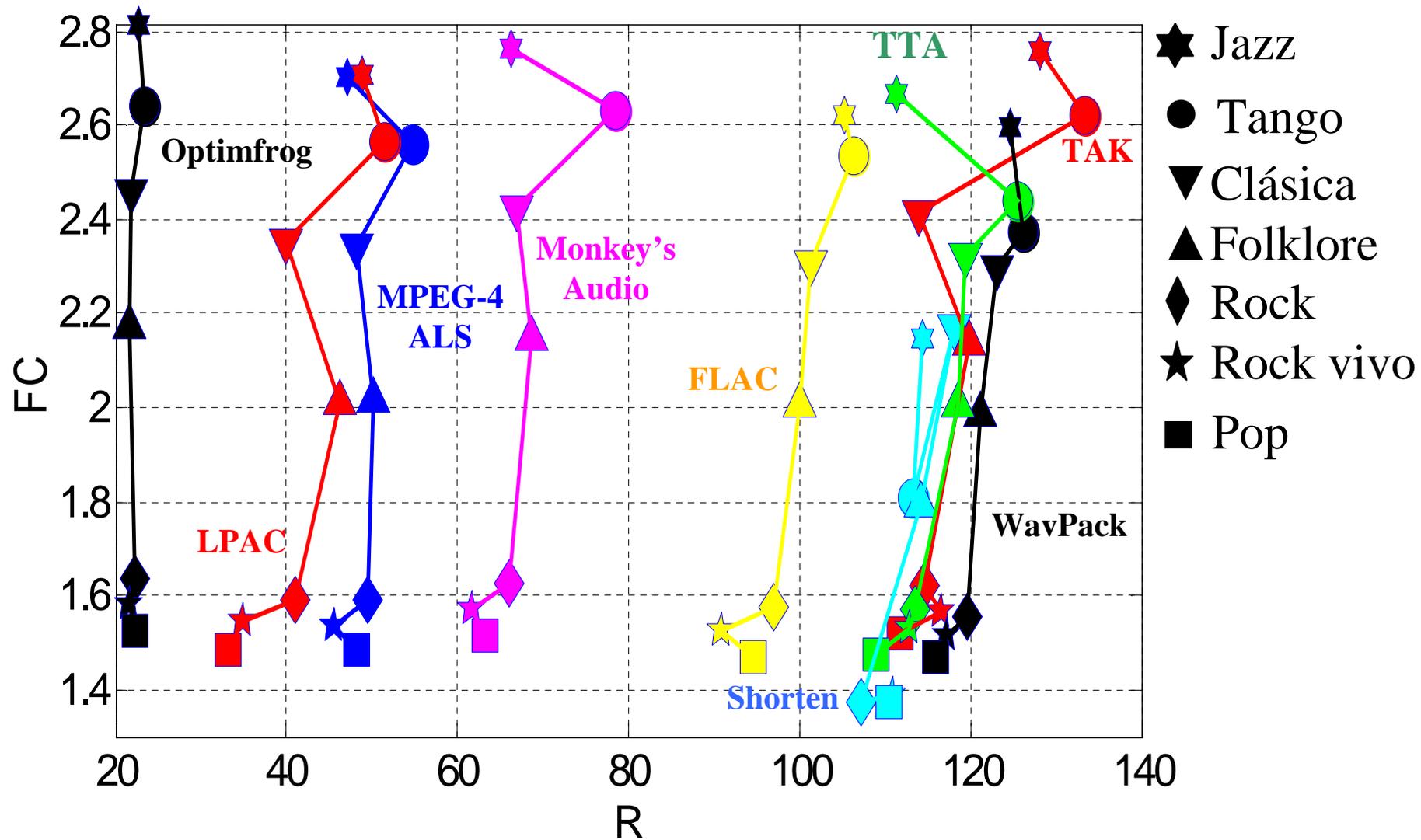
Test!

Compresión por género y codec



Media	2,64	2,46	2,34	2,03	1,57	1,53	1,48
-------	------	------	------	------	------	------	------

FC vs Rate por codec y género



Cambios porcentuales de FC

“Música para los Reales Fuegos Artificiales – Allegro” de Handel

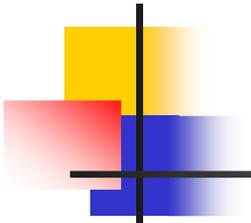
En modo normal, FC~2

	Óptimo	Rápido
TAK	4,16	-1,96
WavPack	2	-1,92
TTA	-	-
Shorten	3,57	-
FLAC	0,85	-5,5
Monkey's Audio	4,16	-1,96
MPEG-4 ALS	8,51	-
LPAC	1	-8
OptimFrog	6,52	-2

Rate codificación/decodificación

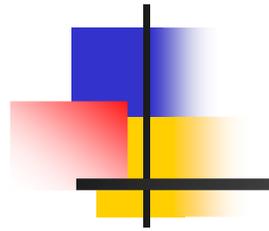
“Música para los Reales Fuegos Artificiales – Allegro” de Handel

	Óptimo	Rápido
TAK	36 / 88	120 / 113
WavPack	3 / 91	150 / 114
TTA	116 / 111 (único modo)	
Shorten	89 / 116	122 / 146 (modo normal)
FLAC	33 / 196	90 / 211
Monkey's Audio	11 / 11	85 / 66
MPEG-4 ALS	1 / 11	28 / 58 (modo normal)
LPAC	20 / 48	85 / 50
OptimFrog	0,44 / 3	31 / 44

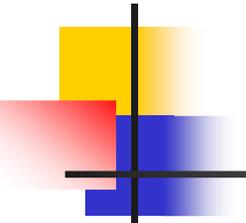


Comparación del Rate en todos los géneros musicales

	Óptimo	Normal	Rápido
TAK	36,76	119,77	205,58
WavPack	2,59	121,12	142,52
TTA	-	115,72	-
Shorten	82,92	112,63	-
FLAC	31,59	99,37	140,84
Monkey's Audio	11,11	67,33	84,16
MPEG-4 ALS	1,37	49,09	-
LPAC	23,90	42,21	-
OptimFrog	0,44	22,25	29,43



Método de Frontera



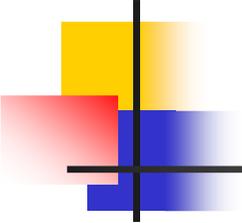
¿Por qué proponer un nuevo tipo de análisis?

El anterior estudio es:

- Engorroso
- No permite comparar resultados directamente
- No considera la posibilidad de simplificar

Queremos:

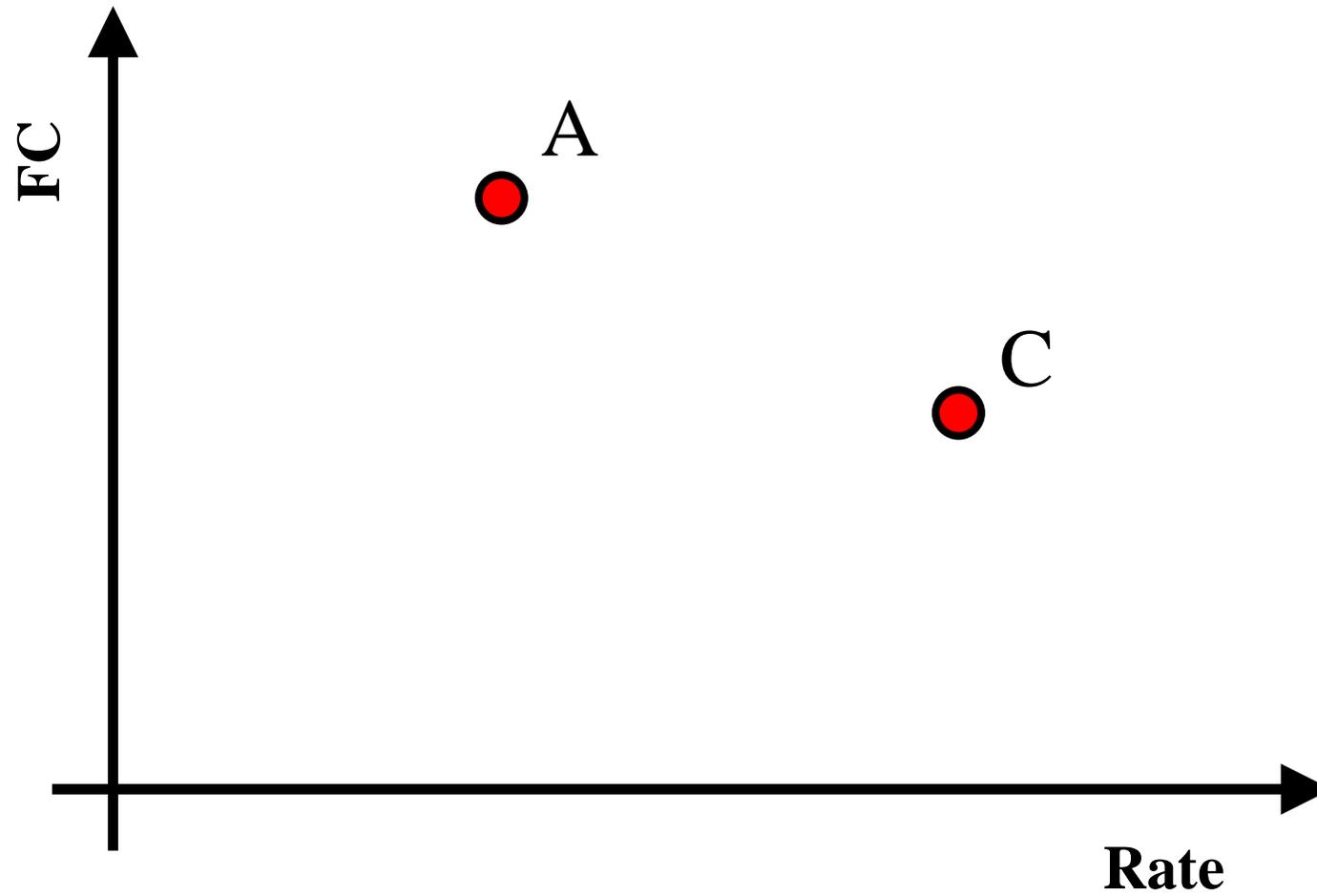
- Algo orientado al usuario
- Que permita buscar una solución de compromiso de forma simple e intuitiva
- Que permita comparar resultados independientemente del archivo y hardware ensayado
- No tiene por qué ser exacto

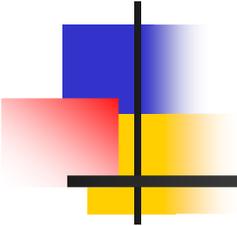


Puntos a tener en cuenta

- Se incluyen todos los codec y modos (C-M) en el análisis.
- La selección de temas está compuesta por dos temas representativos de cada género.
- Se requiere de algún tipo de normalización para poder comparar resultados.
- Es necesario establecer un criterio de simplificación para descartar C-M.

Ejemplo de C-M eliminable

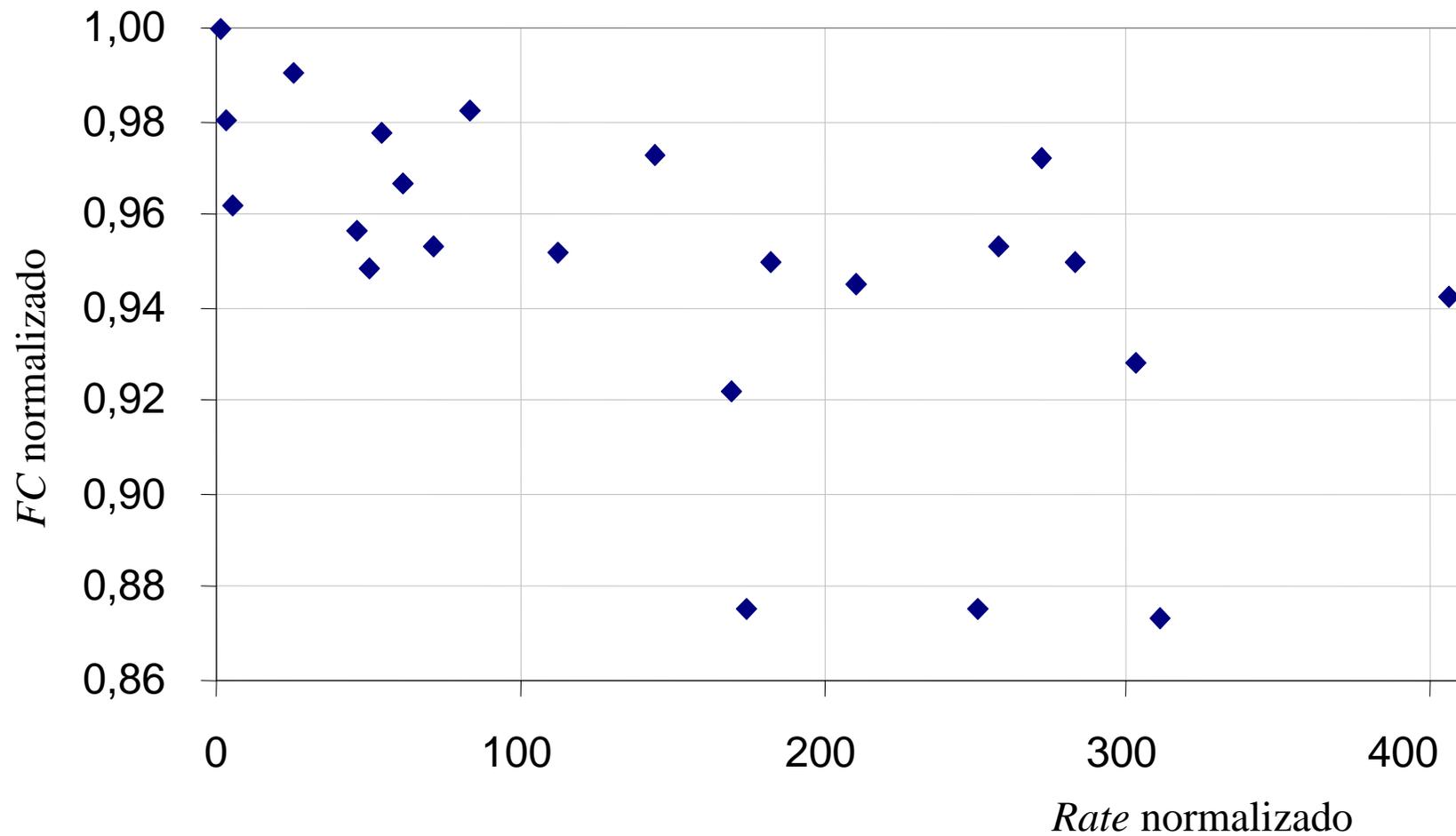




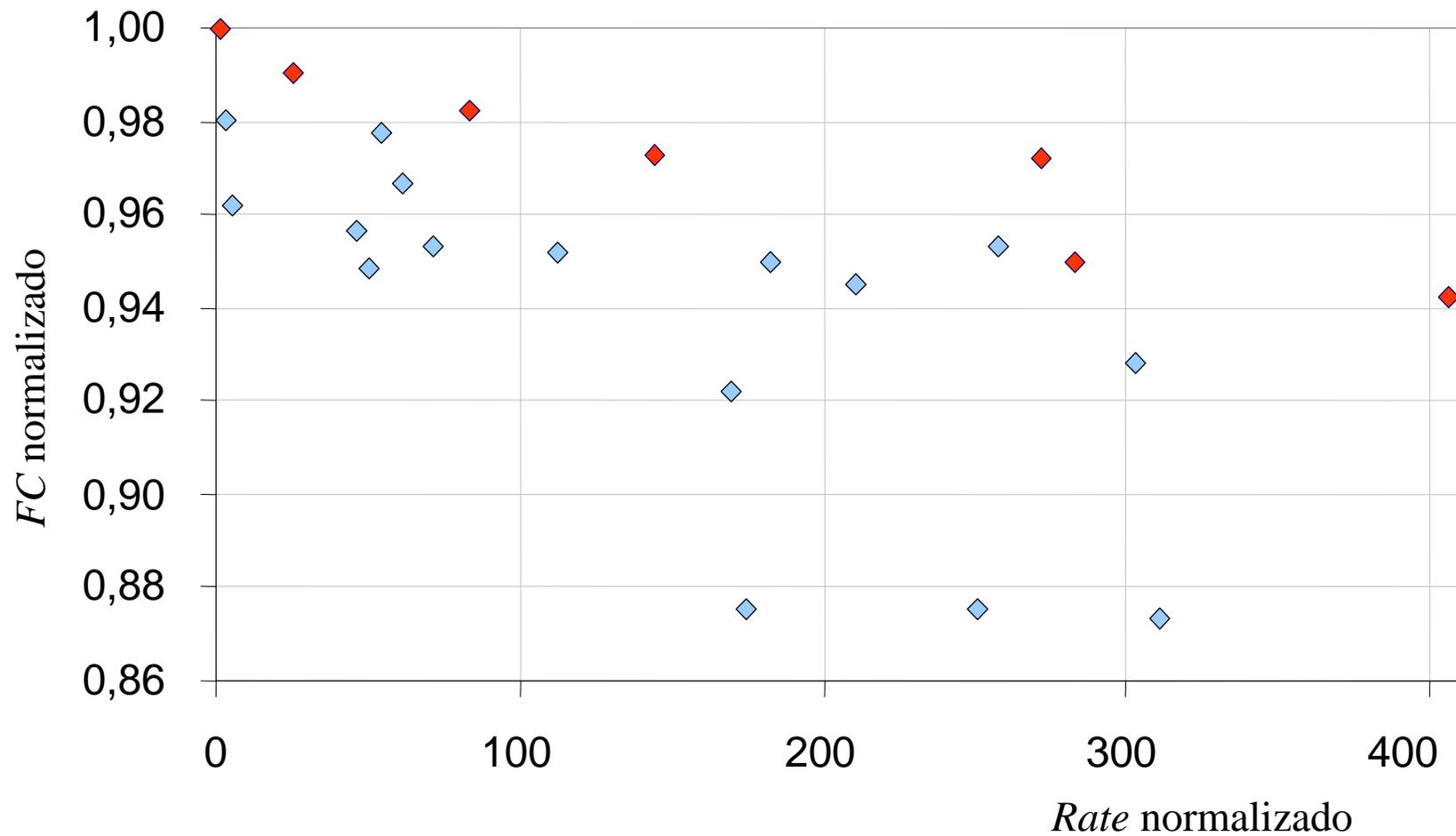
Aplicación del método con un solo archivo

Pieza "Sledgehammer" de Peter
Gabriel

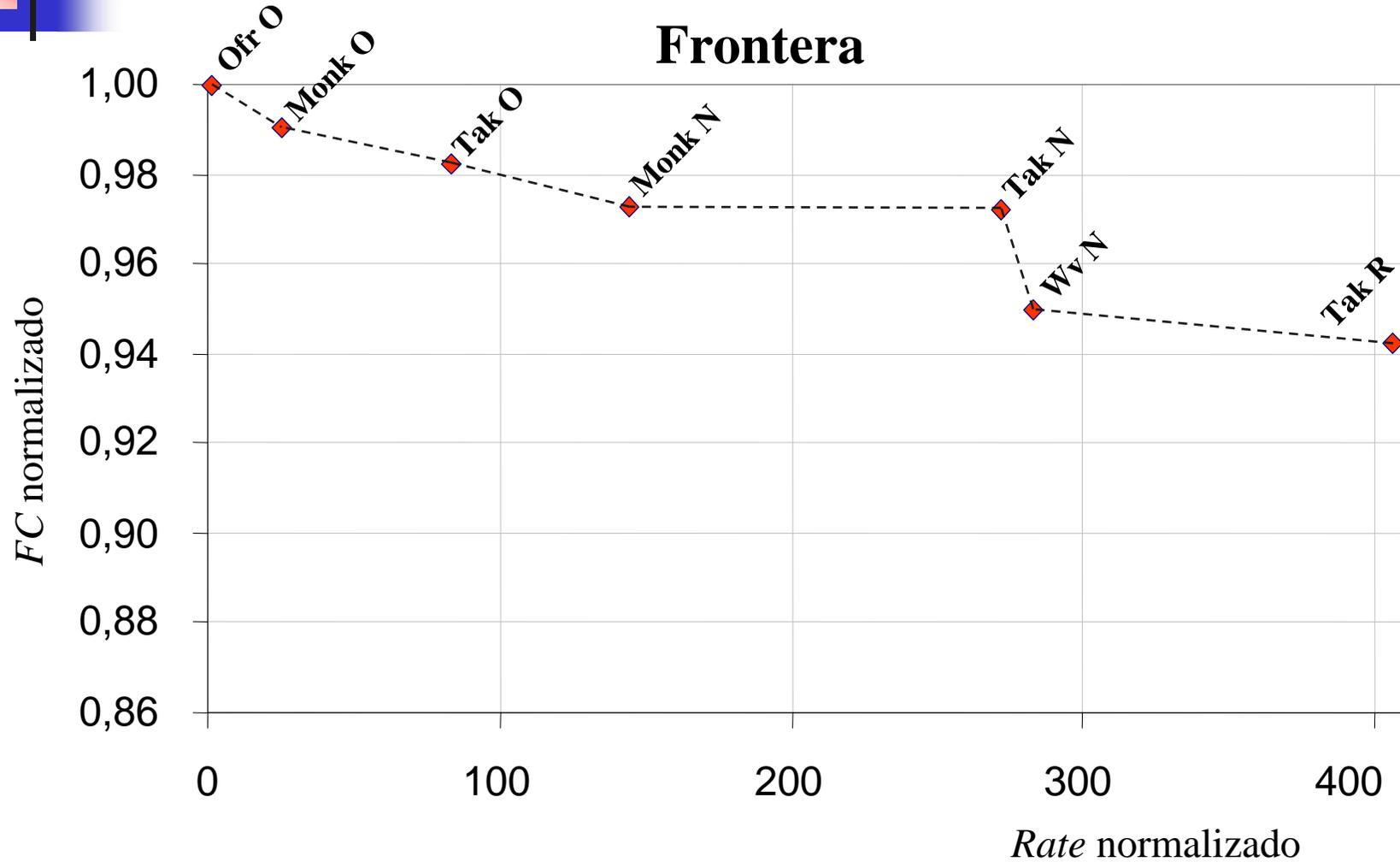
Gráfica de resultados normalizados

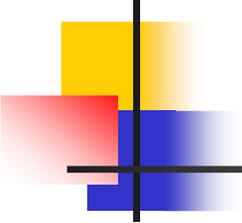


Gráfica de resultados normalizados



Gráfica de resultados normalizados





Extensión a múltiples archivos

- Frecuencia de aparición en la Frontera
 - Selección de los C-M en base a la cantidad de archivos en donde aparecen en la frontera.
 - Los resultados son relativamente inmunes a la selección de temas, pero si la dispersión es alta, el método puede fallar.
- Promedios ponderados
 - Aplica por cada C-M y antes de normalizar, un promedio ponderado por el tamaño del archivo, sobre todos los archivos.
 - Permite realizar estimaciones del tiempo de codificación, ya que representa un caso "promedio". Elimina el problema visual de superposición de nubes.

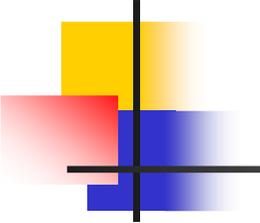
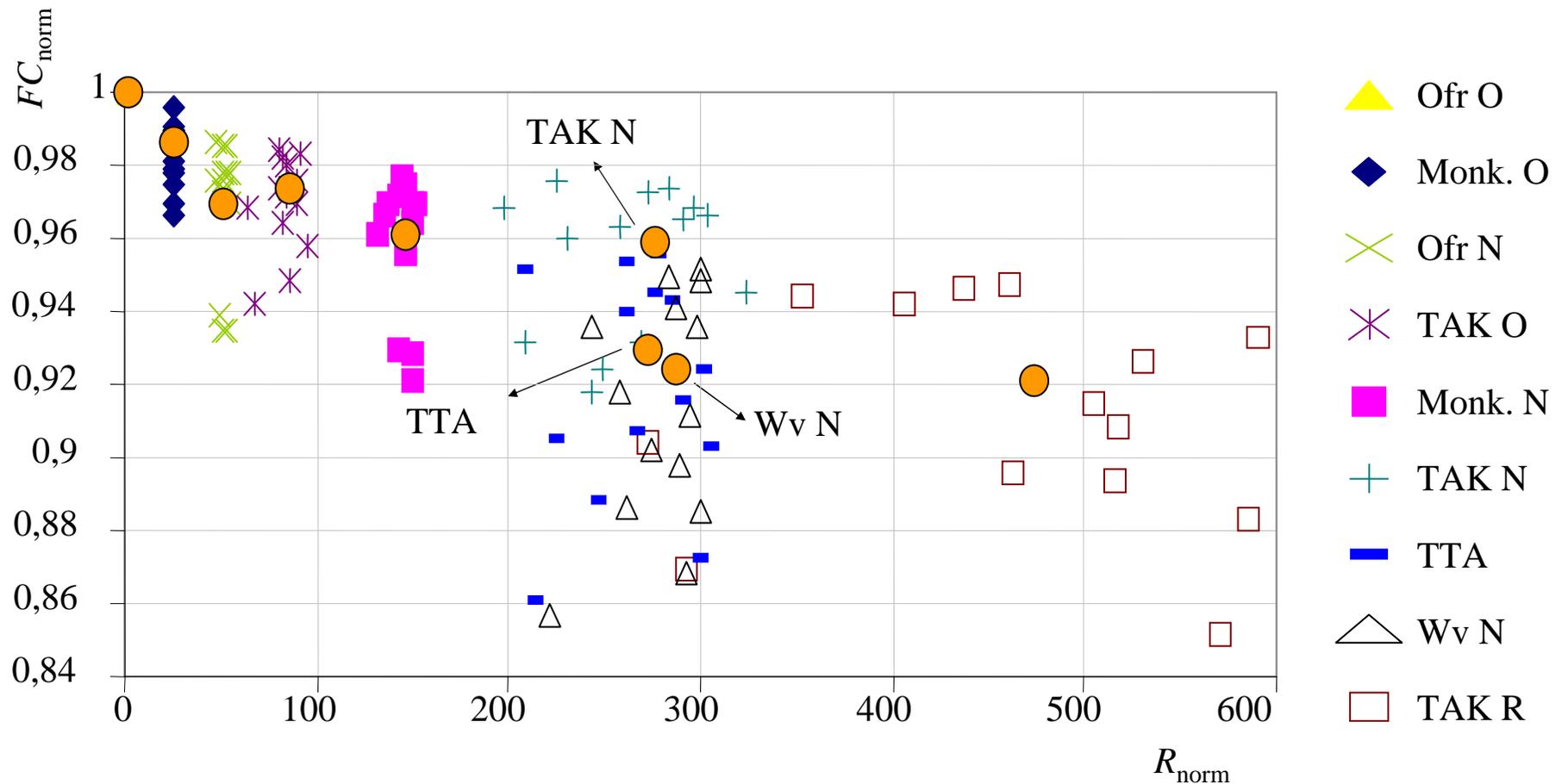


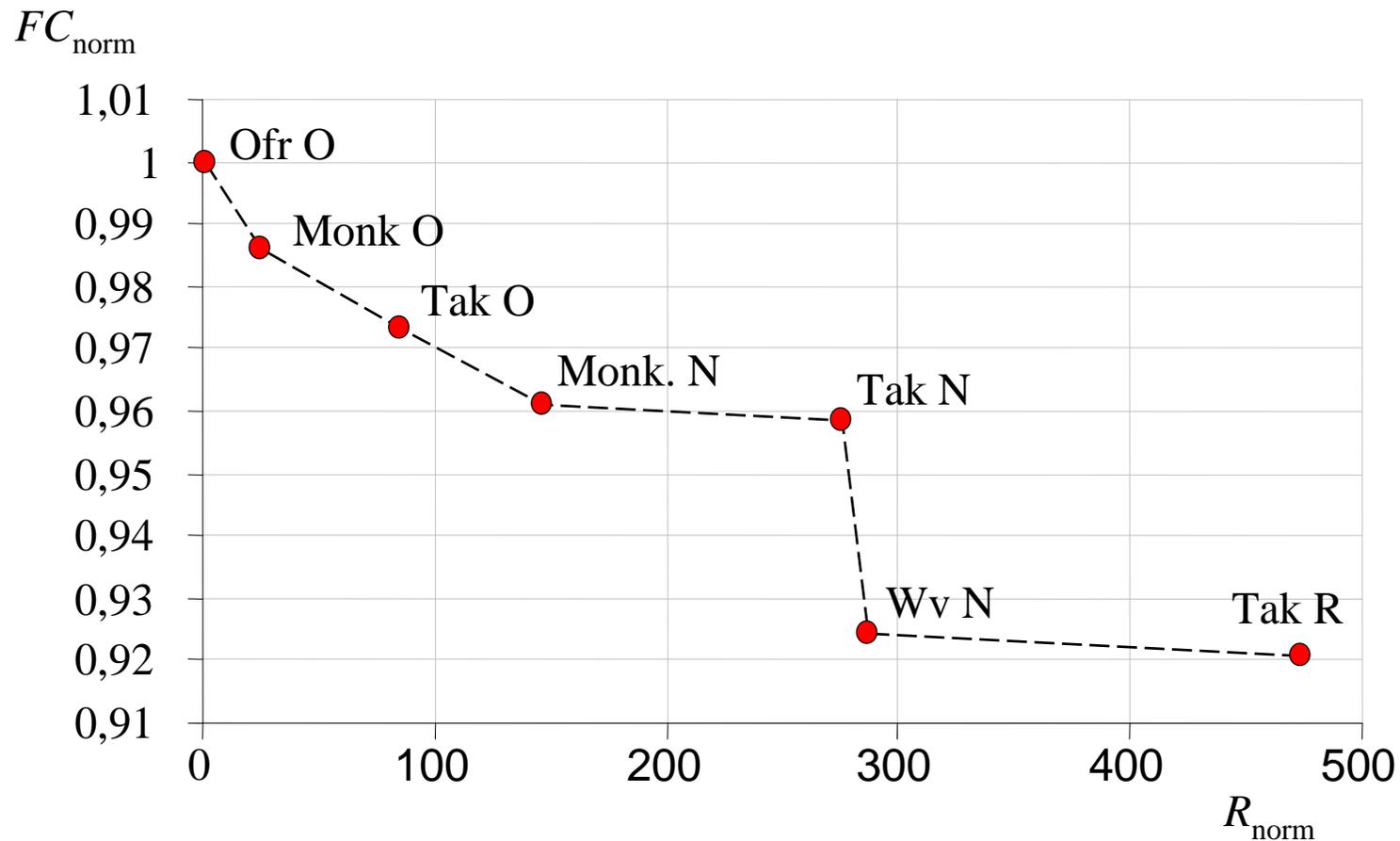
Tabla de apariciones

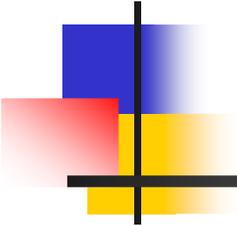
	Óptimo	Normal	Rápido
FLAC	0	0	1
LPAC	0	0	0
Monkey's	13	10	1
MP4	2	0	-
Ofr	14	7	3
Shorten	0	0	-
TAK	11	14	14
TTA	-	5	-
WavPack	0	5	2

Superposición de los C-M, sobre los 14 archivos



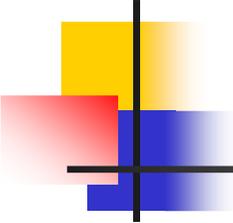
Frontera unificada final





Ejemplo de estimación del tiempo de codificación

Tema "Sledgehammer" de Peter Gabriel

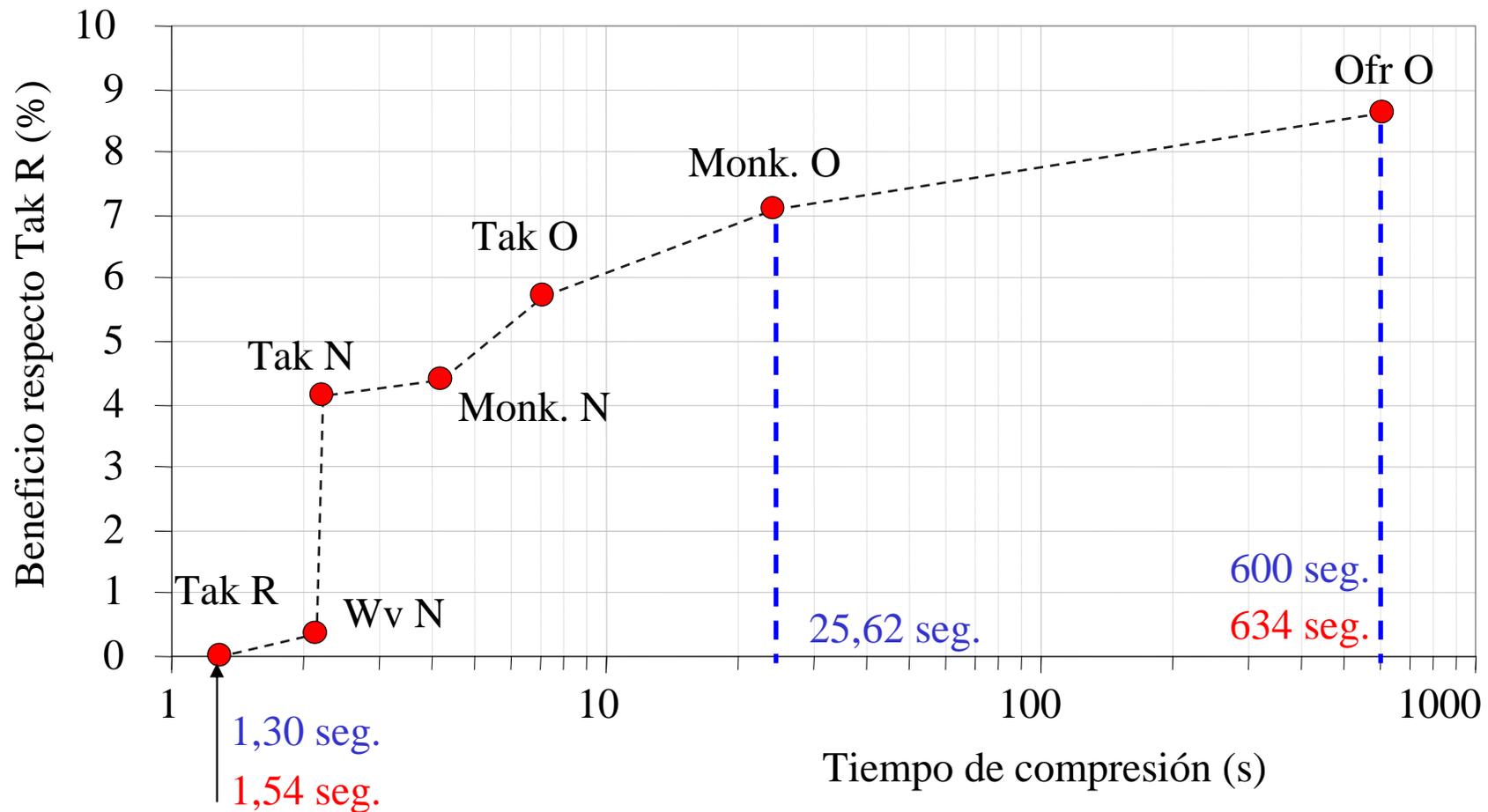


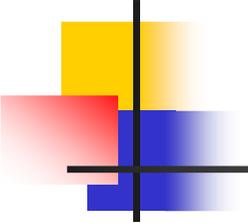
Ejemplo de estimación de tiempo

DATOS

- Duración del tema: 283 seg.
- Tiempo de compresión con Monkey's-O: 25,62 seg.
- Tiempo compresión con OFR-O: 634 seg.
- Tiempo compresión con Tak-R: 1,56 seg.

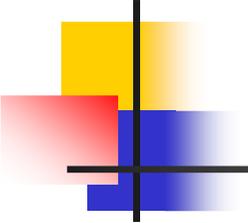
Ejemplo de estimación de tiempo





Conclusiones

- La compresión es más sensible al género musical que al codec.
- En jazz y tango se puede compactar más del 60 %.
- Las piezas de pop y rock se reducen un poco más de 35 %.
- El ruido presente en las grabaciones en vivo empeoró la compresión en 10 %.
- Se pudo caracterizar cada codec por una zona operativa de velocidad, independientemente de la PC utilizada.
- La codificación en modo óptimo supera a la del modo normal en menos del 10 %, con fuertes incrementos de demora de procesamiento. Este tiempo suele ser inferior al de reproducción por más de un orden de magnitud, excepto en MPEG-4 ALS y en Optimfrog (este caso es impráctico).



Conclusiones

- La codificación en modo rápido sacrifica típicamente menos de 5 % de compresión, pero reduce el tiempo de procesamiento a 1 % del tiempo de reproducción.
- Los tiempos de decodificación son por lo general menores que los de codificación, excepto en el caso de Monkey's Audio (simétrico).
- Adicionalmente al software de análisis de codec, se desarrolló un método gráfico orientado al usuario común, basado en la selección de C-M óptimos. Esto permite determinar el par C-M más conveniente según las necesidades de compresión y velocidad. Su empleo es extensible a cualquier PC de escritorio, ya que aproximadamente se conserva la zona operativa de cada codec.
- El método gráfico se diseñó utilizando un conjunto de codec y piezas, pero se puede extender su uso incluyendo otros codificadores y géneros musicales.



¡Gracias!

Cerro de 7 colores, Purmamarca, Jujuy