
83GIIN - Compresión y Recuperación de Información Multimedia

Actividad 3 - Portafolio

Gagliardo Miguel Angel

20 de Abril de 2025

Introducción

En esta actividad se implementará un sistema de codificación y decodificación diferencial aplicado a señales de audio de 4 archivos distintos, con el objetivo de reducir la cantidad de datos necesarios para representar la señal. Este esquema se basa en calcular las diferencias entre muestras sucesivas y cuantizarlas mediante un cuantizador uniforme.

Además de desarrollar el sistema, se realizará un **análisis previo** sobre los audios y sus señales, incluyendo medidas de dispersión, entropía y características de cada archivo.

Luego, se aplicará el codificador diferencial con distintos niveles de cuantización y se evaluará su desempeño mediante métricas objetivas como Signal-to-Noise Ratio (SNR) y distorsión, así como también una evaluación subjetiva mediante la escucha de los audios reconstruidos.

Finalmente, se discutirán y analizarán los resultados con una interpretación sobre la escucha de los audios y la eficiencia del esquema de compresión utilizado en esa actividad.

Aclaraciones

Todo el código incluido en este trabajo se puede encontrar o bien en el .zip incluido en la actividad o en el github gist:

- <https://gist.github.com/mgagliardo/7d4e1557e4ec3620988e38312ffdf1a0>

En cuanto a las librerías de python requeridas para esta actividad, son:

- notebook (para el Jupyter Notebook)
- math, numpy y pandas (cálculos y estimaciones)
- scipy (lectura de archivos .wav)
- matplotlib (gráficas)

Para los cálculos se han utilizado las siguientes fórmulas:

- **Entropía**

$$H(X) = -\sum p(x) \log_2 p(x); \text{ Fórmula de Shannon}$$

- **Varianza**

$$\text{Var}(X) = \frac{1}{N} \sum (x_i - \bar{x})^2$$

- **Desviación Estandar**

$$\sigma = \sqrt{\text{Var}(x)}; \text{ Donde var}(x) \text{ es la varianza de } x$$

- **Rango**

$$\text{rango}(x) = \max(x) - \min(x); \text{ Diff entre el valor más alto y más bajo}$$

- **Signal-to-Noise Ratio (SNR)**

$$\text{SNR} = 10 * \log_{10} \frac{\sum (x^2)}{\sum (x - x_{qr})^2}$$

- **Distorsión**

$$distorsion = \frac{\sum (x - xqr)^2}{len(x)}$$

Análisis Previo

Archivo	GolpeoPingPong.wav	PreparacionOrquesta.wav	Semaforo1.wav	Vitamina.wav
---------	--------------------	-------------------------	---------------	--------------

Tamaño del archivo	435608	1766940	88280	170032
Sample Rate (bytes)	44100	44100	44100	44100
Entropía (bits/símbolo)	5.4425	14.6286	12.2649	10.9227
Entropía de las diferencias (bits/símbolo)	4.6789	4.6789	11.3684	11.3684
Varianza Original	320938.31	320938.31	35237552.64	35237552.64
Desvío Original	566.51	566.51	5936.12	5936.12
Rango Original	64990	64990	40084	40084
Varianza de las Diferencias	7753.93	7753.93	9526787.62	9526787.62
Desvío de las Diferencias	88.06	88.06	3085.55	3086.55
Rango de las Diferencias	64990	64990	40084	40084

Aplicación del Codificador y Decodificador Diferencial

Aclaración: A la hora de estimar las distorsiones se ha utilizado un redondeo, por lo que

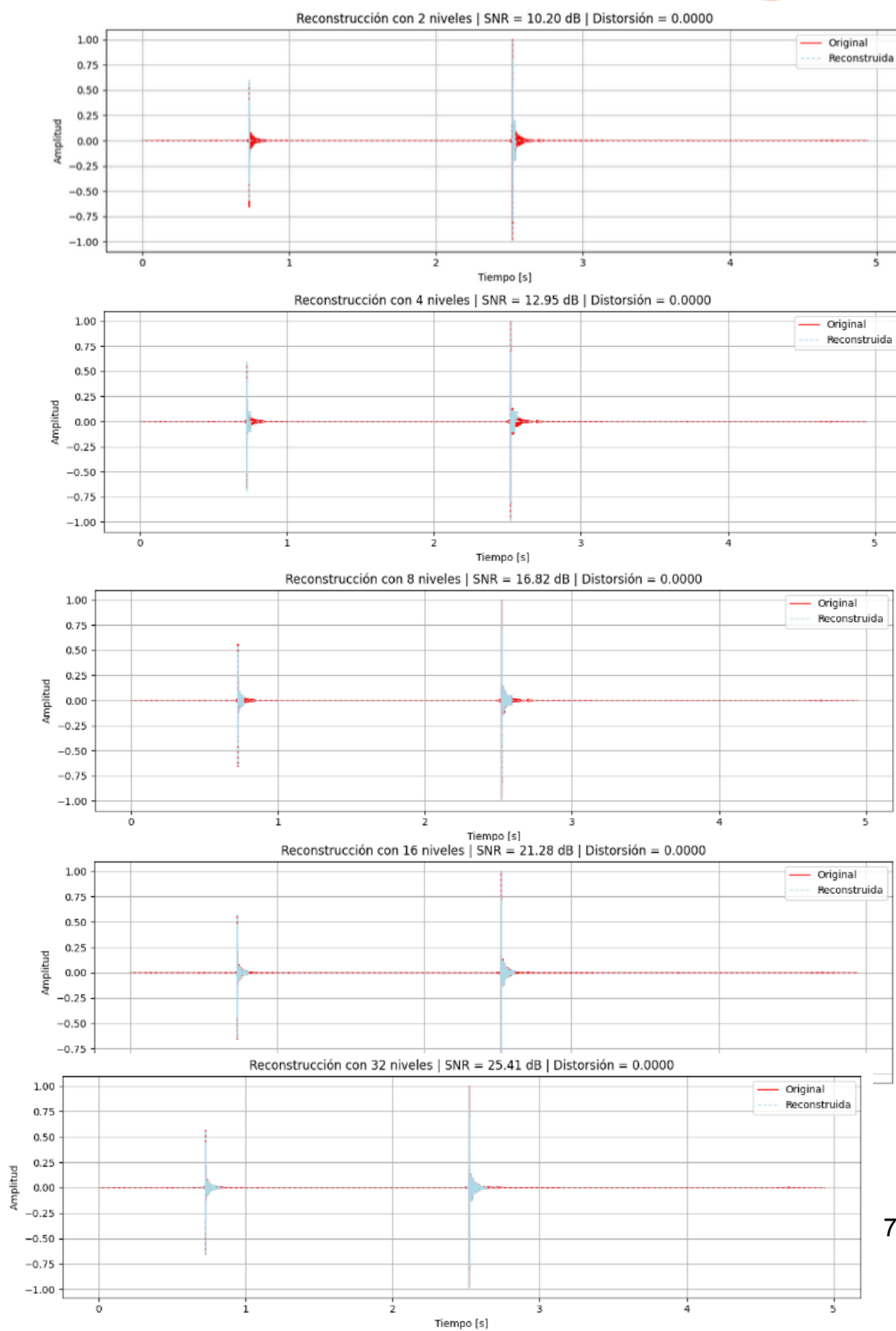
números muy pequeños (como $N \times 10^{-5}$) serán contabilizados como un 0 (cero).

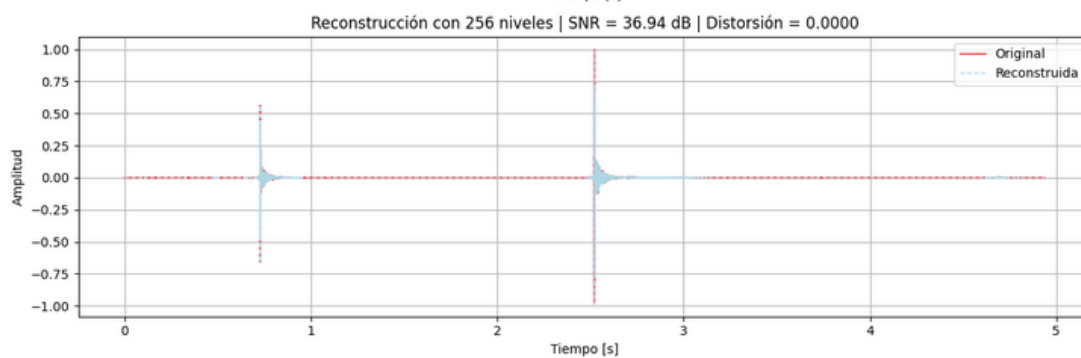
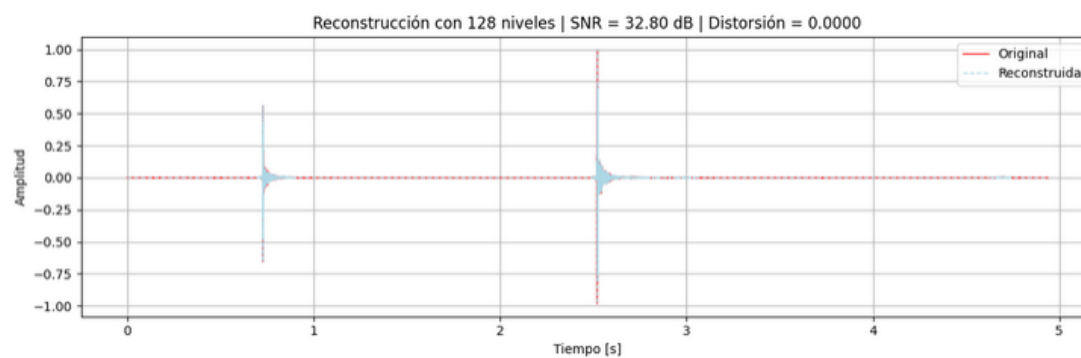
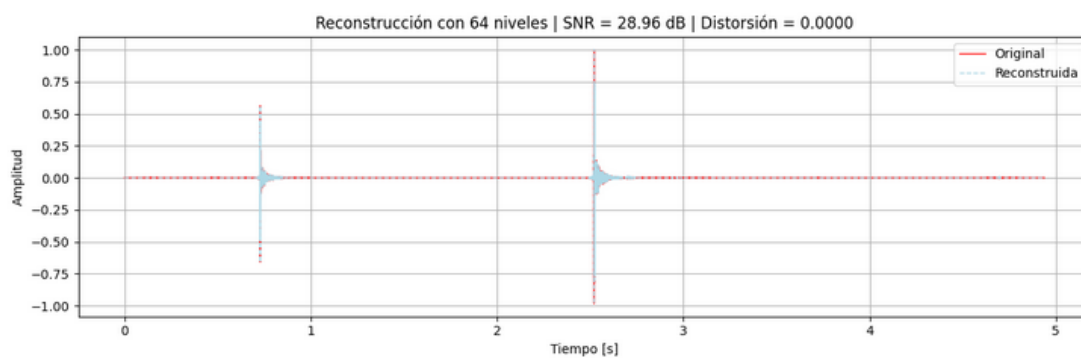
Archivo	Intervalo (N)	GolpeoPingPong.wav	PreparacionOrquesta.wav	Semaforo1.wav	Vitamina.wav
Distorsión Introdúcida	2	0	0.0017	0.0192	0.0024
	4	0	0.0004	0.0053	0.0008
	8	0	0.0001	0.0014	0.0002
	16	0	0	0.0003	0.0001
	32	0	0	0.0001	0
	64	0	0	0	0
	128	0	0	0	0
	256	0	0	0	0
SNR (dB)	2	10.20	16.71	6.56	6.29
	4	12.95	22.73	12.15	11.06
	8	16.82	28.75	17.98	16.34
	16	21.28	34.73	24.06	21.69
	32	25.41	40.71	29.95	27.72
	64	28.96	46.72	35.74	33.84
	128	32.8	52.73	41.58	39.85
	256	36.94	58.73	47.33	45.93

Los audios están incluidos y se pueden reproducir utilizando el archivo jupyter incluido en el zip o bien en el gist antes mencionado.

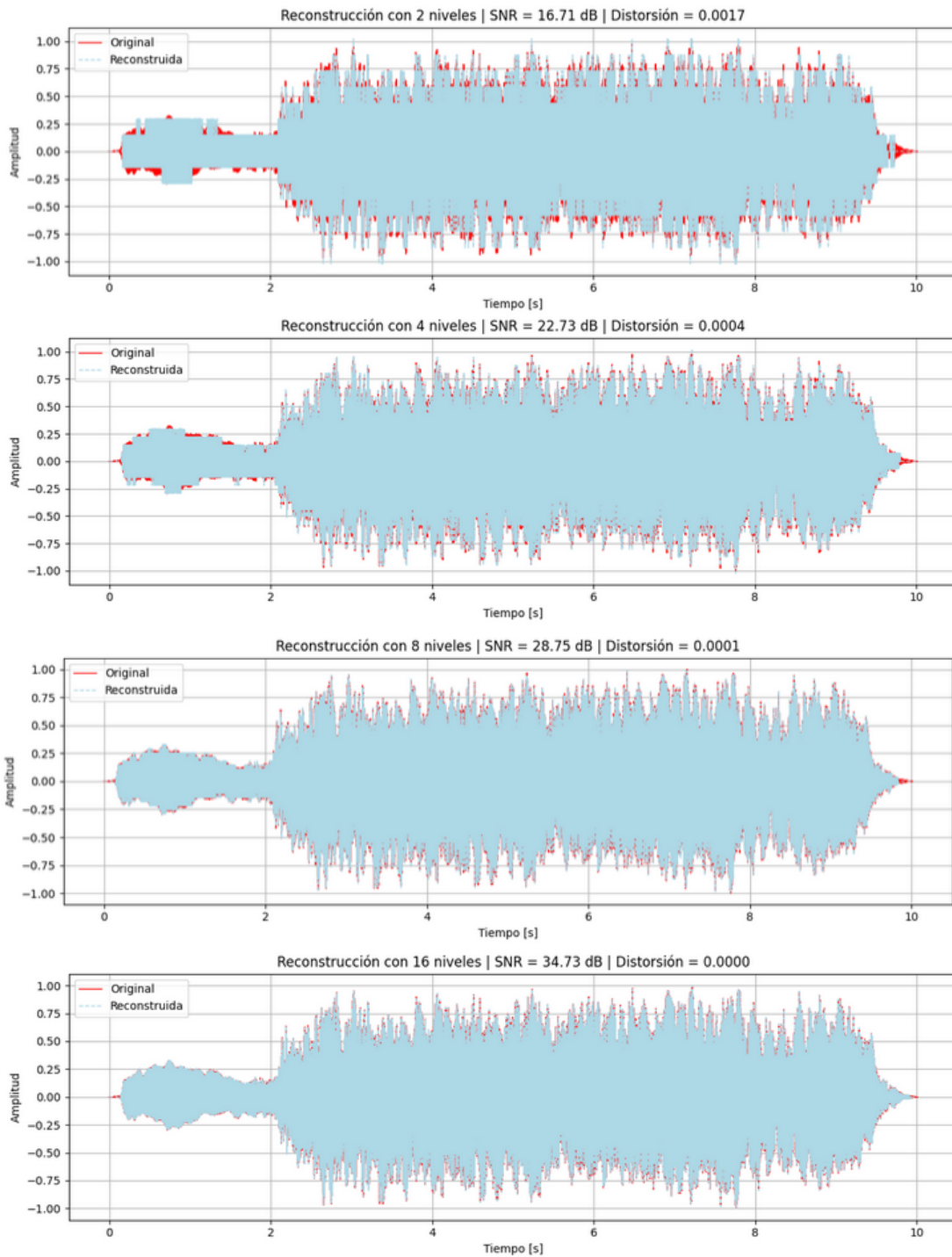
Gráficos de señales reconstruidas

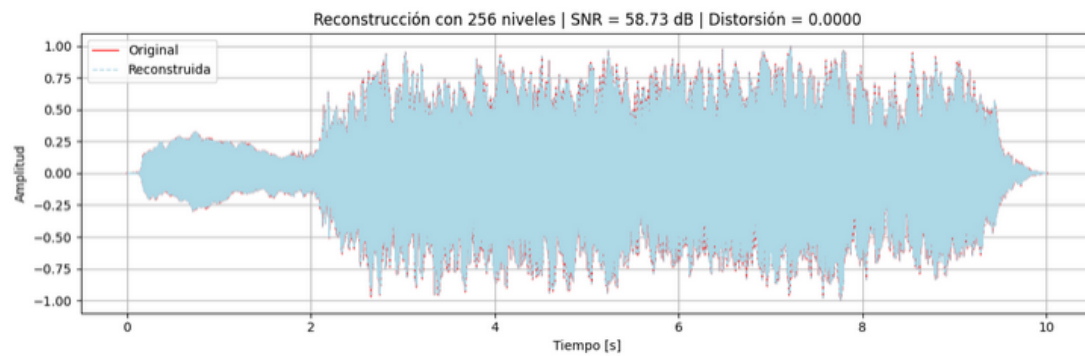
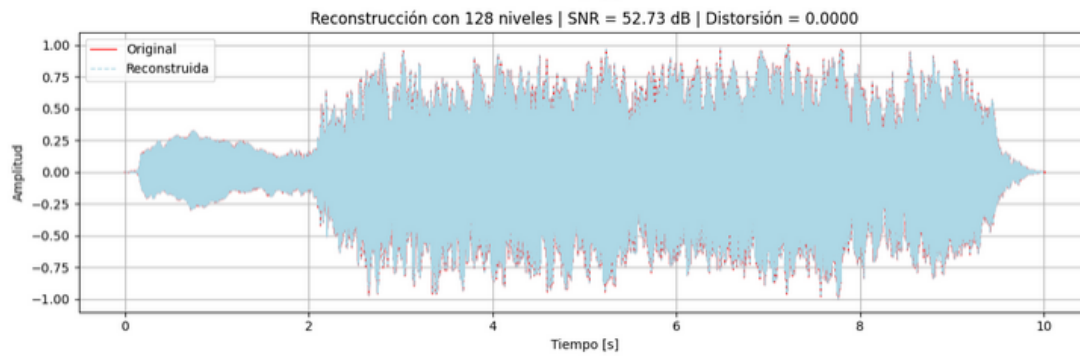
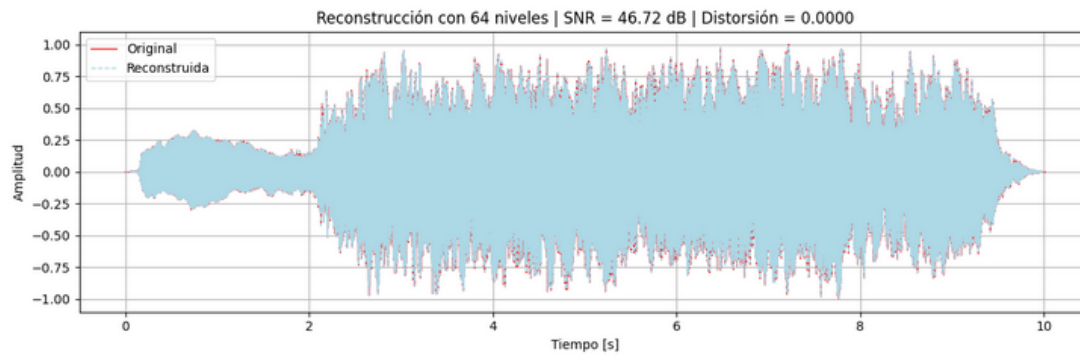
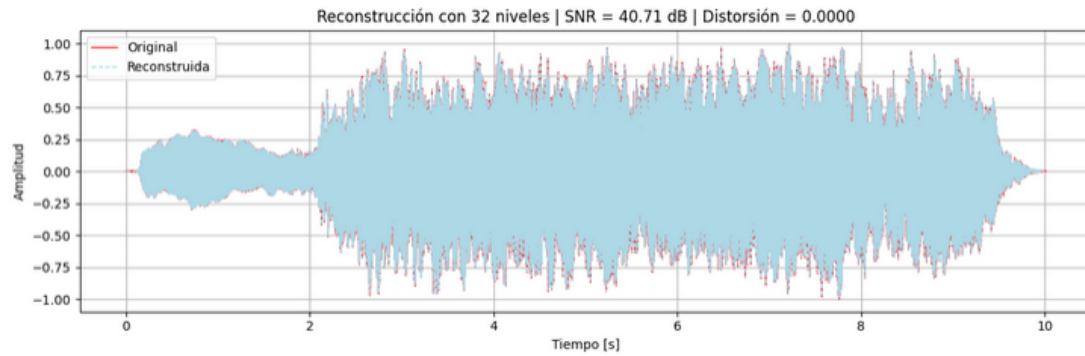
GolpeoPingPong.wav



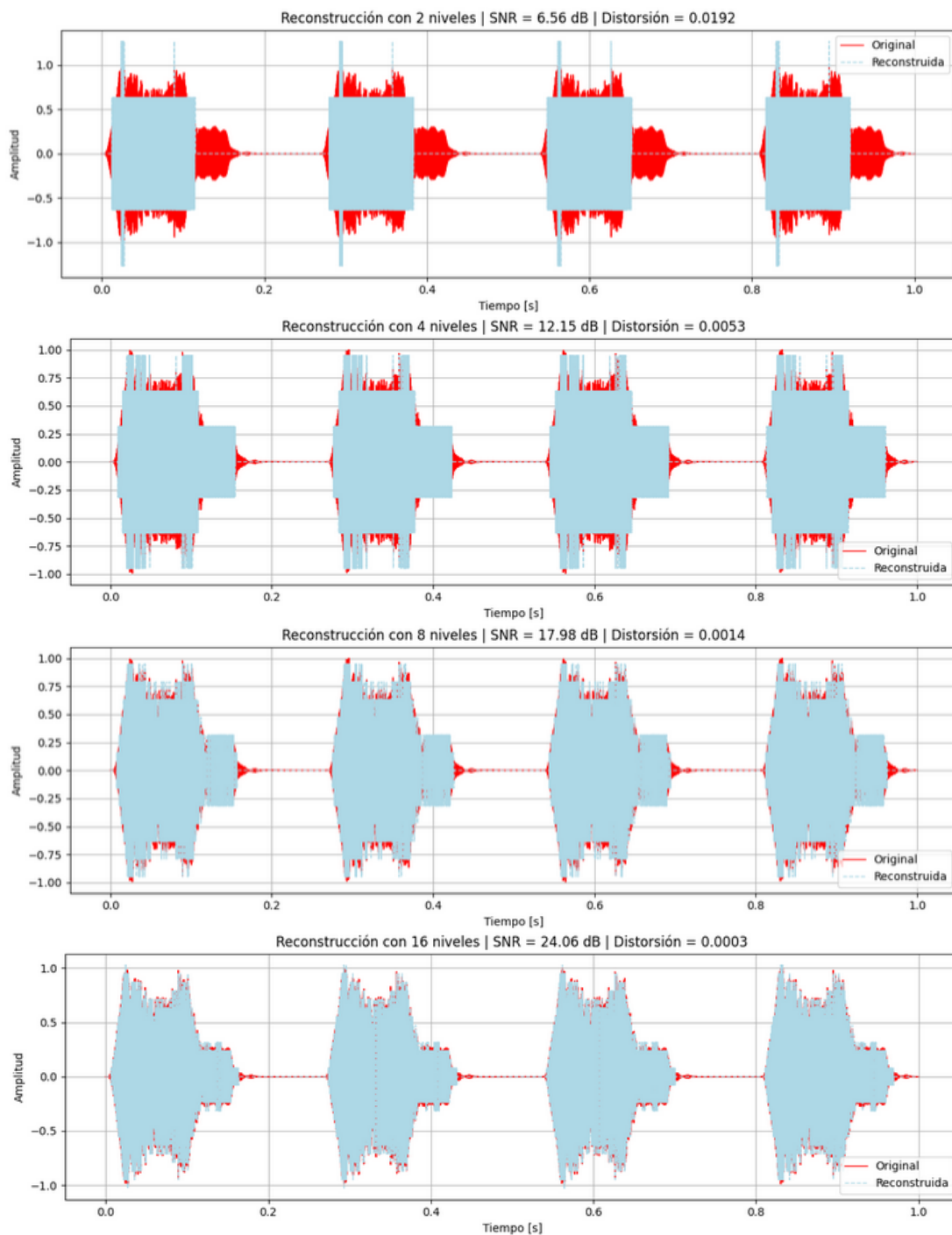


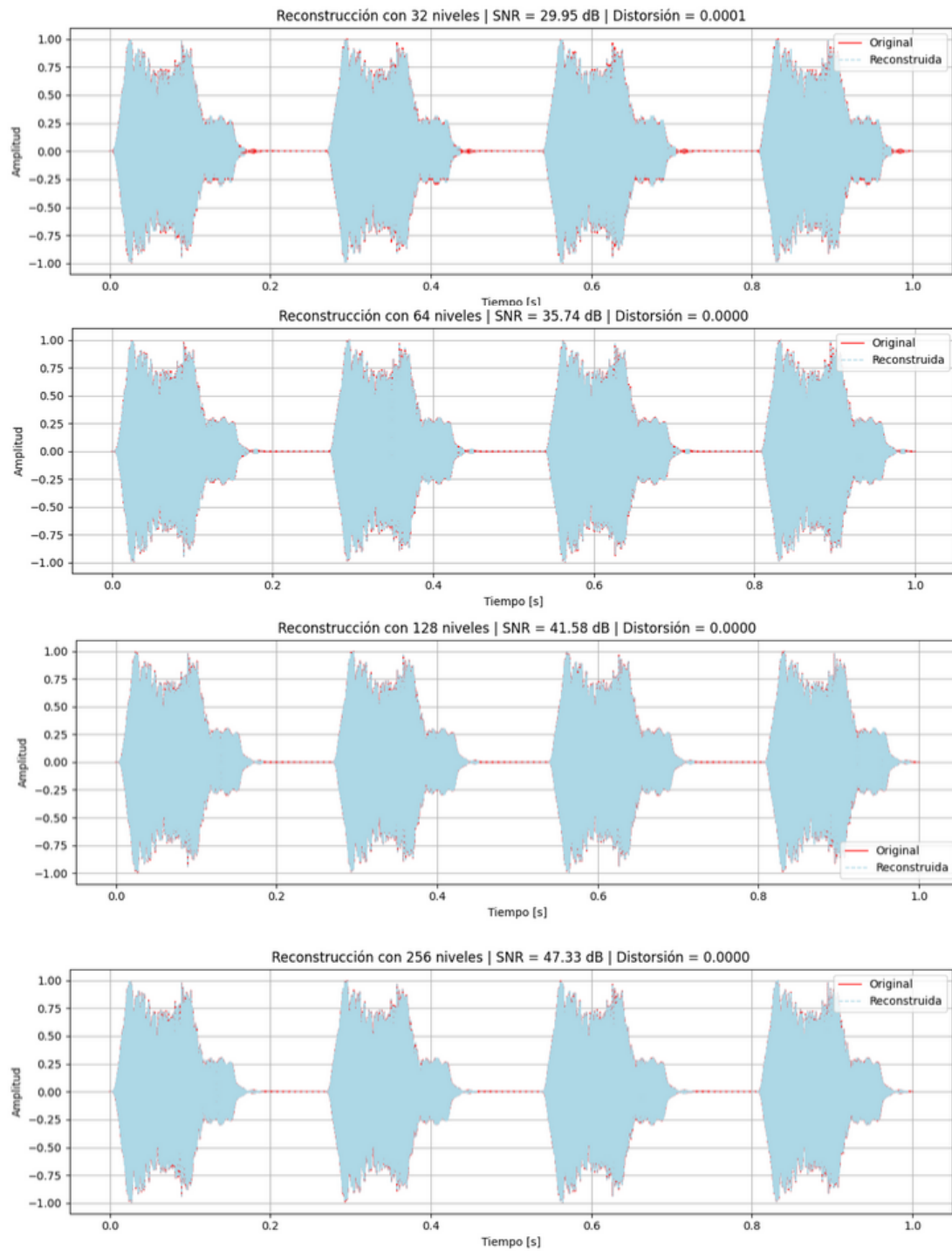
PreparacionOrquesta.wav



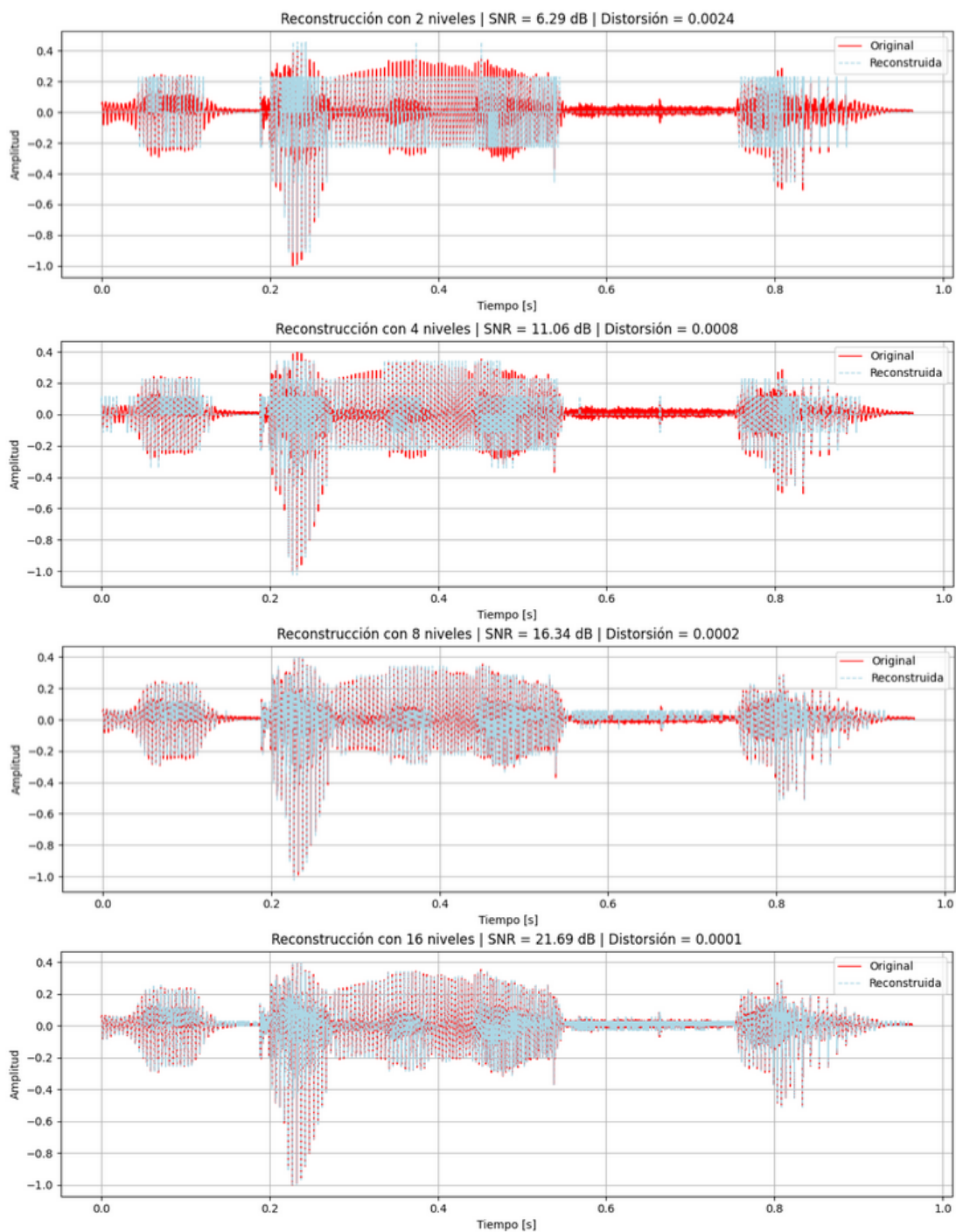


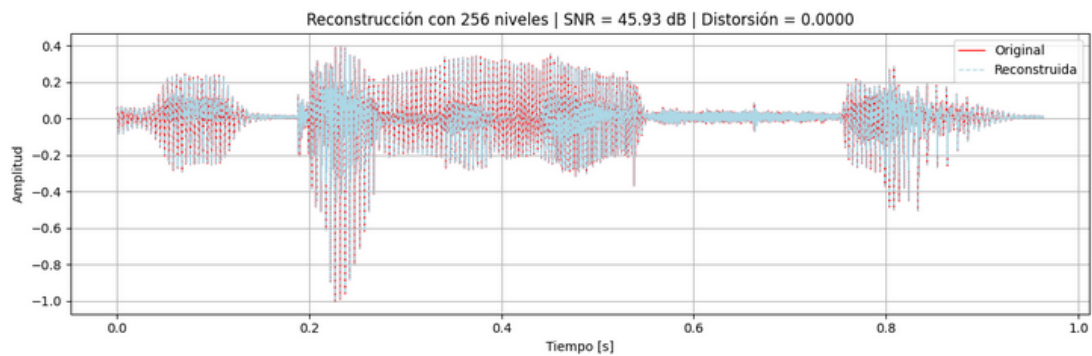
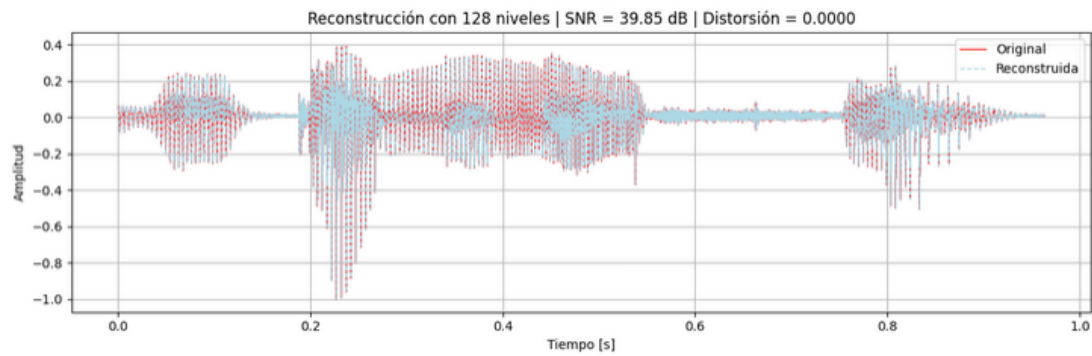
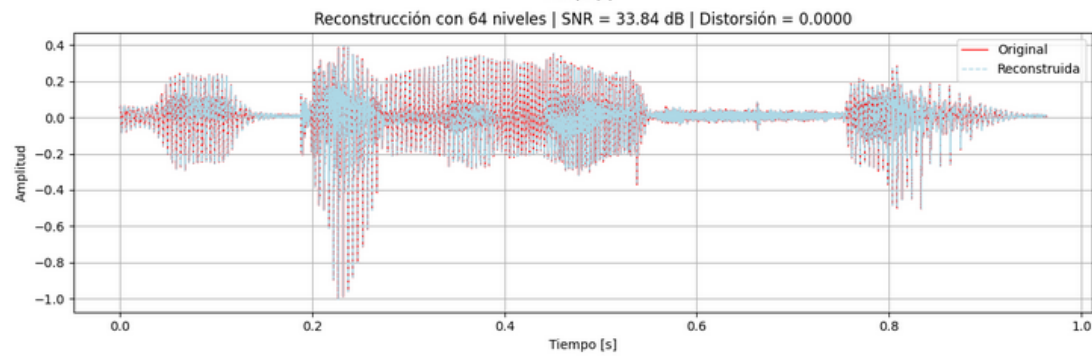
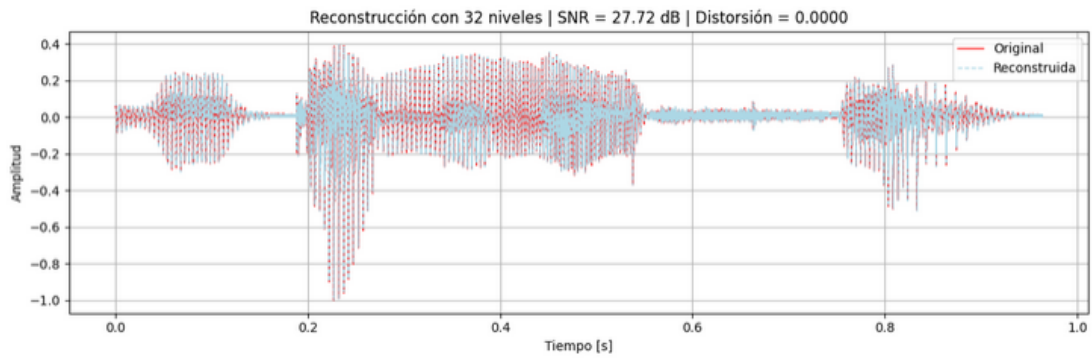
Semaforo1.wav





Vitamina.wav





Conclusión

En este trabajo se ha implementado un **codificador y decodificador diferencial**, se analizaron 4 archivos distintos y se han evaluado sus estadísticas básicas (peso del archivo, varianza, desvío estándar y entropía) y luego, aplicando distintas configuraciones del cuantizador (con intervalos de 2 a 256 usando $N=2^n$), se observó:

- Se observa **en todos los casos** que **al aumentar los intervalos disminuye la distorsión**, dado que **al tener menor cantidad de intervalos (N)** podemos representar menores valores de diferencia (\hat{d}_n) y **las aproximaciones se vuelven más “bruscas”**.
Por lo tanto al aumentar los intervalos aumenta la capacidad de representar dichos valores, ergo su error es menor o bien tenemos menor “ruido” o pérdida de calidad/fidelidad.
- En caso contrario, **al aumentar los intervalos aumenta el SNR (signal-to-Noise Ratio)**, dado que al disminuir la distorsión disminuye el denominador en la ecuación (ver sección “aclaraciones”), o sea el cociente en la misma aumenta como así también su logaritmo y por tanto los decibeles (dB), o sea el SNR finalmente, aumenta.
- En cuanto a lo **auditivo/perceptivo**, dado que al utilizar intervalos más pequeños (menor N) podemos representar finalmente la señal de una manera “menos fidedigna” se sienten saltos más bruscos (dado que el \hat{d}_n es menor) y por eso el ruido se nota más y la calidad del audio es menor.