

Trabajo Práctico de Señales y Sistemas

Juan Manuel Pérez

19 de junio de 2017

1. Ejercicio 1 y 2

Recordemos que un espectrograma es una representación visual donde se muestra el espectro de una señal en función del tiempo. Para esto, se separa el dominio del tiempo en ventanas y para cada una se aplica una Transformada de Fourier de Tiempo Corto (*STFT*).

La *STFT* tiene los siguientes parámetros:

- Tipo de ventana
- Tamaño de ventana
- Step
- Tamaño de las FFTs

La elección de estos parámetros nos define la resolución tanto en el tiempo como en el espectro: a mayor tamaño de ventana (mayor N), menos resolución temporal y mayor resolución en las frecuencias; inversamente, si hago decrecer el N obtengo mejor resolución en el tiempo pero peor en las frecuencias.

Por lo tanto, es necesario entender el trade-off entre ambas y tomar una decisión al respecto. En nuestro caso, al estar analizando piezas de piano, nos interesa obtener una resolución en frecuencias tal que nos permita distinguir entre dos notas. Así mismo, también queremos tener una resolución temporal que nos permita distinguir temporalmente la ocurrencia de cada nota.

1.1. Elección de parámetros

De la tabla de frecuencias de las notas observamos que la distancia mínima entre semitono es de 31hz, con lo cuál necesitamos tener tal resolución como mínimo.

Haciendo una cuenta con las DFTs, obtenemos que el tamaño de la ventana debe ser mayor a 34ms.

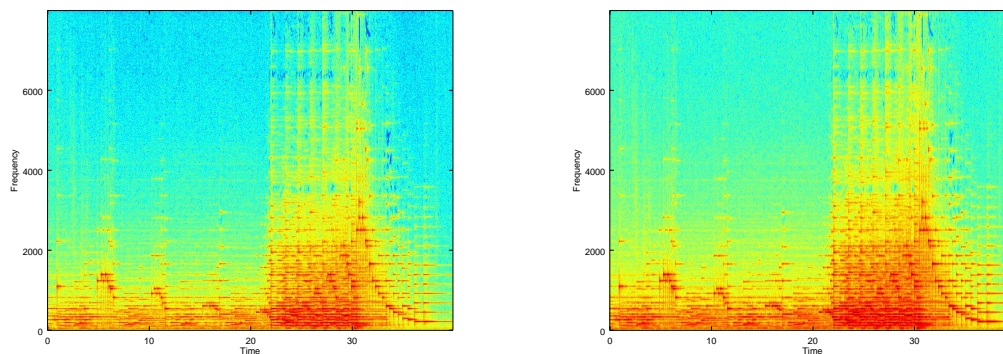


Figura 1.1: Espectrograma de ventana 100 ms y step de 50 ms

2. Ejercicio 3

En este primer intento de duplicar el largo del audio, interpolamos entre muestras del audio.

Como es de esperarse, esto baja a la mitad el pitch de cada nota, ya que podemos pensar que la onda que antes tardaba N muestras en ciclar ahora le va a tomar $2N$. A su vez, la interpolación lineal agrega armónicos de alta frecuencia, como podemos observar en el espectrograma de la figura 2.1.

3. Ejercicio 4

En este punto se nos pide reducir el audio a la mitad de la duración usando una decimación de orden dos, es decir

$$Y[n] = X[2n] \quad (3.1)$$

Sabemos por propiedades de la transformada de Fourier en tiempo discreto que esto genera un ensanchamiento del espectro, con lo cuál debemos aplicar un filtro pasa bajos para evitar tener aliasing. El tipo de filtro elegido fue un Butterworth de orden 10 y frecuencia de corte $\frac{\pi}{2}$, es decir, la mitad de la frecuencia de Nyquist.

En la figura 3.1 podemos observar el espectrograma. Como es de esperarse, las frecuencias de las notas se multiplican por dos (suben una octava) producto de que son reproducidas a “mayor velocidad”.

4. Ejercicio 5

Este punto nos pide reconstruir una señal en base a su espectrograma. Este problema se conoce como la transformada inversa de Fourier de tiempo corto.

En ¹, uno de los primeros trabajos en el área, se discuten dos alternativas para esta tarea. La que decidimos utilizar fue similar al método llamado OverLap Addition (OLA): de cada rebanada espectral, reconstruir un pedazo de la señal y luego concatenarlos.

Para esto,

- Rellenamos con los conjugados simétricos
- Calculamos la *ifft* de cada rebanada
- Dividimos por una ventana de *hanning*
- Concatenamos los distintos segmentos con un *offset*

El último paso surgió como medida paliativa ya que notábamos que los segmentos iniciales y finales de cada pieza reconstruída resultaban sensiblemente atenuados.

En el gráfico 4.1 se puede observar la diferencia entre el audio reconstruído y el audio original. En este caso, la reconstrucción es casi perfecta (salvando principio y final del audio).

¹<http://ieeexplore.ieee.org/sci-hub.cc/abstract/document/1455039/#>

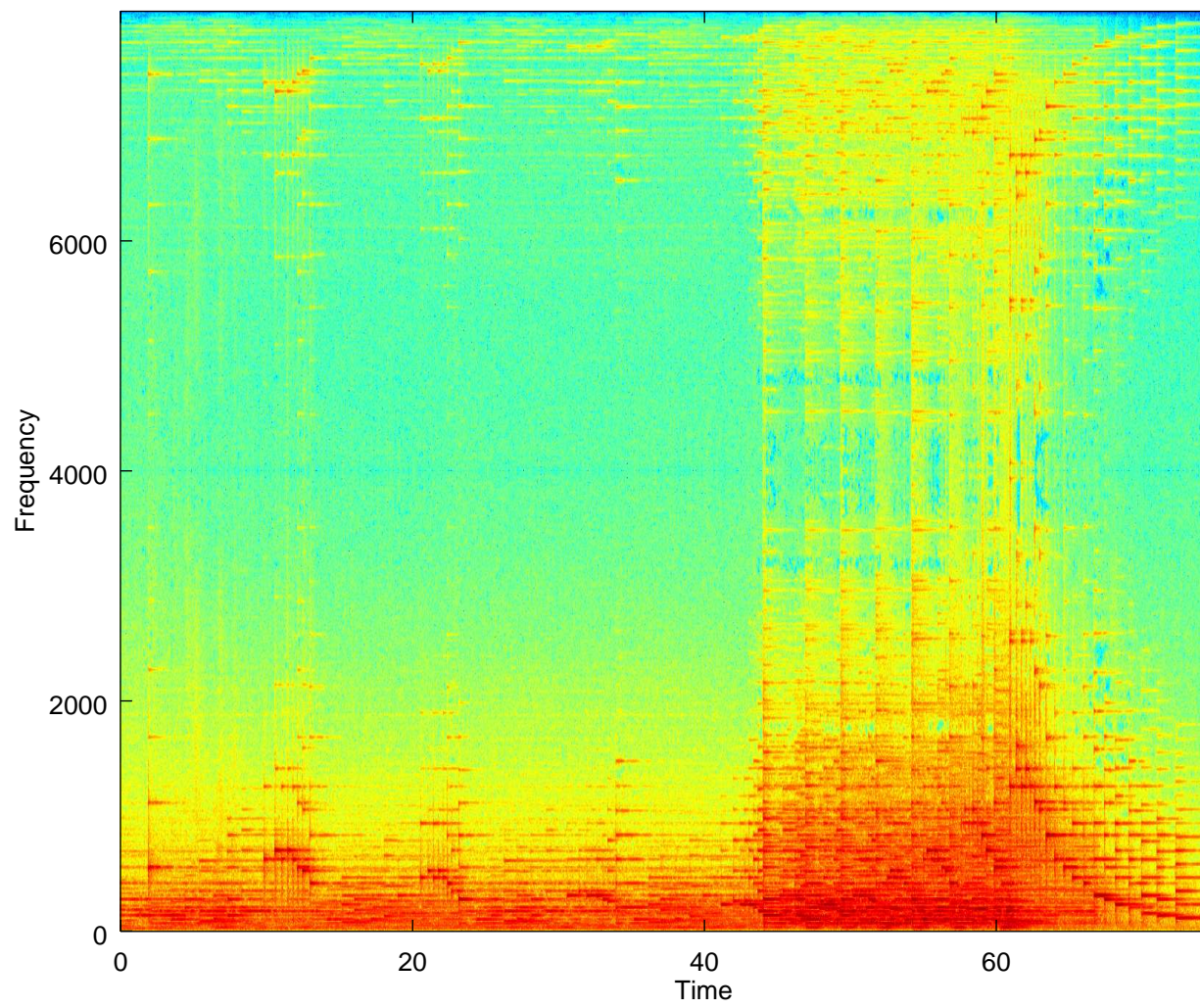
5. Ejercicio 6

En este ejercicio, utilizando la función de síntesis del anterior punto, realizamos la misma idea que en el punto 3 pero esta vez interpolando el espectrograma, con la esperanza de obtener una canción que sea el doble de larga pero con el mismo espectro que la original. Lamentablemente, si bien el grueso energía del espectro se mantiene en las mismas frecuencia (es decir, son las mismas notas) en el proceso de síntesis se agregan artefactos de alta frecuencia que no supimos cómo remover. Estimamos que dichos artefactos se producen ya que no “empalmamos” adecuadamente los distintos segmentos de la señal.

6. Ejercicio 7

Análogamente a lo hecho en el ejercicio 4, queremos acortar el audio a la mitad pero esta vez manteniendo el contenido espectral. Para ello, aplicamos la reconstrucción sobre una decimación de orden 2 del espectro.

De la misma manera que nos ocurrió en el ejercicio anterior, si bien mantenemos el grueso de la energía en la misma zona de frecuencias, tenemos muchos artefactos (que suenan como un “traqueteo”) que no pudimos resolver. Podemos observar el espectrograma en



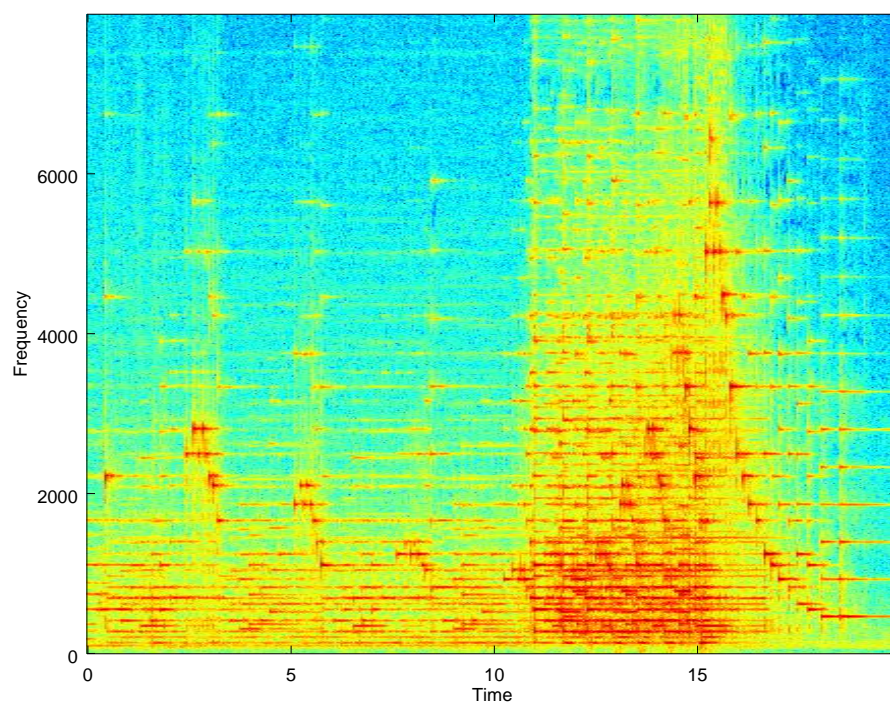


Figura 3.1: Espectrograma del audio decimado

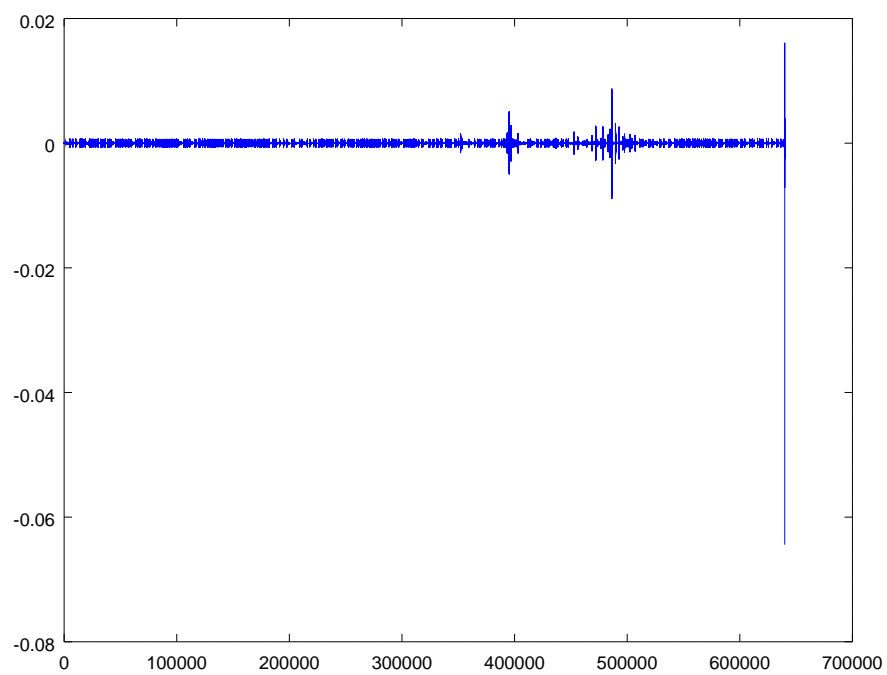


Figura 4.1: Diferencia entre audio y audio reconstruido espectralmente

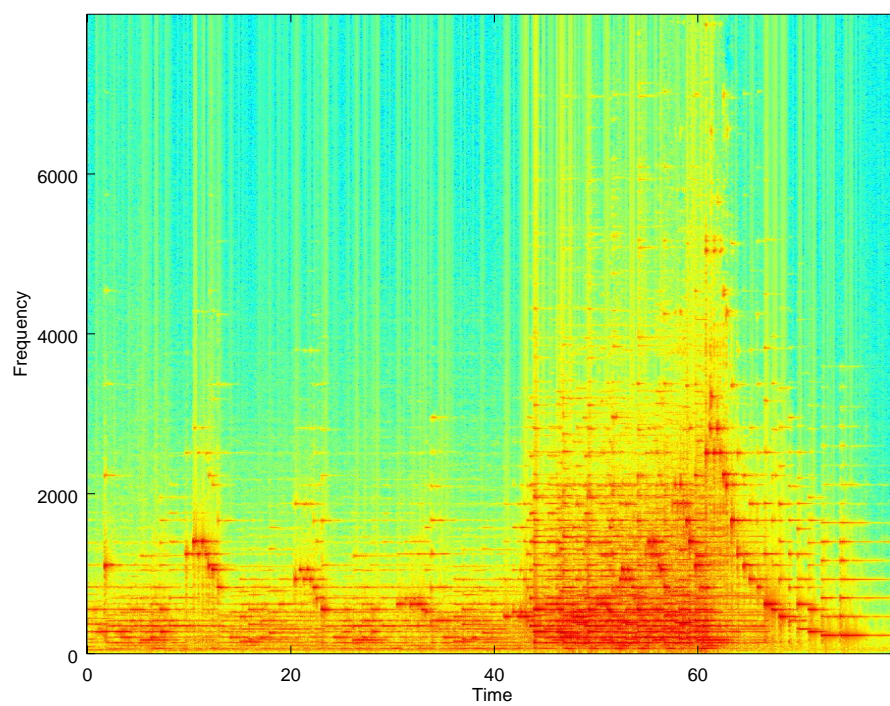


Figura 5.1: Espectrograma del audio interpolado espectralmente

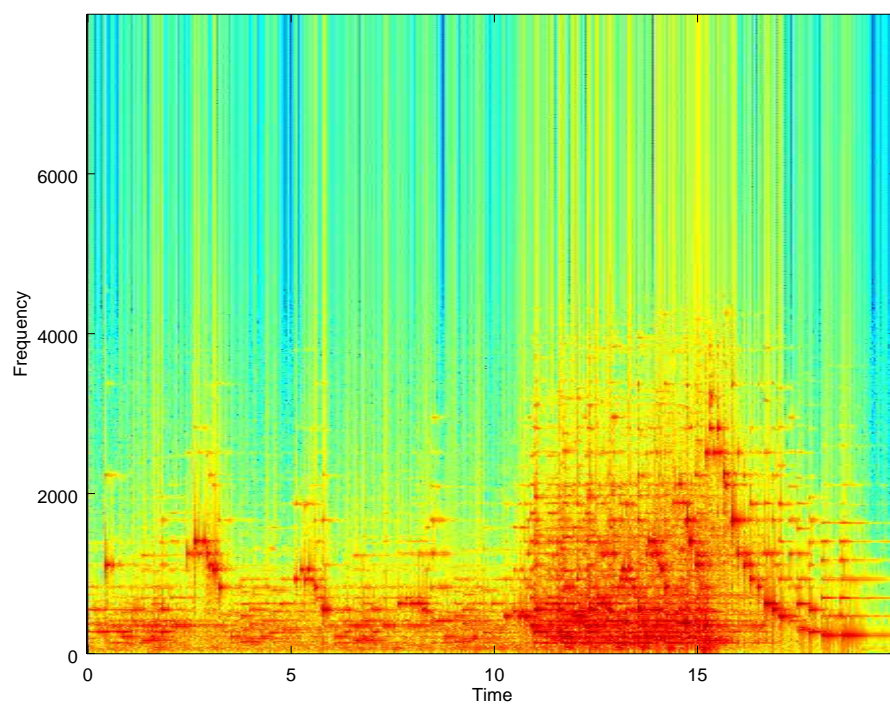


Figura 6.1: Espectrograma del audio decimado espectralmente