

TB065-86.05 - SEÑALES Y SISTEMAS

TRABAJO PRÁCTICO ESPECIAL 1: ANÁLISIS DE LA SEÑAL DE HABLA

🐍 = Python.

La señal de habla está compuesta de secciones de propiedades cambiantes. Si pensamos a la señal de habla como la salida de un sistema, podemos atribuir dichas variaciones a dos causas: cambios en la excitación o cambios en la configuración del tracto vocal, es decir en el sistema. Si la entrada se comporta como un tren de impulsos cuasi-periódicos, la salida será uno de los posibles sonidos vocálicos (/a/, /e/, /i/, /o/, /u/, /m/, /n/, /l/). Si la entrada en cambio es un generador de ruido blanco, el sonido obtenido será un fonema fricativo (/s/, /f/, /sh/). La distinción entre los fonemas de la misma clase se produce por la forma que va tomando el tracto vocal para cada uno de ellos. La variación de la transferencia del sistema se supone que es suficientemente lenta como para considerar que la señal de habla es la concatenación de porciones de señales que se originan como salida de un sistema LTI. Por esto aparecerán bien representados en un espectrograma. Los sonidos explosivos (/p/, /k/, /t/) en cambio tienen una naturaleza distinta, y son más parecidos a un transitorio que a un sonido estacionario. Un esquema del modelo de producción de la voz se muestra en la Figura 1.

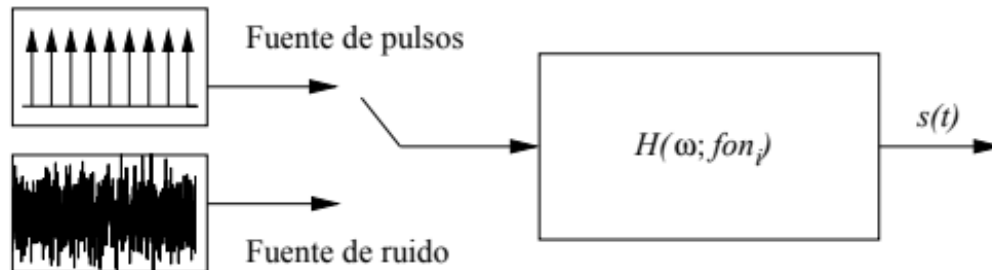


Figura 1: Modelo de producción de la voz

1. Grabar 2 muestras de su propia voz diciendo la palabra “Picasso”. Una rápida y otra lenta, de forma que una tenga aproximadamente el doble de duración que la otra. Graficar en 🐍 las señales de voz ubicando porciones de señales periódicas y no periódicas.

2. Realizar una segmentación 🐍 de la señal lenta localizando los segmentos de muestras donde aparece una [a] y una [s]. Describir qué diferencias hay entre estos fonemas. Graficar 🐍 los segmentos cuasi periódicos de la señal lenta y estimar el periodo y la frecuencia. Calcular lo mismo para la señal rápida.

Los sonidos sonoros son producidos forzando el aire a través de la glotis o a través las cuerdas vocales. La tensión de las cuerdas vocales se ajusta de manera tal que vibre en forma oscilatoria. La interrupción periódica del flujo de aire subglotal resulta en un soplido casi periódico de aire que excita el tracto vocal. El sonido producido por la laringe es llamado sonoro o con fonación. Este tipo de sonido consiste en una frecuencia fundamental (F_0) y sus componentes armónicos producidos por las cuerdas vocales. El tracto vocal modifica esta señal de excitación causando el formante. El término formante se utiliza para indicar el centro de estas frecuencias de resonancia, es decir, los picos de la envolvente del espectro de la señal de voz que representan las frecuencias de resonancia del tracto vocal. Cada formante tiene una frecuencia central, amplitud y un ancho de banda, y son usualmente denotadas F_1 , F_2 , F_3 ,..., comenzando con la menor frecuencia. Las frecuencias a las que se producen los primeros formantes son muy importantes para reconocer o sintetizar la voz. En la siguiente figura pueden verse representados los 3 primeros formantes de una señal de voz.

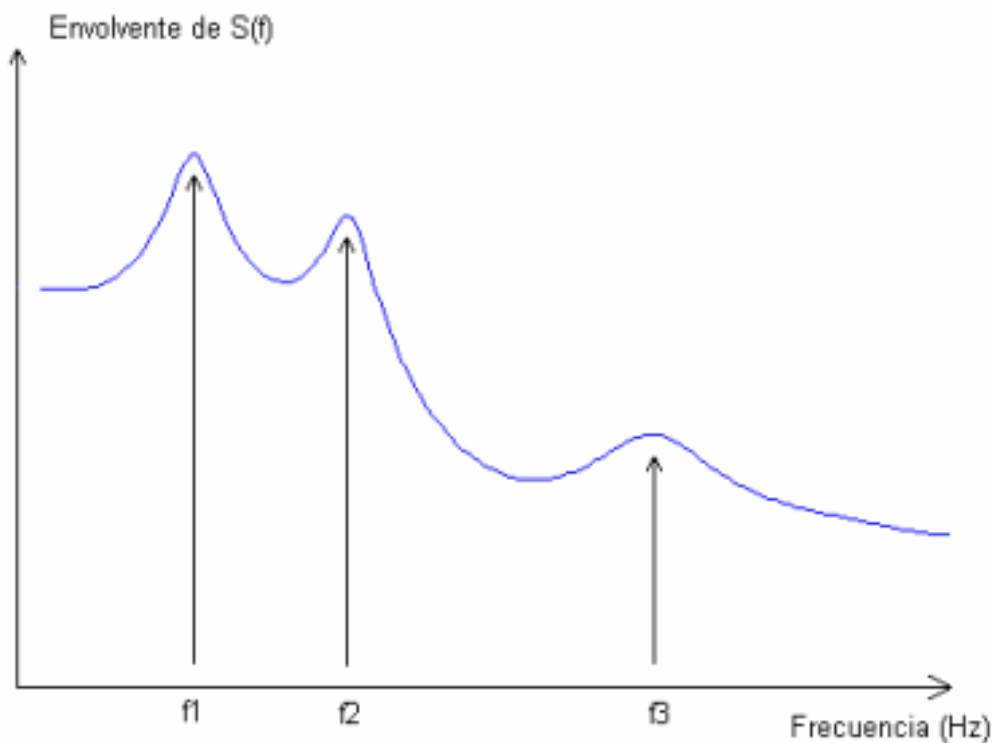



Figura 2: Envolvente del espectro de una vocal.

3. Utilizando la FFT  grafique los coeficientes de Fourier de las porciones correspondientes a las vocales que hay en la señal. Hacer el cálculo tomando varios períodos de la vocal y también tomando un solo período para las 2 señales. Identificar los primeros máximos en las envolventes de estos espectros y estimar los valores de frecuencia en que se producen.