

Práctica 4. Cuantización de parámetros del modelo LPC de producción de voz

1 Cuantización de los parámetros del modelo LPC de producción de voz

Diseñe un cuantizador escalar para el factor de ganancia G y los coeficientes a_i usando el algoritmo de Lloyd descrito en clase. Utilice 8 bits para G y en lugar de cuantizar directamente los coeficientes a_i transfórmelos primero en los coeficientes de reflexión usando la siguiente fórmula:

$$k_i = a_i^{(i)}$$

$$a_j^{(i-1)} = \frac{a_j^{(i)} + a_i^{(i)} a_{i-j}^{(i)}}{1 - k_i^2} \quad 1 \leq j \leq p-1$$

donde $i = p, p-1, \dots, 1$ e inicialmente

$$a_j^{(p)} = a_j, \quad 1 \leq j \leq p$$

Son los coeficiente de reflexión los que ha de cuantizar con la siguiente distribución de bits: 7 bits para k_1 , 6 bits para k_2 , 5 bits para k_3 y 4 bits para cada k_4, \dots, k_{10} . Utilice cuatro de los ficheros del directorio **8k** como base de datos para diseñar los cuantizadores.

Una vez diseñados los cuantizadores de los parámetros G y k_i obtenga la señal de voz sintetizada resultante de cuantizar los parámetros G y k_i de cada segmento de voz de 15 ms (utilice el fichero de voz que no usó para diseñar los cuantizadores) utilizando además la señal

$u(n)$ no cuantizada. Para ello ha de volver a determinar los a_i a partir de los k_i cuantizados usando la siguiente fórmula:

$$a_i^{(i)} = k_i$$

$$a_j^{(i)} = a_j^{(i-1)} - k_i a_{i-j}^{(i-1)}, \quad 1 \leq j \leq i-1$$

con $i = 1, 2, \dots, p$. La solución final viene dada por

$$a_j = a_j^{(p)}, \quad 1 \leq j \leq p$$

Realice un estudio de la degradación introducida por el proceso de cuantización de los parámetros G y k_i sobre la señal de voz tanto objetiva como subjetivamente.

Obtenga también la señal de voz resultante si utiliza el modelo LPC simplificado para la excitación y evalúe asimismo la calidad de la señal de voz obtenida tanto objetiva como subjetivamente. ¿Si utilizásemos 7 bits para cuantizar el pitch y 1 bit para la decisión sonoro/sordo, cuantos kilobytes ocuparía el fichero de voz ahora? ¿Qué factor de compresión hemos conseguido respecto al tamaño del fichero de voz original?