

Práctica 1. Determinación de la frecuencia fundamental en una señal de voz mediante parámetros en el dominio del tiempo

1 Fundamentos teóricos

1.1 Introducción

Los parámetros que más se utilizan para caracterizar a la señal de voz en el dominio del tiempo son la energía E_n , los cruces por cero Z_n y la autocorrelación R_n . La definición de dichos parámetros para una ventana de tamaño N viene dada por

$$E = \sum_{n=0}^{N-1} s^2(n)$$

$$Z = \frac{1}{2} \sum_{n=1}^{N-1} |sgn[s(n)] - sgn[s(n-1)]|$$

$$sgn[s(n)] = 1 \quad s(n) \geq 0$$

$$sgn[s(n)] = -1 \quad s(n) < 0$$

$$R(k) = \sum_{n=|k|}^{N-1} s(n)s(n-|k|)$$

En esta práctica vamos a utilizar algunos de dichos parámetros para la determinación de la frecuencia fundamental de una señal de voz.

1.2 Algoritmo para la determinación del pitch usando la función de autocorrelación

Existen muchos algoritmos para la determinación del pitch utilizando la función de autocorrelación. Uno de ellos es el que se describe a continuación, en el cual se utiliza la técnica de "center clipping" descrita en clase.

Pasos del algoritmo:

- 1. La señal muestreada a 8kHz es filtrada con un filtro pasa-baja a 900 Hz.
- 2. Se toman segmentos de longitud 30 ms (240 muestras) a intervalos de 15 ms. Por lo tanto los segmentos se solapan en 15 ms.
- 3. El nivel de "clipping" se determina como un porcentaje fijo (68%) del mínimo de los valores máximos absolutos en las primeras y las últimas 80 muestras del segmento de voz.
- 4. Se determina la función de autocorrelación de la señal transformada mediante "center clipping" sobre un rango que abarque los posibles valores de periodo de pitch (de 20 a 120 muestras).
- 5. Se localiza el mayor pico de la función de autocorrelación y dicho valor de pico es comparado con un umbral (el 30% de $R(0)$). Si el valor de pico cae por debajo del umbral el segmento es clasificado como sordo y si cae por encima como sonoro, siendo en este último caso el periodo de pitch la localización del pico más alto.

Implemente este algoritmo y determine qué segmentos son sordos y cuales son sonoros y en este último caso determine su periodo de pitch para una de las frases que se le suministran. Una vez realizado esto chequee manualmente si el algoritmo ha cometido algun error e indique el % de error del algoritmo.