



Arquitectura de comunicaciones Móviles

Codificación de la voz

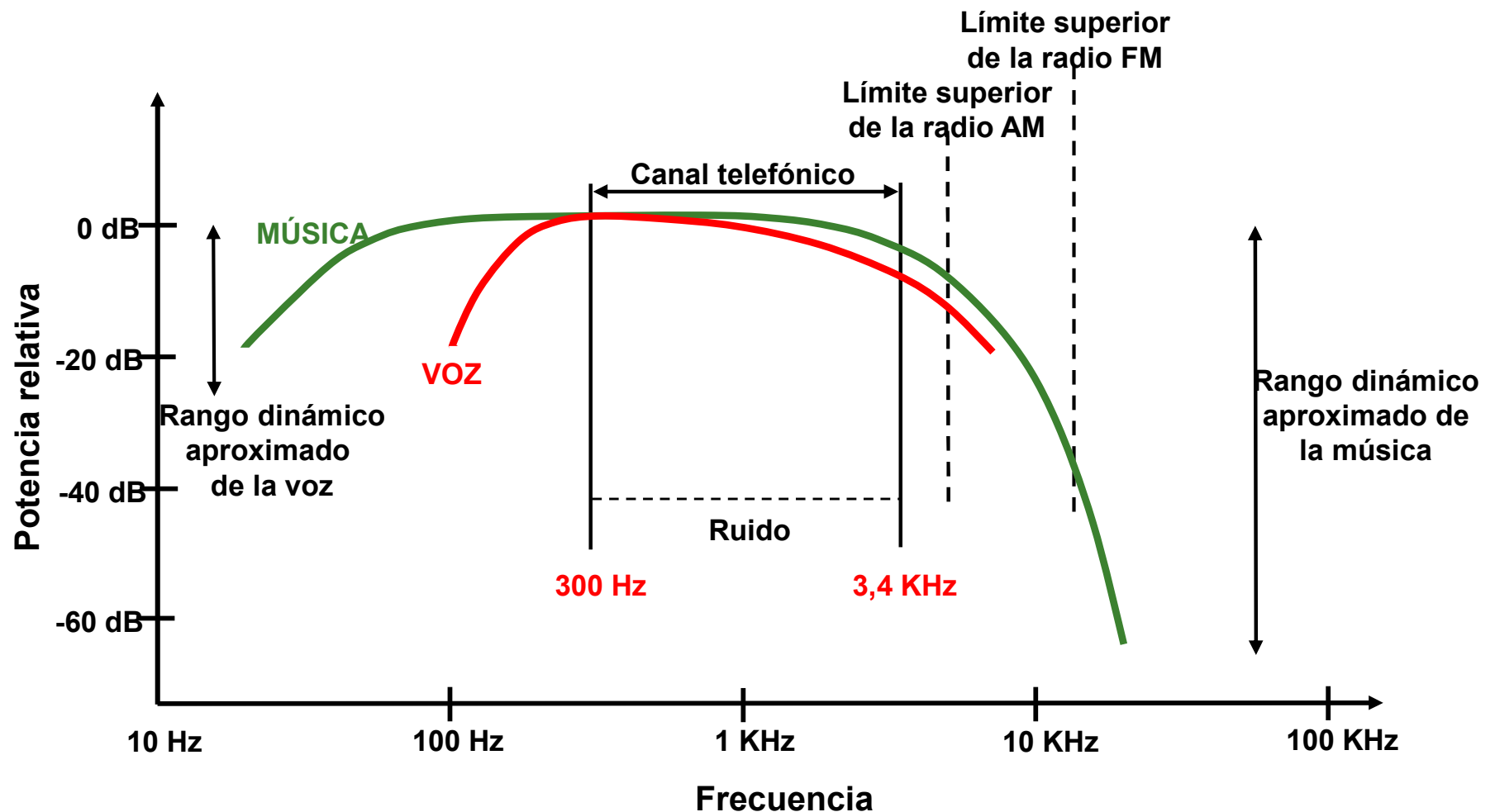
Ing. Anibal Pose

Rango de frecuencias de la voz humana

El espectro de frecuencia de la voz humana se extiende desde aproximadamente 50 Hz a 8 kHz, puede entenderse bien en este ancho de banda tan limitado. Con la digitalización de la voz y las tecnologías actuales el ancho de banda analógico es solo de 300 Hz a 3,4 kHz.

Con la introducción de telefonía de voz sobre IP (VoIP) y la alta capacidad de transmisión de las redes y terminales IP, es posible permitir un mayor espectro de frecuencia en las llamadas telefónicas. El audio de banda ancha aumenta el rango de frecuencia de 50 Hz a 7,0 kHz. Gracias a esto, se pueden transferir los tonos más bajos y más altos cuando se habla por teléfono.

Espectro acústico de la voz y la música



Potencia relativa=Potencia/Potencia máxima

Propiedades

- La codificación de la voz en sistemas digitales se basa en las siguientes propiedades:
 - ❑ Función densidad de probabilidad no uniforme en amplitud
 - ❑ Autocorrelación diferente de cero entre diferentes muestras de voz
 - ❑ Espectro no plano
 - ❑ La existencia de segmentos con voz y con silencio en la señal
 - ❑ Cuasi periodicidad de la señal de voz.

Propiedades

En una señal cuando se transmite, la capacidad que posee para transportar información, o bien viene limitado por la propia señal (que es lo visto anteriormente, una señal con frecuencia máxima f_m) y su ancho de banda, o bien viene limitado por el ancho de banda del canal en la que es transmitida.

En resumen, o el ancho de banda lo fija la fuente o bien el canal.

Propiedades

- Sin embargo, la propiedad principal es que la voz tiene un ancho de banda limitado, por lo tanto se puede digitalizar.
- Función densidad de probabilidad: Una buena aproximación de la pdf de una señal de voz es:

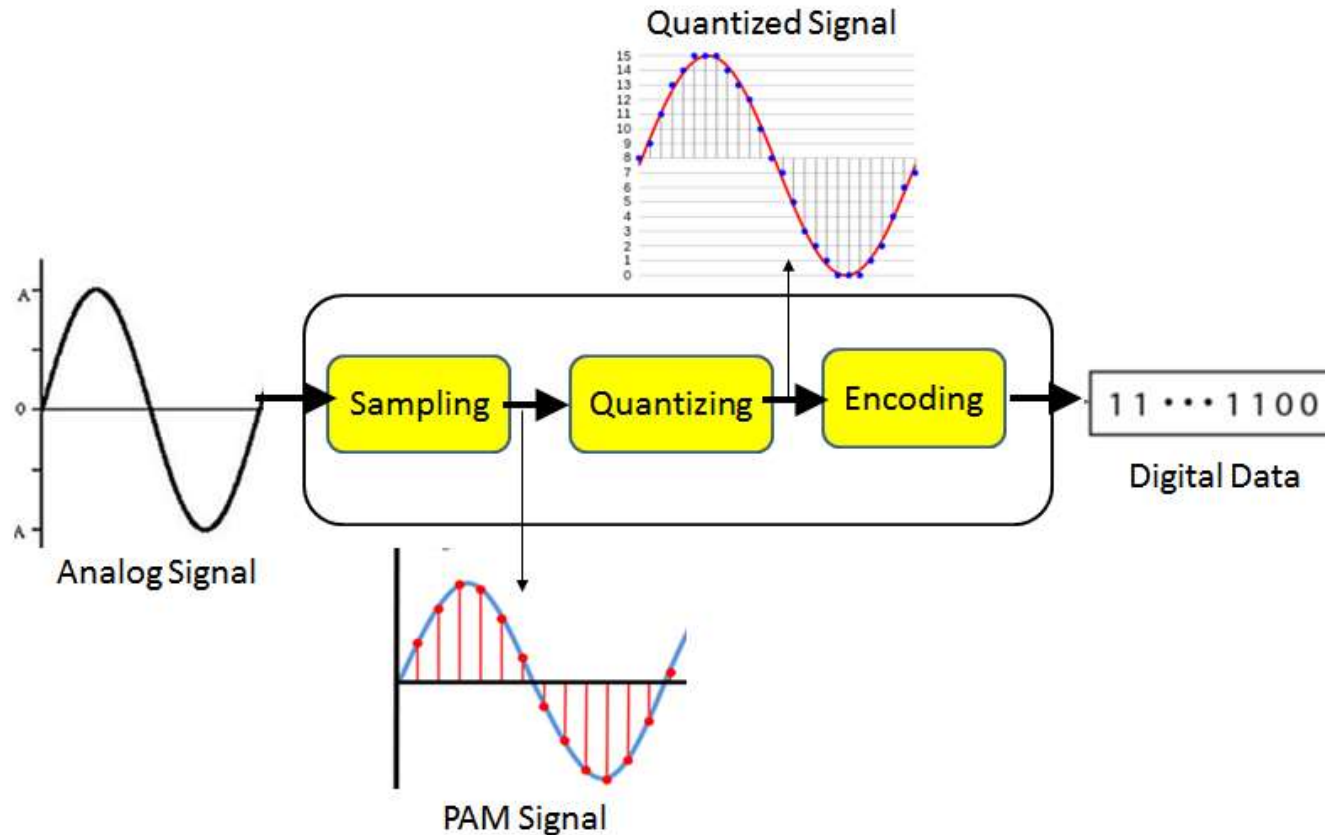
$$p(x) = \frac{1}{\sqrt{2\sigma_x}} \exp(-\sqrt{2}|x| / \sigma_x)$$

Ej. La voz humana, tiene un BW >4KHz, pero los circuitos de las centrales operan hasta 4KHz.

Predictibilidad de la voz

- Función de autocorrelación
- Tal vez la segunda propiedad más empleada por los codificadores es la existencia de una correlación muy alta entre muestras consecutivas de la señal vocal.
- Esto significa que a partir de una muestra, es posible predecir gran parte de la señal que sigue.

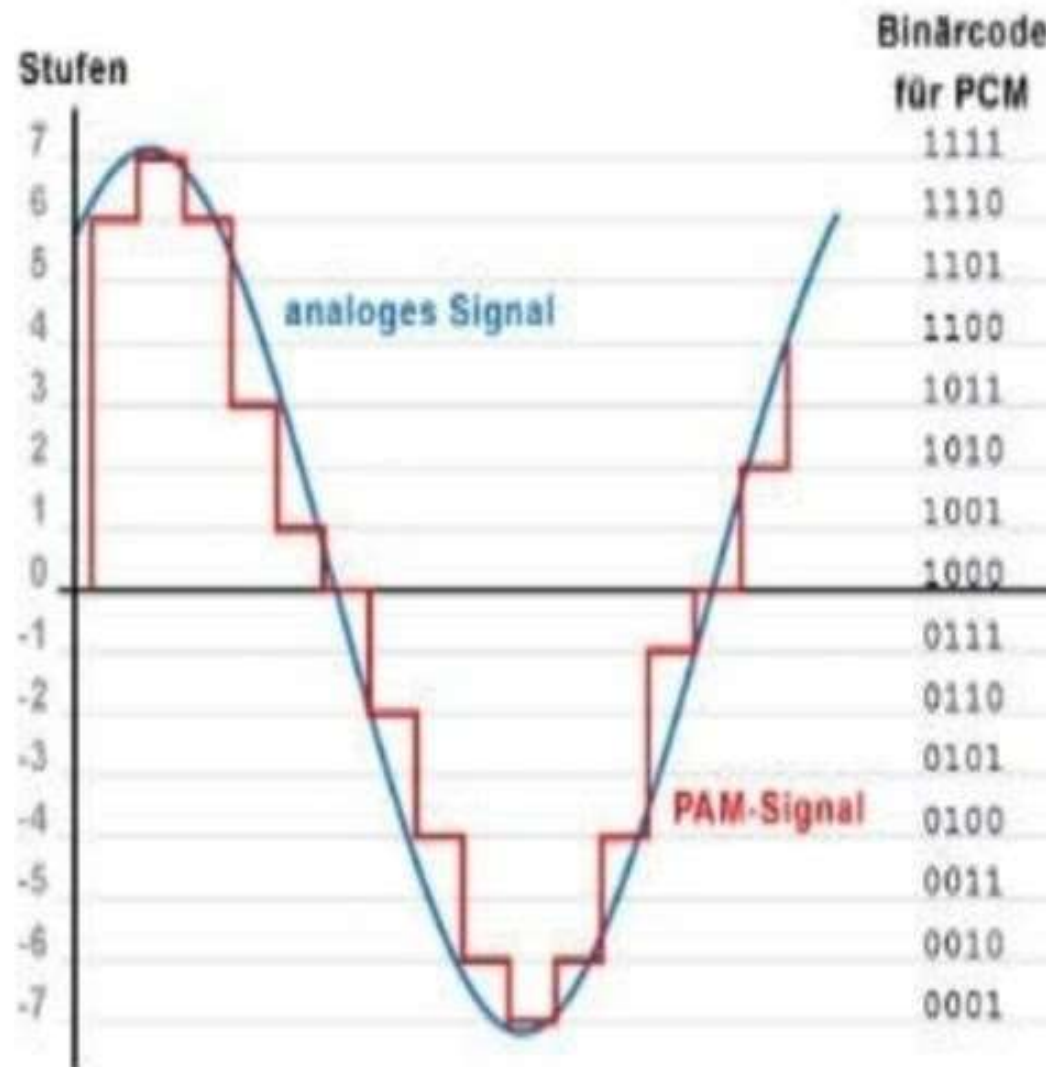
Digitalización de señales analógicas



Procesos básicos de digitalización de señales analógicas

PAM (Pulse amplitud Modulation) y PCM (Pulse Code Modulation)

Digitalización de señales analógicas



Muestreo de señales (sampling)



El muestreo digital es una de las partes del proceso de digitalización de las señales. Consiste en tomar muestras de una señal analógica a una frecuencia o tasa de muestreo constante, para cuantificarlas posteriormente.

El muestreo está basado en el teorema de muestreo, que es la base de la representación discreta de una señal continua en banda limitada. Es útil en la digitalización de señales y en la codificación del sonido en formato digital.



En una señal cuando se transmite, la capacidad que posee para transportar información, o bien viene limitado por la propia señal (que es lo visto anteriormente, una señal con frecuencia máxima f_m) y su ancho de banda, o bien viene limitado por el ancho de banda del canal en la que es transmitida. En resumen, o el ancho de banda lo fija la fuente o bien el canal.

Ej. La voz humana, tiene un $BW > 4\text{KHz}$, pero los circuitos de las centrales operan hasta 4KHz .



Teorema de muestreo de Nyquist-Shanon

Teorema del Muestreo: La frecuencia mínima de muestreo necesaria para evitar el “aliasing” debe ser :

Frecuencia de muestreo $> 2.BW$

Siendo

BW: ancho de banda de la señal a muestrear
($BW = f_{max} - f_{min}$)

Para señales con $f_{min} = 0$, se puede expresar como
Frecuencia de muestreo $> 2.f_{max}$ de la señal

Cuantificación de señales

- Técnicas de cuantificación: La cuantificación es el proceso de convertir un rango continuo de amplitudes, en un grupo finito de valores digitales.
 - ❑ Cuantificación uniforme
 - ❑ Cuantificación no uniforme: Ley $-A$, ley $-\mu$.
 - ❑ Cuantificación adaptativa: Las señales de voz no son estacionarias, por lo tanto se pueden tener cuantificadores adaptativos.
 - ❑ Cuantificación vectorial

Predictibilidad de la voz

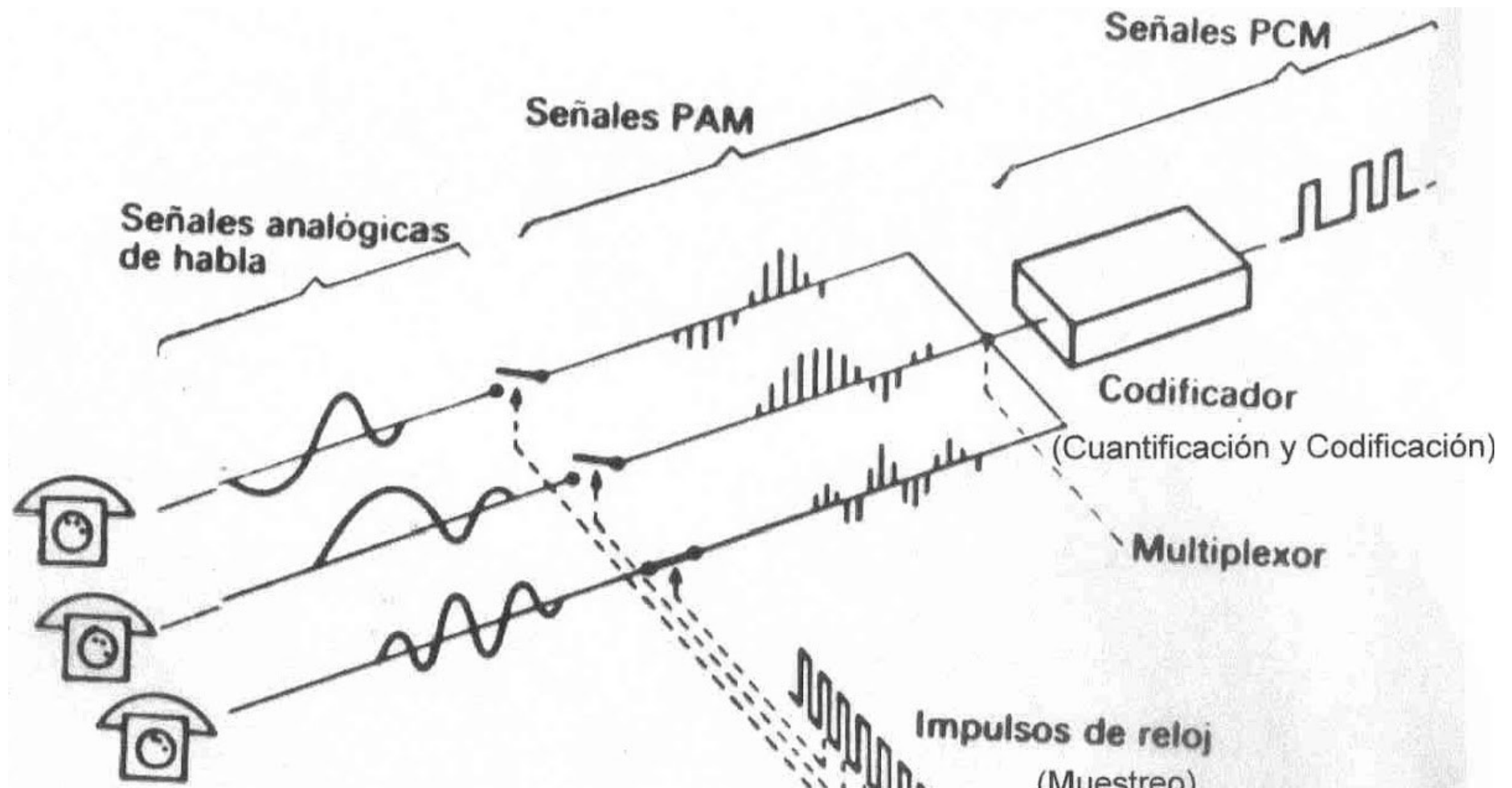
- La función de autocorrelación está definida como:

$$C(k) = \frac{1}{N} \sum_{n=0}^{N-k-1} x(n)x(n+|k|)$$

- $x(k)$ representa la k -ésima muestra de voz..

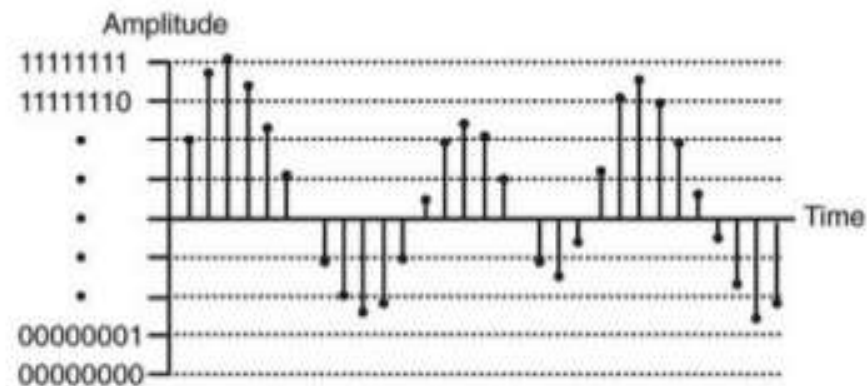


Etapas de codificación





Codificación – G711



Uniformly Quantized PCM Signal

- 8 bits por muestra permiten codificar 256 niveles de señal en la amplitud de la señal de audio
- Tomamos muestras con una frecuencia de 8 Khz, es decir, cada 125 microsegundos

1 Muestra cada 125 microsegundos= 8000 muestras por segundo
 $8000 \text{ muestras} * 8 \text{ bits} = \underline{\underline{64 \text{ Kbps/seg.}}}$



Definición de Codec

Códec es la abreviatura de codificador-decodificador. Describe una especificación desarrollada en software, hardware o una combinación de ambos, capaz de transformar un archivo con un flujo de datos (stream) o una señal.

Los códecs pueden codificar el flujo o la señal (a menudo para la transmisión, el almacenaje o el cifrado) y recuperarlo o descifrarlo del mismo modo para la reproducción o la manipulación en un formato más apropiado para estas operaciones. Los códecs son usados a menudo en videoconferencias y emisiones de medios de comunicación.

Algoritmos de codificación

- Tres técnicas de codificación
 - PCM: codificadores de forma de onda
 - Codifica cada muestra del conversor AD
 - Ej: G.711 (RTC)
 - ADPCM: Adaptativo diferencial PCM
 - Codifica la diferencia entre cada muestra y la anterior
 - Menor rango de codificación \Rightarrow tasas de bit menores
 - LPC: predictivo lineal
 - Codificación por bloques (Representación parametrizada de las características del conjunto)

Tipos codificación

- Codificación en el dominio de la frecuencia
 - Codificación Sub-banda: Se divide la señal de voz en varias sub-bandas (4 u 8), y cada banda se muestrea a la tasa de Nyquist y se codifica con diferente precisión.
 - Codificación por Transformación Adaptativa (ATC): Emplea la Transformada Discreta del Coseno (DCT)

Algoritmos de codificación

- LPC excitado por múltiples pulsos (MPE-LPC):
 - La excitación del codificador con un solo pulso produce distorsión de audio, por lo que este sistema utiliza varios pulsos de excitación por período.
 - Produce mejor calidad que el LPC simple.

Algoritmos de codificación

- LPC excitado por código (CELP):
 - ❑ El codificador y el decodificador tienen un libro de códigos aleatorios predeterminado.
 - ❑ Para cada señal, el transmisor busca el código más adecuado para excitar el LPC.
 - ❑ Son dispositivos muy complejos

Algoritmos de codificación

- LPC con Excitación Residual (RELP):
 - Una vez se calculan los parámetros del dispositivo, se sintetiza la voz en el transmisor y se resta a la señal original para formar una señal residual.
 - La señal residual se cuantiza, codifica y envía, junto con los parámetros del LPC.

Codificadores

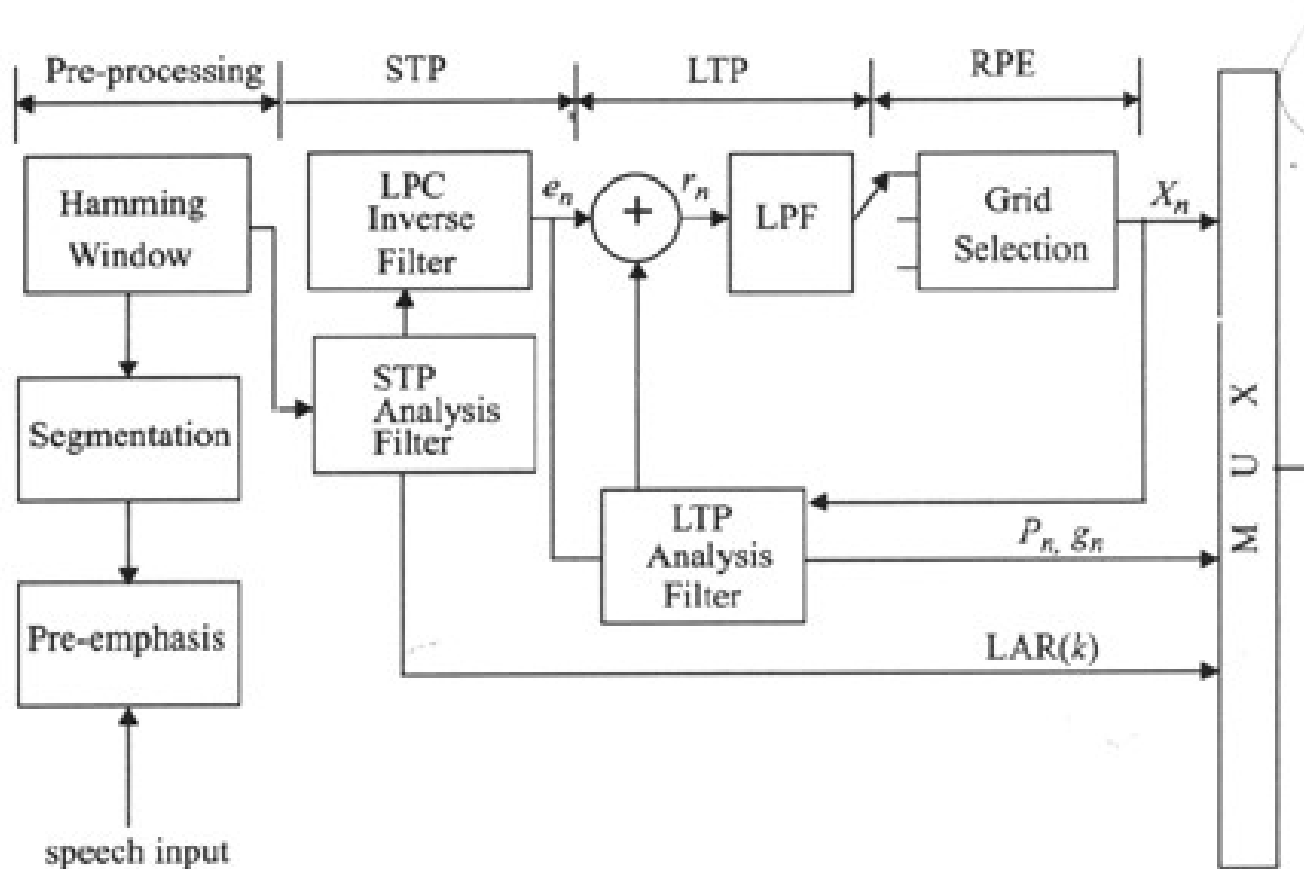
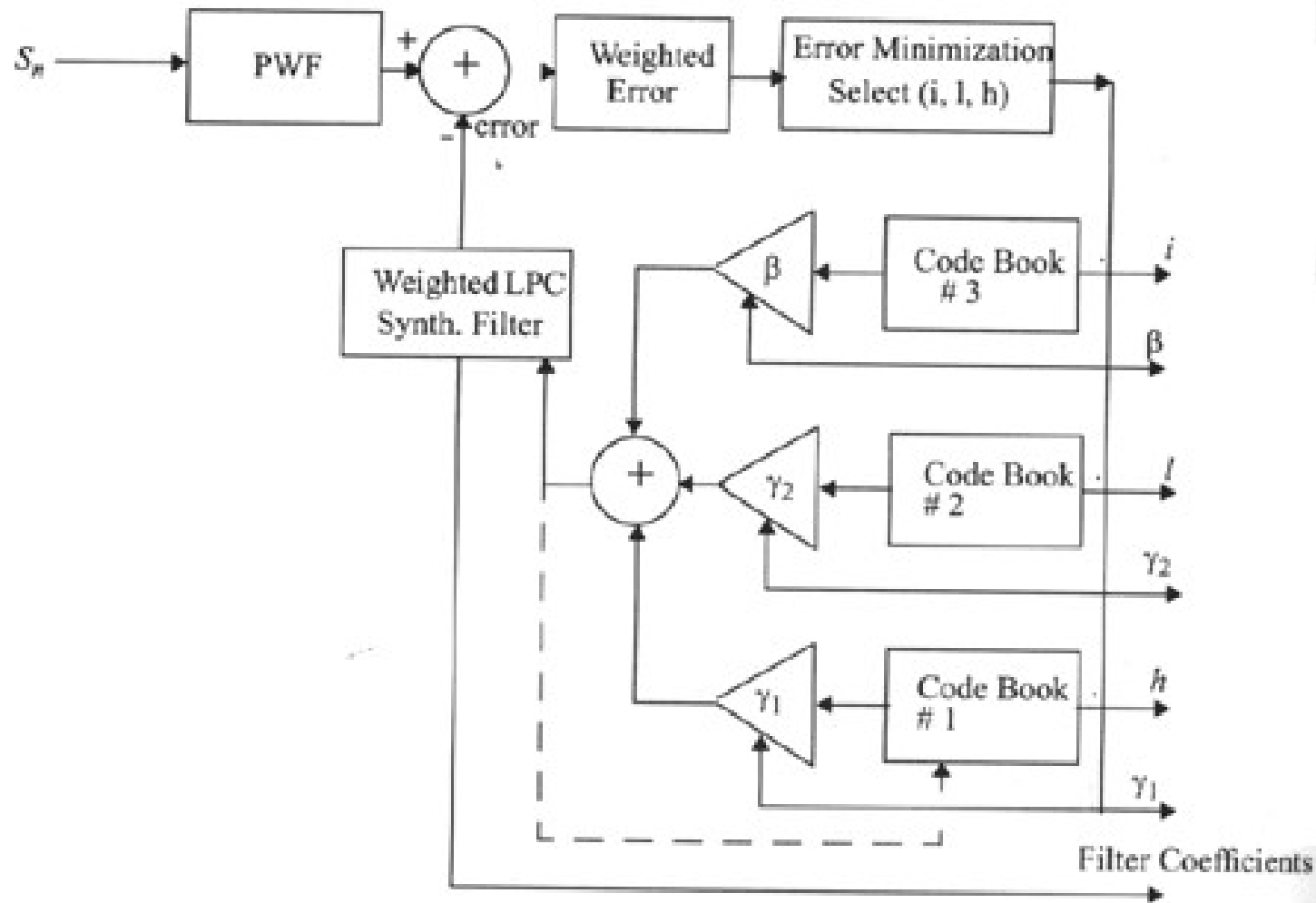


Figure 7.10
Block diagram of GSM speech encoder.

Codificadores



Digitalización de la voz

Codecs



Coding Standard	Algorithm	Data Rate
G.711	PCM (Pulse Code Modulation)	64 kbps
G.726	ADPCM (Adaptive Differential Pulse Code Modulation)	16, 24, 32, 40 kbps
G.728	LD-CELP (Low Delay Code Excited Linear Prediction)	16 kbps
G.729	CS-ACELP (Conjugate Structure Algebraic CELP)	8 kbps
G.723.1	MP-MLQ (Multi-Pulse Maximum Likelihood Quantization)	6.3 kbps 5.3 kbps
	ACELP (Algebraic Code Excited Linear Prediction)	6.3 kbps 5.3 kbps