

23
2 Gen



Universidad Nacional Autónoma de México

FACULTAD DE INGENIERIA

CODIFICACION DIGITAL DE LA
VOZ

T E S I S

Que para obtener el Título de:

ING. EN COMPUTACION

P r e s e n t a

JESUS SAVAGE CARMONA



México, D. F.

Mayo 1985



Universidad Nacional
Autónoma de México



UNAM – Dirección General de Bibliotecas Tesis Digitales Restricciones de uso

DERECHOS RESERVADOS © PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL

Todo el material contenido en esta tesis está protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

INDICE

1. Introducción
 1. Codificación Digital de la Voz
 2. Criterio de fidelidad
2. Codificadores de Forma de Onda
 1. Introducción
 1. Conversión A / D
 1. Prefiltrado
 2. Muestreo
 3. Cuantización
 1. Cuantización Uniforme
 2. Medidas de Ejecución
 3. Cuantización no Uniforme
 4. Cuantización Robusta
 4. Codificación
 2. Codificadores en el dominio del tiempo
 1. PCM
 2. PCM Logarítmica
 3. APCM
 1. Adaptación con realimentación adelantada
 2. Adaptación con retroalimentación
 4. DPCM
 1. Modulación Delta
 5. ADPCM
 3. Codificadores en el dominio de la Frecuencia
 1. SBC
3. Codificadores Fuente (Vocoders)
 1. Introducción
 2. Vocoders en el dominio de la frecuencia
 1. Channel Vocoder

3. Vocoders en el dominio del tiempo
4. Experimentación
5. Resultados

CAPITULO 1
INTRODUCCION

1.1 Codificación Digital de la Voz

Ya que las tecnologías digitales han evolucionado aceleradamente, y los costos de los dispositivos han sido reducidos a lo largo de estos últimos años, un nuevo interés ha sido enfocado en los métodos para la codificación digital y la transmisión de la voz digitalmente.

El objetivo fundamental no es nuevo, es el de transmitir la voz con la más alta calidad posible, y con el menor costo. Pero el intento para acompletar esta tarea a través de métodos digitales nuevos y sofisticados, es nueva, teniendo como premisa el hardware digital.

Típicamente el costo de la codificación de la voz es positivamente correlacionado con la complejidad del codificador y la complejidad, es por su parte, positivamente correlacionada con la eficiencia del código y la utilización del canal.

En otro tiempo las investigaciones relacionadas con la codificación digital habían sido rechazadas, principalmente debido a su alto costo. Pero los avances en la escala de integración de los dispositivos ha cambiado totalmente esta tendencia. El objetivo de esta tesis es el de perfilar un entendimiento y las capacidades que puede tener la "voz digital".

El diseño completo de cualquier sistema de transmisión requiere una selección óptima (en algún sentido) de una combinación de factores tales como calidad de la señal, razón en la transmisión de bits, y costo del codificador.

La selección apropiada depende mucho de el medio ambiente de la transmisión (por ejemplo, alambrado terrestre, fibras ópticas, o radio)

1.2 Criterio de Fidelidad.

Cualquier medida para medir la calidad señal implica una medida de fidelidad. Para la mayoría de los sistemas de comunicación esta medida es difícil de especificar cuantitativamente principalmente por que involucra percepciones humanas.

La calidad de la voz es tradicionalmente medida por el criterio de que la persona que escucha entienda que se esta diciendo y quien lo esta diciendo.

CAPITULO 2.

CODIFICADORES DE FORMA DE ONDA

2.1 Introducción

Una clase amplia de codificadores de voz es llamado codificadores de forma de onda. Como su nombre lo implica, estos codificadores esencialmente se esfuerzan para seguir la forma de onda de la señal.

En un principio ellos son diseñados para ser independientes de la señal, por este motivo ellos pueden codificar bien una variedad de señales-voz musica, tonos.

2.1.1 Conversión Analógica Digital

El procesamiento y la transmisión de señales digitales rápidamente se aproxima a jugar un papel dominante en los sistemas de comunicación.

A pesar, que el origen físico de la mayoría de las señales (voz, imagen, telemetría, sísmica, etc) es intrínsecamente por naturaleza analógica y continua en el tiempo. Por esta razón, una efectiva interface entre el medio ambiente analógico y digital es de crucial importancia en el procesamiento moderno de señales.

La conversión analógica digital puede ser vista como una operación que consta de cuatro etapas: prefiltrado, muestreo, cuantización, y codificación.

2.1.1.1 Prefiltrado -

Ya que una señal consiste de diferentes componentes de frecuencia ,podemos hablar acerca de una porción de la señal ,referida a alguna de sus componentes de frecuencia. Por ejemplo el oído humano puede detectar solo las componentes de frecuencia de audio de una señal ,esto es ,aquellas componentes de frecuencia que estan debafo de los 10,000 Hz.

Para poder extraer de la señal alguna de sus componentes de frecuencia ,consideremos un sistema lineal invariante cuya función de transferencia es

$$H(\omega) = \begin{cases} 1 & -\omega_c < \omega < \omega_c \\ 0 & \text{caso contrario} \end{cases}$$

como se muestra en la siguiente figura :

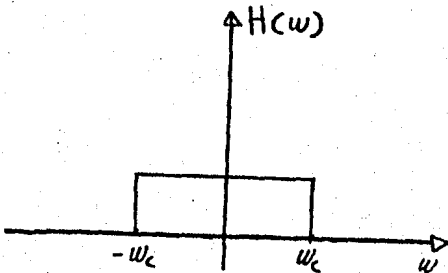


Fig 2.1.1.1.1 Filtro.

Un sistema es llamado filtro si permite el paso de solo de algunas componentes de frecuencia de la señal de entrada y rechaza las otras componentes de frecuencia. El sistema lineal e invariante mostrado en la figura anterior es un filtro. En

particular este filtro es conocido como un filtro paso bajas por que el permite el paso solo de las componentes de baja frecuencia de la señal de entrada.

El ancho de banda para la voz es aproximadamente de 0 a 3200 Hz ,por esta razón para realizar la conversión digital analógica es necesario quitar todas aquellas componentes de frecuencias indeceables que nuestra señal de entrada porta. Por esta razón antes de realizar la conversión digital analógica de nuestra señal es necesario filtrarla con un filtro paso bajas cuyas frecuencias de corte sean de 0 y 3200 Hz

2.1.1.2 Muestreo -

Considerese una señal $f(t)$, que varía continuamente, la que se desea convertir a la forma digital. Esto se logra en primer lugar al muestrear $f(t)$ a una velocidad de F_c muestras por segundo.

Aunque en la practica este procedimiento de muestreo podria presumiblemente llevarse a cabo por medios electrónicos, conmutando la señal entre encendido y apagado a la señal deseada, el proceso de muestreo se muestra en la siguiente figura en forma conceptual, utilizando un interruptor mecánico que gira.

Gira F_c veces por segundo.

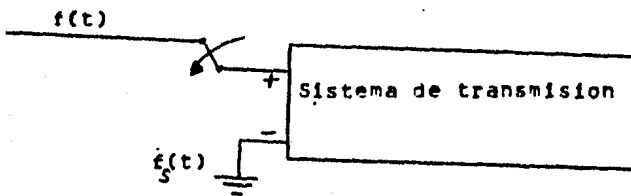


Fig. 2.1.1.2.1 Muestreo de una señal analogica.

Supongase que el interruptor permanece en línea de $f(t)$ durante T segundos cuando se encuentra girando a la velocidad deseada de $F_c = 1/T$ veces por segundo ($T \ll T$). La salida del interruptor $f_s(t)$ es entonces una versión muestreada de $f(t)$.

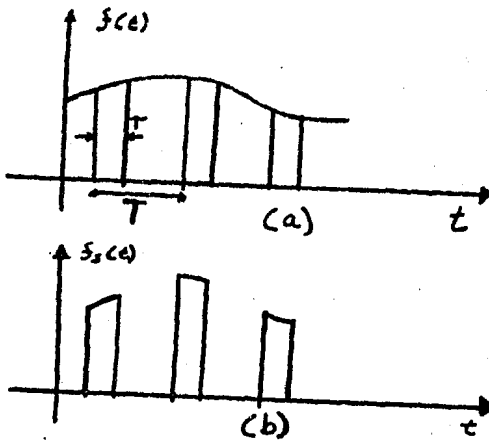


Fig 2.1.1.2.2. Proceso de muestreo ($T =$ tiempo de muestreo)

$T=1/F_c$:intervalo de muestreo.). a) entrada

$f(t)$. b) Salida muestreada $f_s(t)$.)

Una pregunta que surge de inmediato es la siguiente: ¿ Que valor debiera tener la velocidad de muestreo ? ¿ Habra alguna limitación en la velocidad a la velocidad que se pueda muestrear?.

Podría tenerse una idea intuitiva de que el proceso de muestreo distorciona irremediabilmente la señal original $f(t)$. Este proceso se introduce para convertir la señal $f(t)$ a la forma digital y poder después continuar con el procesamiento y la transmisión. Con seguridad y en algun punto lejano ,en la mayoría de los casos ,se deseara recuperar la función $f(t)$ nuevamente. ¿ Se habrá perdido información valiosa en el proceso de muestreo ?.

La respuesta a esta última pregunta ,y también a las demás ,es que con una simple suposición ,la señal muestreada $f_s(t)$

contiene toda la información de $f(t)$. Además a partir de $fs(t)$ se obtiene $f(t)$.

Este resultado sorprendente y no muy obvio, puede demostrarse por medio del uso del análisis de Fourier.

Se supondrá en primer lugar que la señal $f(t)$ es de banda limitada a ω_m hertz, lo cual quiere decir que se encuentra libre de componentes de frecuencia por encima de $f = \omega_m$. La transformada de Fourier $F(\omega)$ de este tipo de señal se muestra en la siguiente figura.

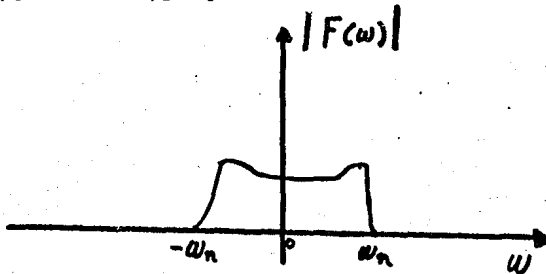


Fig. 2.1.1.2.3 Señal de banda limitada

Las señales físicas que se presentan normalmente no tienen la característica de corte abrupto de frecuencia que se supone en la banda limitada. Con excepción de algunos casos especiales, las señales reales contienen componentes de frecuencia para toda la gama de frecuencias. También se sabe, que el contenido de frecuencia de las señales decae rápidamente después de ancho de banda definido. Esta aproximación de limitación de banda de las señales reales no introduce un error significativo en el análisis y, por lo tanto, es al que se considera en la realidad. En la práctica, se utilizan filtros paso bajo de corte abrupto, los que se introducen con frecuencia antes del proceso de muestreo para asegurar la condición de limitación de banda se

cumpla con la aproximación deseada.

Con la señal $f(t)$ limitada en banda a ω_m hertz, se demuestra fácilmente que muestreando la señal no destruye ningún contenido de información de la misma siempre que la velocidad de muestreo sea $F_c \geq 2\omega_m$. La velocidad mínima de muestreo de $2\omega_m$ veces por segundo se denomina velocidad de muestreo de Nyquist y $1/2\omega_m$ se llama intervalo de muestreo de Nyquist.

El teorema del muestreo se puede probar con la ayuda del teorema de convolución en frecuencia. Considerese el muestreo de $f(t)$ designado como $f_s(t)$. La $f(t)$ que se muestra se puede representar por la multiplicación de $f(t)$ por un tren de impulsos unitarios de período T , igual al intervalo de muestreo. Es decir:

$$f_s(t) = f(t) \cdot \sum_{n=-\infty}^{\infty} \delta(t - nT)$$

Ahora se obtendrá el espectro de $f_s(t)$. Se sabe, por el teorema de convolución en la frecuencia, que $F_s(\omega)$ es la convolución de $F(\omega)$ y la transformada de Fourier de $1/2\pi$ veces el tren de impulsos. Es decir,

$$F_s(\omega) = F(\omega) * \omega_0 \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0)$$

2

donde $\omega_0 = 2\pi/T$. Así, substituyendo por ω_0 , se obtiene el espectro de los valores muestreados como:

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} (F(\omega) * \delta(\omega - n\omega_0))$$

$$= \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_0)$$

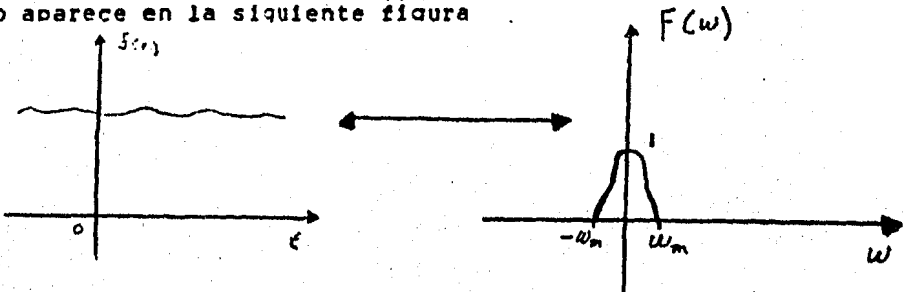
El miembro derecho de la expresión anterior representa la función $F(\omega)$ repetida cada ω_0 rad/seg. Si el ancho de banda de las frecuencias distintas de cero en $F(\omega)$ es menor que entre las repeticiones de $F(\omega)$, entonces dichas repeticiones no se traslaparan. Es decir, $F(\omega)$ se repetirá periódicamente en el dominio de la frecuencia sin traslaparse, suponiendo que $\omega_0 > 2\omega_m$, y esto implica que

$$\frac{2\pi}{T} \geq 2(2\pi f_m)$$

o

$$T \leq \frac{1}{2f_m}$$

Mientras el muestreo de $f(t)$ se realice a intervalos no mayores de $1/2f_m$ segundos, $F_s(\omega)$ será la réplica periódica de $F(\omega)$. Este resultado también se puede demostrar gráficamente como aparece en la siguiente figura



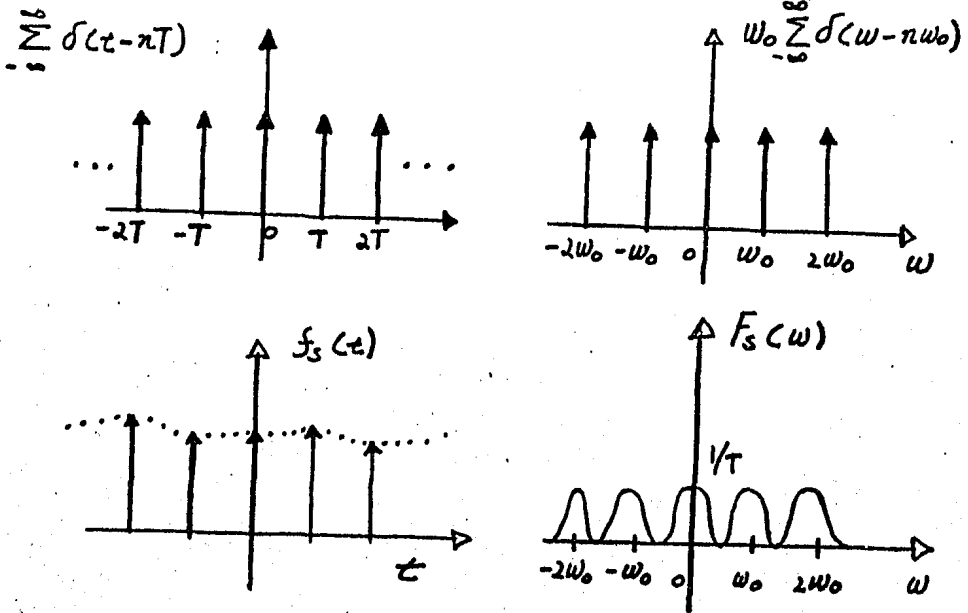


Fig 2.1.1.2.4 Muestreo

El espectro de $F_s(\omega)$ es la convolución del tren de impulsos

∞

$$\left(\frac{\omega_0}{2\pi}\right) \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0) \text{ y } F(\omega)$$

$-\infty$

La imagen espejo es igual a la función original por que el tren de impulsos $(1/T) \sum_{n=-\infty}^{\infty} \delta(\omega - n\omega_0)$ es función par de ω . Para realizar la convolución, se desplaza el tren de impulsos sobre $F(\omega)$. Los impulsos están separados una distancia ω_0 . Así, las repeticiones de $F(\omega)$ que forman $F_s(\omega)$ también están separadas una distancia ω_0 rad/seg. De esta forma se obtiene el espectro de $F_s(\omega)$ que se ilustra en la figura anterior. Hay que observar que el espectro repetido está multiplicado por $1/T$.

$F(\omega)$ (y por lo tanto $f(t)$) se puede obtener fácilmente de $F_s(\omega)$ anulando mediante filtración todas las componentes de frecuencias abajo de ω_m rad/seg. Es posible lograr esto si se emplea un filtro paso baja que permita la transmisión de todas las frecuencias arriba de ω_m .

El intervalo $T=1/2f_m$ se conoce como intervalo de Nyquist. Si $f(t)$ se muestrea a intervalos mayores de $1/2f_m$ seg, la repetición de $F(\omega)$ que forma $F_s(\omega)$ se trasladara y al recobrar $F(\omega)$ de $F_s(\omega)$ se tendra cierto error. Por supuesto, esto es lo que nos indicaría la intuición. Si las muestras estuvieran demasiado espaciadas, es lógico que la señal $f(t)$ no se pueda recuperar a partir de sus muestras.

El proceso de recobrar $f(t)$ a partir de su versión muestreada $f_s(t)$ se realiza pasando $f_s(t)$ a través de un filtro paso bajo. Esto se puede demostrar matemáticamente si se utiliza el teorema de convolución en el tiempo. Sea $T=1/2f_m$ el espaciamiento entre las muestras y por lo tanto $\omega_0 = 2\pi/T = 4\pi f_m = 2\omega_m$. El espectro de $f_s(t)$ esta dado por

$$F_s(\omega) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - n\omega_0) = \frac{1}{T} \sum_{n=-\infty}^{\infty} F(\omega - 2n\omega_m)$$

El proceso de filtrar equivale a multiplicar $F_s(\omega)$ por una función de ω que es uno para $\omega < |\omega_m|$ y 0 para otros casos; es decir una función pulso rectangular $G_{2\omega_m}(\omega)$

$$G_{2\omega_m}(\omega) = \begin{cases} 1, & \omega < |\omega_m| \\ 0, & \omega > |\omega_m| \end{cases}$$

asi

$$\frac{F(\omega) = F_s(\omega)G_{2\omega m}(\omega)}{T}$$

o'

$$F(\omega) = TF_s(\omega)G_{2\omega m}(\omega)$$

Aplicando a la expresión anterior el teorema de convolución en el tiempo se llega a

$$f(t) = Tfs(t) * \frac{\omega m}{\pi} \text{sinc } \omega m t$$

$$= fs(t) * \text{sinc } \omega m t$$

$$\begin{aligned} & \overset{\infty}{=} \sum_{n=-\infty}^{\infty} f(nT) \delta(t - nT) * \text{sinc } \omega m t \end{aligned}$$

$$\begin{aligned} & \overset{\infty}{=} \sum_{n=-\infty}^{\infty} f(nT) \text{sinc } (\omega m(t - nT)) \end{aligned}$$

Por lo general, la función sinc $[\omega m(t - nT)]$ recibe el nombre de función de interpolación, debido a que con ella se pueden interpolar los valores muestra $f(nT)$ para obtener la $f(t)$ para toda t . En la siguiente figura se muestra este resultado:

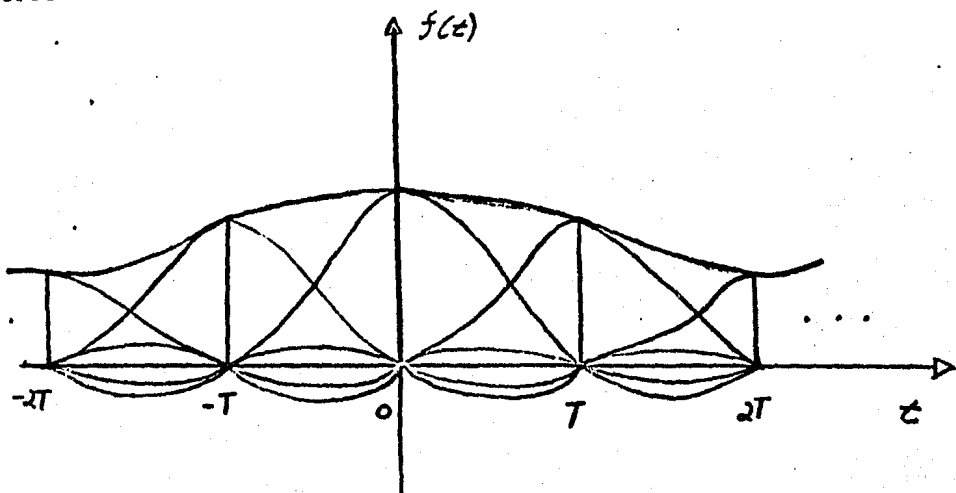


Fig 2.1.1.2.5. Funcion de interpolacion.

Cada valor muestra esta multiplicado por la función sinc centrada en el valor muestreado. Entonces, estas funciones se suman para obtener la forma de onda original.

2.1.1.3 Cuantización -

La cuantización empieza cuando se cuenta las muestras analógicas. Cada muestra puede ser tomada en general sobre cualquier valor de amplitud, desde menos infinito a infinito.

El cuantizador reemplaza cada uno de estos valores con un valor de salida el cual es una aproximación de la amplitud original. La principal característica de cada valor de salida, es que este número pertenece al conjunto de los números reales. Una palabra binaria (n -bits) puede estar asociada a cada valor de salida si el conjunto de valores de salida no contiene más de 2 al n miembros. Con este procedimiento, una secuencia de muestras analógicas puede ser transformada en una secuencia de palabras binarias apropiadas para ser guardadas, transmitidos o alguna otra forma de procesamiento digital.

Un receptor teniendo una tabla de los valores de salida asociado con el conjunto de palabras binarias, puede entonces reconstruir una aproximación a la secuencia original de muestras. El proceso de reconstrucción es llamado conversión digital analógica (D/A).

2.1.1.3.1 Cuantización Uniforme -

La forma más simple y más común de cuantizar es el cuantizador de memoria cero. En este caso el valor de salida es determinado por el cuantizador solamente con la muestra de entrada.

Un cuantizador de memoria cero de N puntos, puede ser definido especificando un conjunto de $N+1$ niveles de decisión X_0, X_1, \dots, X_n y un conjunto de N puntos de salida Y_1, Y_2, \dots, Y_n .

Cuando el valor de la muestra de entrada cae en el i ésimo intervalo llamado:

$$R_i = (X_{i-1} < X < X_i)$$

el cuantizador produce un valor de salida Y_i .

Ya que Y_i es utilizada para aproximar muestras contenidas en el intervalo en el intervalo R_i , Y_i es escogida de tal modo que sea un valor en el intervalo R_i . Los niveles X_0 y X_n son escogidos igual al más pequeño y al más grande valor, respectivamente, que las muestras de entrada puedan tener. Los N puntos de salida siempre tienen valores finitos. Si $N=2^n$, una palabra binaria única puede estar asociada con cada punto de salida, obteniéndose un cuantizador de n -bit.

La función $Q(x)$ entrada-salida de un cuantizador tiene una forma de escalera, como lo muestra la siguiente figura.

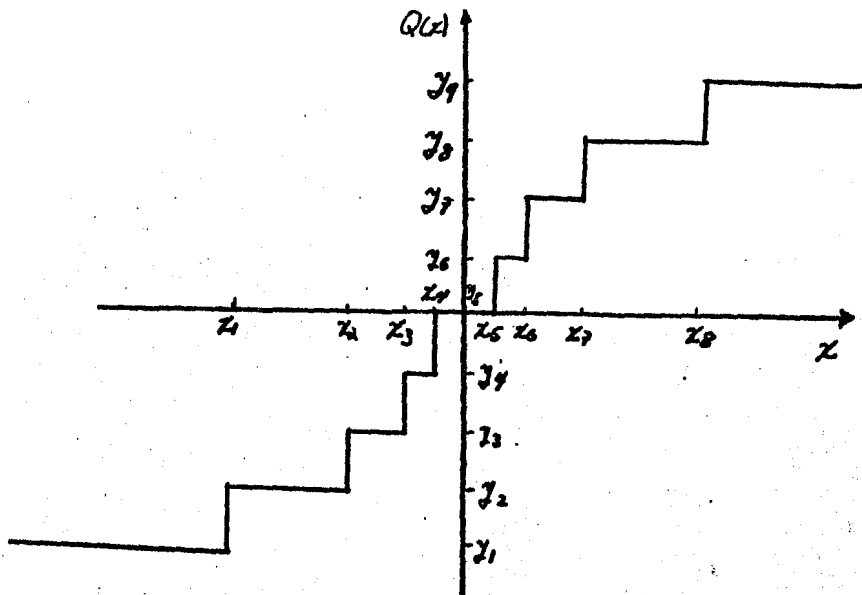


Fig 2.1.1.3.1 Característica de entrada salida de un Cuantizador con $N=9$

Un cuantizador es simplemente una no linealidad sin memoria, cuya característica puede ser vista como una aproximación escalera a la operación identidad $Y=X$.

Cuando la muestra de entrada se encuentra en las regiones R_1 o R_n , se dice que el cuantizador se encuentra en la región de sobrecarga. Todos los demás intervalos de R_i son de tamaño finito.

Es fundamental para un estudio analítico, de la cuantización el reconocer que las muestras de entrada pueden ser vistas como una función aleatoria. Las muestras de entrada no son conocidas y por esto pueden ser observadas como portadoras.

de información.

Se define $P(x)$ como la función de densidad de probabilidad (PDF) de cada muestra de entrada a el cuantizador. Asumir por conveniencia que el valor medio de las muestras de entrada es cero y que $P(x)$ tiene simetría alrededor de cero, con la suposición de simetría la función del cuantizador $Q(x)$ es ordinariamente escogido a tener simetría impar. El proceso de cuantización puede ser modelado como al adición de una componente de ruido $e=Q(x)-x$ Como se indica en la siguiente figura:

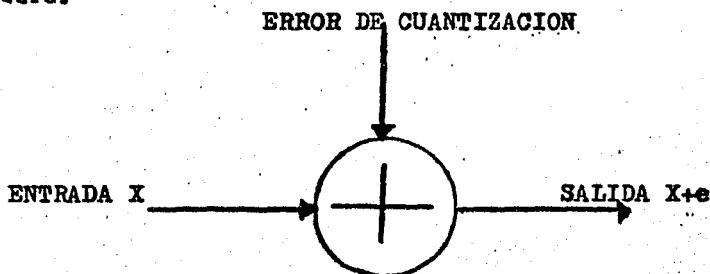


Fig 2.1.1.3.2. Adición de ruido

El ruido es dependiente sobre la amplitud de la señal. El ruido de cuantización puede ser visto como la respuesta de un sistema no lineal cuando la muestra es aplicada a su entrada.

$$U(x)=Q(x)-x$$

como se muestra en la siguiente figura.

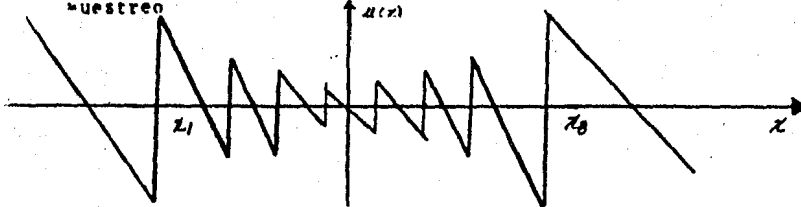


Fig 2.1.1.3.3. Error de cuantización

Quando la muestra de entrada cae dentro del intervalo $x_i < x < x_{i+1}$ el ruido de salida es descrito como ruido granular y es limitado en magnitud. Cuando la entrada cae fuera de ese intervalo la salida es descrita como ruido de sobrecarga y la amplitud es no limitada. Un diseño efectivo de un cuantizador debera tomar en cuenta la función de densidad de probabilidad de las muestras de entrada.

En particular para un número fijo de niveles L la apropiada selección de los niveles x_i y x_{i+1} controla la cantidad de ruido de sobrecarga y granular.

2.1.1.3.2 Medidas de Efectividad -

Ya que el error de cuantización es modelado como una variable aleatoria, una medida de efectividad de un cuantizador debera estar basada en alguna función de error. La más común es la medida de distorsión media cuadrada D :

$$D = \int_{-\infty}^{\infty} (q(x) - x)^2 p(x) dx \dots \dots \dots (1)$$

Frecuentemente es mas útil describir la efectividad de un cuantizador por medio de la razón de señal a ruido (SNR).

Frecuentemente definida como :

$$SNR = 10 \log (\overline{V^2} / D) \dots \dots \dots (2)$$

donde $\overline{V^2}$ es la variancia de la muestra de entrada. En la mayoría de las aplicaciones de cuantización el número N es muy grande, así que una SNR grande es obtenida. Una fórmula útil para el error cuadrático medio puede entonces utilizarse. Así la ecuación (1) puede ser escrita de la forma:

$$D = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (Y_i - X)^2 p(X) dx$$

Haciendo la integración en intervalos separados R_i y notando que $0(X) = Y_i$ cuando X esta en R_i . Para una N grande cada intervalo R_i puede hacerse bastante pequeño, con la excepción de los intervalos de sobrecarga R_1 y R_n los cuales no estan limitados. Entonces es razonable aproximar la densidad de probabilidad $p(X)$ constante dentro del intervalo R_i y aproximando $p(x)=0$ para X en las regiones de sobrecarga, entonces:

$$Y_i = \frac{x_i + x_{i-1}}{2}$$

$$p(X) = p(Y_i) = cte$$

$$D = \sum_{i=1}^N \int_{x_{i-1}}^{x_i} (Y_i - X)^2 p(X) dx =$$

$$\sum_{i=1}^N -p(x) \int_{x_{i-1}}^{x_i} (y_i - x)^2 \sigma(y_i - x) =$$

$$\sum_{i=1}^N -p(x) \frac{1}{3} (y_i - x)^3 \Big|_{x_{i-1}}^{x_i} =$$

$$\frac{-1}{3} \sum_{i=1}^N p(x) \frac{(x_i + x_{i-1} - x)^3}{2} \Big|_{x_{i-1}}^{x_i}$$

$$-1 \sum_{i=1}^N p(x) \left(\frac{(x_i + x_{i-1} - 2x)^3}{2} - \frac{(x_{i-1} + x_{i-1} - 2x)^3}{2} \right)$$

$$-1 \sum_{i=1}^N p(x) \left(\frac{(-x + x_{i-1})^3}{2} - \frac{(x + x_{i-1})^3}{2} \right)$$

$$= -1 \sum_{i=1}^N p(x) \left(\frac{(-\Delta x)^3}{2} - \frac{(\Delta x)^3}{2} \right)$$

$$\begin{aligned}
 &= \frac{-1}{3} \sum_{i=1}^N p(x) \left(\frac{-\Delta x}{8} - \frac{\Delta x}{8} \right) \\
 &= \frac{1}{3} \sum_{i=1}^N p(x) \frac{\Delta x}{8} \\
 &= \frac{1}{12} \sum_{i=2}^{N-1} p(y_i) \Delta_i^3 \\
 D &= \frac{1}{12} \sum_{i=2}^{N-1} p(y_i) \Delta_i^3
 \end{aligned}$$

donde $\Delta_i = x_i - x_{i-1}$ es el tamaño del intervalo R_i .

En el caso especial de una cuantización uniforme en la que los niveles de decisión están igualmente espaciados, así que los intervalos R_i son de tamaño constante $\Delta_i = \Delta$. Con lo que el error puede simplificarse:

$$D = \frac{\Delta^2}{12} \sum_{i=2}^{N-1} p(y_i) \Delta$$

Pero
asi $\sum p(y_i) \Delta \approx \int p(s) ds = 1$

$$D = \frac{\Delta^2}{12}$$

El error rms es entonces $\Delta / \sqrt{12} = \Delta / 2\sqrt{3}$ volts ,lo cual representa el ruido rms que se observa en la salida del sistema.

La relación señal a ruido de cuantización (snr) puede definirse de dos maneras ; Por una parte se definirá una snr en terminos de la señal pico ,v volts ,v por la otra se definirá una segunda snr en terminos de la potencia promedio de la señal. Como $v = \Delta\pi / 2$ es la excursión pico de la señal ,la relación entre el voltaje máximo de la señal v el ruido rms sera de :

$$\frac{Sov}{D} = \frac{v}{\Delta/2\sqrt{3}} = \sqrt{3} n$$

donde Sov = señal y D = ruido

La relación de potencia correspondiente

$$\left(\frac{S_o}{D} \right)^2 = 3 N^2$$

o en decibeles

$$\left(\frac{S_o}{D} \right)_{db} = 4.8 + 20 \log N$$

La relación de potencia crece rapidamente con el cuadrado de el número de niveles. El mejoramiento que experimenta el snr en decibeles con N aparece en la siguiente tabla :

MEJORAMIENTO DE LA SNR DE CUANTIZACION CON EL NUMERO DE NIVELES

So/D db	N	Ancho de banda relativo
11	2	1
17	4	2
23	8	3
29	16	4
35	16	5
41	64	6
47	128	7

También se indica el ancho de banda relativo que se obtiene a partir del análisis siguiente:

Como N es el número de niveles empleados, determina el número de pulsos con los cuales se codifica la señal cuantizada antes de ser transmitida, al aumentar N se incrementa el número de pulsos de codificación y por lo tanto el ancho de banda.

Es posible entonces relacionar el ancho de banda con la snr, lo cual se realiza considerando que $N = m \cdot n$, donde n es el número de pulsos del grupo de codificación y m es el número de niveles de codificación.

Con esta relación las ecuaciones se transforman respectivamente en:

$$\left(\frac{S_0}{D}\right)^2 = 3 \cdot m^{2 \cdot n}$$

$$y \left(\frac{S_0}{D}\right)_{db}^2 = 4.8 + 20 n \log m$$

En particular para un código binario ($m = 2$)

$$\left(\frac{S_0}{D}\right)_{db}^2 = 4.8 + 6 n$$

2.1.1.3.3 Espaciamiento no uniforme - Cuantización uniforme no es en general, la forma más efectiva de obtener una buena realización. Para un número dado de intervalos de cuantización, tomando en cuenta la densidad de probabilidad en la entrada, el espaciamiento no uniforme de los niveles de decisión puede llevarnos a un ruido de cuantización menor al que tuviéramos si fuera espaciamiento uniforme, también nos lleva a una menor sensibilidad a las variaciones de la estadística de la señal de entrada. Se supone entonces el siguiente modelo :

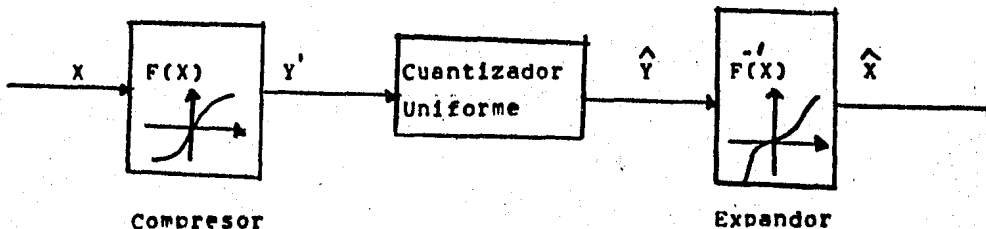


Fig 2.1.1.3.3.1 Cuantización no lineal. es el modelo de un cuantizador como una función $F(x)$ no lineal, el compresor

seguido por un cuantizador uniforme como se muestra en la figura anterior.

La no linealidad extiende los valores de las muestras de amplitudes bajas sobre un nivel mas grande ,mientras que la región de las amplitudes grandes la acorta.

El efecto es el de colocar mas niveles de cuantización a las amplitudes bajas ,las que generalmente tienen una probabilidad mas alta de ocurrir , y menos niveles de cuantización a las amplitudes altas ,que son las que ocurren con menos frecuencia.

La señal comprimida es cuantizada uniformemente y los valores de salida son la entrada a la función no lineal inversa

-1

$F(x)$

produciendo una aproximación de la señal original aplicada al compresor.

El esquema anterior recibe el nombre en ingles de "Companding" ,que es término obtenido de las palabras inglesas "compressing" y "expanding".

La función $F(x)$ es una función monótona creciente ,teniendo simetría impar ,tomando valores desde desde $-V$ a V y con $F(V)= V$ y $F(0)= 0$

Esta operación siendo no lineal y monótona es completamente invertible. Esto es ,una muestra de entrada X podría ser recuperada aplicando el valor $Y = F(X)$ a la no linealidad inversa del expander $F^{-1}(Y)$ y obtener X de nuevo; por este motivo no hay pérdida debido a la operación no lineal.

Considerar una señal cuantizada $Y(t)$ como se muestra en la siguiente figura.

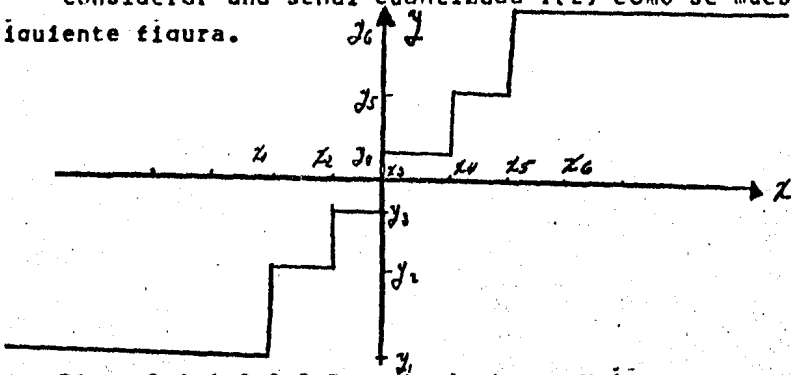


Fig. 2.1.1.3.3.2 Espaciamento no Uniforme

Si los niveles se definen de una forma simétrica con respecto al nivel de voltage cero, pero puestas en una manera arbitraria en el intervalo $(-V, V)$, el problema de minimizar la distorsión de la señal cuantizada por medio de un espaciamento no uniforme, se considerara a continuación.

Denotando los niveles por $Y_{-n}, Y_{-n+1}, \dots, Y_0, Y_1, \dots, Y_{n-1}$ donde $Y_k = Y_{-k}$ y $Y_0 = 0$. Además, asumir que un valor de la señal Y que satisface :

$$Y_{k-1/2} < Y < Y_{k+1/2} \dots \dots \dots (1)$$

es transmitido como nivel Y_k

La medida de distorsión que es adoptada, es la distorsión media cuadrada definida como:

$$F_k = \int_{Y_{k-1/2}}^{Y_k} (Y - Y_k)^2 p(Y) dY \quad \dots\dots\dots(2)$$

donde F_k es la distorsión en el késimo nivel. El error de la señal transmitida es $(Y - Y_k)$ y $P(Y)$ la densidad de probabilidad de la señal Y .

Primero derivar una relación entre las diferentes Y 's en (2) así que F_k en késimo nivel es mínima.

Suponer que los niveles están tan cercanos que la probabilidad $P(Y)$ puede ser considerada como una constante sobre la región de integración e igual a $P(Y_{av})$, donde :

$$Y_{av} = Y_{k+1/2} + Y_{k-1/2}$$

$$F_k = \frac{P(Y_{av})}{3} [(Y_{k+1/2} - Y_k)^3 + (Y_k - Y_{k-1/2})^3] \quad \dots\dots\dots(3)$$

diferenciando F_k con respecto a Y_k se obtiene :

$$\frac{d F_k}{d Y_k} = P(Y_{av}) [(Y_{k+1/2} - Y_k)^2 + (Y_k - Y_{k-1/2})^2] = 0$$

o

$$Y_{k+1/2} - Y_{k-1/2} = Y_{av}$$

entonces

$$\lambda = 2 U_k^2 = 2 U_{k-1}^2 = 2 U_{k-2}^2 = \dots = 2 U_{k-n}^2$$

Por lo tanto (9) es mínimo cuando

$$\mu_{-n} = \mu_{-n+1} = \dots = \mu_{n-1} = \frac{K}{2N+1} = \mu_k$$

de esto sigue :

$$\frac{1}{3} \Delta(Y_k) \Delta Y_k = \frac{K}{2n+1}$$

o

$$F_k = \frac{2}{3} \frac{K^3}{(2n+1)^3}$$

la potencia de la distorsión total mínima

$$F_m = \frac{2}{3} \frac{K^3}{(2n+1)^3}$$

$$= \frac{1}{12(2n+1)^2} \left(\int_{-V}^V p^{1/3}(Y) dY \right)^3$$

Por esto la condición para hacer la \overline{E}_k mínima es que Y_k caiga en medio de $Y_{k+1/2}$ y $Y_{k-1/2}$

$$Y_{k+1/2} = Y_k + \Delta Y_k \dots\dots\dots(5)$$

$$Y_{k-1/2} = Y_k - \Delta Y_k$$

sustituyendo esos valores en tres

$$\overline{E}_k = \frac{2}{3} p(Y_k) \Delta Y_k^3 \dots\dots\dots(6)$$

$$= \frac{2}{3} \sum_{-n}^n p(Y_k) \Delta Y_k^3 \dots\dots\dots(7)$$

El total de la distorsión media cuadrada \overline{E}_a es mínimo cuando \overline{E}_k es constante, independiente de késimo nivel.

Por la definición de integral, podemos escribir :

$$2 \sum_{-n}^n p(Y_k) \Delta Y_k^3 =$$

$$= \int_{-V}^V p(Y) dY = 2k \dots\dots\dots(8)$$

donde k es una constante, va que la integral es una función de solo sus límites sea $\mathcal{L}_{k=0}(Y_k)^{1/3} Y_k$ entonces (7) y (8) se hacen respectivamente

$$\sqrt[3]{d} = \frac{2}{3} \sum_{-n}^n U_k^3$$

y

$$k = \sum_{-n}^n U_k$$

El problema ahora se reduce a minimizar la suma de los cubos sujetos a la condición que la suma de las variables es una constante. Utilizando el método de Lagrange de los multiplicadores indeterminados:

$$\phi = \frac{2}{3} \sum_{-n}^n U_k^3 + \lambda \left(k - \sum_{-n}^n U_k \right)$$

$$\phi_k = 2U_k^2 - \lambda = 0$$

$$\phi_{k-1} = 2U_{k-1}^2 - \lambda = 0$$

⋮

$$\phi_{k-n} = 2U_{k-n}^2 - \lambda = 0$$

Ya $P(Y)$ es una función par v haciendo $N=2n+1$ el número total de niveles:

$$E_m = \frac{2}{3N} \left(\int_0^v P^{1/3}(Y) dY \right)^3$$

La razón de la distorsión media cuadrada del valor medio cuadrado de la señal es:

$$D_m = \frac{E_m^2}{E^2}$$

$$= \frac{2}{3n} \frac{\left[\int_0^v P^{1/3}(Y) dY \right]^3}{\int_0^v Y^2 P(Y) dY}$$

$$= \frac{2}{3N} \int_0^v \frac{P(Y) dY}{q(Y) d} \dots\dots(9)$$

La ecuación (9) da la distorsión mínima es el resultado de un espaciamiento óptimo.

Un método aproximado para la obtención de los niveles puede ser obtenido escribiendo:

$$2K = \sum_{-n}^n p^{1/3}(Y_k) \Delta Y_k$$

$$= P^{1/3}(Y_0) \Delta Y_0 + P^{1/3}(Y_1) \Delta Y_1 + \dots + P^{1/3}(Y_n) \Delta Y_n$$

$$Y = \Delta Y_0 + 2\Delta Y_1 + \dots + 2\Delta Y_{k-1} + \Delta Y_k$$

$$= K \left[\frac{1}{1/3} + \frac{2}{1/3} + \dots + \frac{2}{1/3} + \frac{1}{1/3} \right]$$

$$N \quad P(Y_0) \quad P(Y_1) \quad P(Y_{k-1}) \quad P(Y_k)$$

$$= K \left[\frac{1}{1/3} + \frac{2}{1/3} + \dots + \frac{2}{1/3} + \frac{1}{1/3} \right] \frac{2V}{N}$$

$$2V \quad P(Y_0) \quad P(Y_1) \quad P(Y_{k-1}) \quad P(Y_k) \quad N$$

las series pueden ser aproximadas a una integral

$$Y = A \int_0^Z \frac{1}{P(z)} dz$$

donde se ha cambiado la variable en la derecha para evitar confusión. "A" es una constante de proporcionalidad escogida de tal modo que cuando Z=V; Y=V. Así:

$$Y = \int_0^z \frac{1}{P(z)} dz$$

$$Y = \int_0^z \frac{1}{p(z)} dz$$

Variando Z de 0 a V , Y describira una curva como se muestra en la siguiente figura

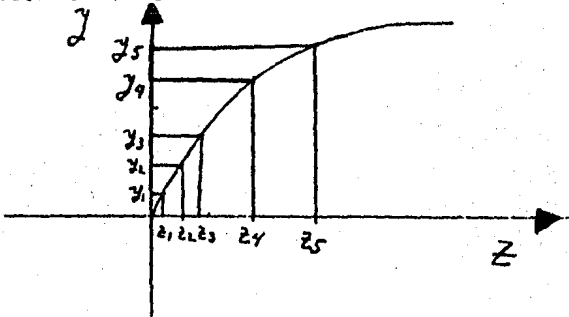


fig.2.1.1.3.3.3 Curva de compresión. para el caso especial donde la función de densidad de probabilidad es laplaciana

$$P(x) = \frac{1}{\lambda} e^{-2|x|/\lambda}$$

de la ecuación (9)

$$q(y) = y^2 P(y)$$

$$= C1 \int_0^s \frac{P(\alpha) d\alpha}{2/3}$$

$$F(s) = C1 \int_0^s P(\alpha) d\alpha$$

$$= C1 \int_0^s \frac{1}{\lambda} e^{-2|\alpha|/3\lambda} d\alpha$$

$$= -\frac{3}{2} C1 E \left. e^{-2|\alpha|/3\lambda} \right|_0^s$$

$$F(s) = \frac{-3}{2} C1 E e^{-2|s|/3\lambda} + \frac{3}{2} C1$$

$$F(v) = \frac{-3}{2} C1 E e^{-2|v|/3\lambda} + \frac{3}{2} C1 = V$$

por lo tanto

$$C1 = \frac{V}{\left(\frac{3}{2} - \frac{3}{2} E\right) \cdot -2|v|/3\lambda}$$

entonces

$$F(\bar{x}) = \frac{-V E \cdot -2|\bar{x}|/3\lambda}{(1 - E) \cdot -2V/3} + \frac{V}{(1 - E) \cdot -2V/3}$$

$$F(\bar{x}) = \frac{V(1 - E) \cdot -2|\bar{x}|/3}{(1 - E) \cdot -2|V|/3}$$

si hacemos $m=2/3$

$$F(\bar{x}) = \frac{V(1 - E) \cdot -m\bar{x}}{(1 - E) \cdot -mV}$$

que es el compresor óptimo para una PDF laplaciana.

2.1.1.3.4 Cuantización Robusta -

En algunas aplicaciones, principalmente en la transmisión de la voz, el cuantizador deberá acomodar señales con niveles de potencia que varían ampliamente. La utilización de cuantizadores "robustos", los cuales son relativamente insensibles a los cambios en la función de densidad de probabilidad de las muestras de entrada, han sido utilizados ampliamente.

Para obtener cuantizador robusto, la SNR del cuantizador deberá ser idealmente independiente de la función de densidad de probabilidad particular de la señal de entrada. Si la pendiente de la curva del compresor es escogida como:

$$a(x) = \frac{v}{B|x|}$$

entonces

$$D = \frac{v^2}{2 \cdot 3N} \int_{-v}^v \frac{P(s) ds}{[a(s)]^2}$$

$$D = \frac{v^2}{2 \cdot 3N} \int_{-v}^v \frac{P(s) ds}{\frac{v^2}{2 \cdot b^2 |s|^2}}$$

$$= \frac{b^2}{2 \cdot 3N} \int_{-V}^V P(x) |x|^2 dx$$

$$D = \frac{b^2}{2 \cdot 3N} \bar{V}^2$$

así que la SNR \bar{V}^2/D se reduce a la constante $3N/b^2$, la cual es en efecto independiente de la PDF particular de la señal de entrada.

Integrando $q(x)$

$$F(x) = \int q(x) dx = \int \frac{V}{bx} dx = \frac{V}{b} \int x^{-1} dx$$

$$F(x) = \frac{V}{b} \log x + C$$

para $x > 0$, donde c es una constante. Este resultado muestra que tal curva de compresión logarítmica debería dar la característica de robustes deseada. Para la expresión:

$$D = \frac{V^2}{2 \cdot 3N} \int_{-V}^V \frac{P(s) ds}{[q(s)]^2}$$

el ruido de sobrecarga es rechazado, así que la SNR no continuará constante si no que empezará a caer, cuando el nivel de potencia se haga suficientemente grande.

La curva de compresión obtenida es en efecto no realizable, ya que $F(0)$ no es finita. Para evitar tal dificultad, una curva de cuantización modificada es entonces utilizada, la cual se comporta bien para valores pequeños X y retiene el comportamiento logarítmico para los de demás valores.

La mitad positiva de esta curva esta dada por:

$$F(X) = \frac{V}{b} \log X + C = V + C \log x - C \log V$$

$$C = V + C \log X - C \log V - \frac{V}{b} \log X; C = \frac{V}{b}$$

$$\Rightarrow C = V - \frac{V}{b} \log V$$

$$F(X) = \frac{V}{b} \log X + V - \frac{V}{b} \log V$$

$$= V + \frac{V}{b} \log (X/V)$$

y finalmente

$$F(X) = \frac{V}{\log(1 + \mu)} \frac{\log(1 + \mu X/V)}{\log(1 + \mu)}$$

μ =parametro de compresion

X =voltage de entrada

F(x)=voltage de salida

V =voltage maximo de entrada

El parametro μ controla el grado de compresión y puede ser escojido tan grande tal que para cambios grandes en la entrada produce cambios pequeños en la salida. F(x) es una función impar, así que $F(x) = -F(-x)$ para x negativas.

2.1.1.4 Codificación -

Ya que contamos con la señal muestreada y cuantizada es muy común que esta se codifique en un grupo o paquete equivalente de pulsos binarios de igual amplitud. La codificación de niveles de amplitud en forma binaria puede realizarse de varias maneras; uno de los procedimientos es emplear la conversión usual entre decimales y binarios, que se ha tabulado para la siguiente tabla:

Dígito	Código binario				Código Gray			
	b1	b2	b3	b4	q1	q2	q3	q4
0	0	0	0	0	0	0	0	0
1	0	0	0	1	0	0	0	1
2	0	0	1	0	0	0	1	1
3	0	0	1	1	0	0	1	0
4	0	1	0	0	0	1	1	0
5	0	1	0	1	0	1	1	1
6	0	1	1	0	0	1	0	1
7	0	1	1	1	0	1	0	0
8	1	0	0	0	1	1	0	0
9	1	0	0	1	1	1	0	1
10	1	0	1	0	1	1	1	1
11	1	0	1	1	1	1	1	0
12	1	1	0	0	1	0	1	0
13	1	1	0	1	1	0	1	1
14	1	1	1	0	1	0	0	1
15	1	1	1	1	1	0	0	1

Tabla 1.

Una de las dificultades que se presentan con la conversión normal de decimal a binario es que, al cambiar un dígito decimal a otro adyacente, el código binario se modifica un número variable de dígitos binarios. Esto es particularmente cierto al cambiar de 3 a 4 y de 7 a 8 en la tabla 1. En el primer caso se modifican tres dígitos binarios, y en el segundo cuatro. En casos donde se utilicen más bits, la variación de dígitos binarios será mayor entre un número y otro, lo cual hace que el código binario sea muy susceptible de errores durante la conversión analógica digital. Sería preferible tener un código

en el cual solo un dígito se modificara a la vez cuando el código decimal correspondiente cambie de un nivel al siguiente. El código Gray, que se muestra en la tabla 1 es un ejemplo de esta clase de códigos.

El código Gray se obtiene facilmente del código binario por medio de formulas de conversión. Supongase que en general se utiliza un código binario de n dígitos. (Esto corresponde a 2^n posibles dígitos decimales, incluyendo el cero). Enumerese los bits sucesivamente por orden, desde el más significativo hasta el menos significativo, $b_1 b_2 b_3 \dots b_n$. Sean los correspondientes dígitos del código Gray $g_1 g_2 g_3 \dots g_n$. Las siguientes ecuaciones convierten el código binario en código Gray :

$$g_1 = b_1$$

$$g_k = b_k \oplus b_{k-1} \quad k > 2$$

donde el símbolo \oplus representa suma módulo 2 o suma or exclusiva de los números binarios.

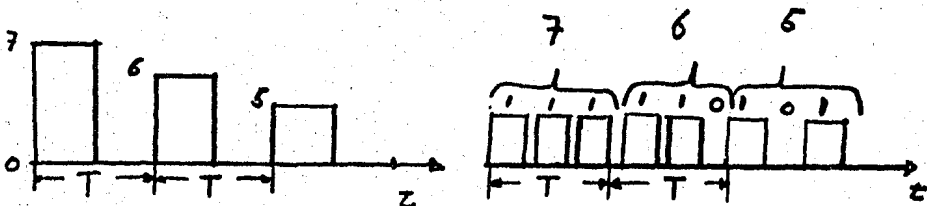


Fig.2.1.1.4.1 Codificación binaria de las muestras:
a) Señal original (va muestreada v cuantizada)
b) Muestras codificadas.

En la figura anterior se muestra un ejemplo de codificación.

binaria. En este caso se transcriben tres pulsos binarios en el intervalo de muestreo original, en el cual el bit uno se presenta por la presencia de un pulso, el bit cero por su ausencia.

2.2 Codificadores en el dominio del tiempo

Para la codificación en el dominio del tiempo contamos con las siguientes técnicas.

- 1.- PCM
- 2.- PCM logarítmica
- 3.- APCM
- 4.- DPCM
- 5.- ADPCM

cada una de estas técnicas se describirán a continuación.

2.2.1 P C M (Modulación por Codificación de Pulsos)

Los sistemas que implican la transmisión de señales digitalizadas y codificadas se denominan comúnmente sistemas de modulación por codificación de pulsos (PCM). Los sistemas digitales binarios constituyen la clase más común de sistemas PCM que se encuentran.

Hay muchas ventajas en la utilización de los sistemas de este tipo:

1. Las señales pueden regenerarse o rearmarse periódicamente durante la transmisión, puesto que la información ya no se encuentra contenida en la amplitud continuamente variable de los pulsos, sino que consiste en símbolos discretos.
2. Toda clase de circuitos digitales puede emplearse durante la totalidad del procesamiento.
3. Las señales pueden ser procesadas digitalmente según convenga.
4. El ruido y la interferencia pueden ser apropiadamente minimizados mediante códigos, etc.

Como se vio en los incisos anteriores (2.1.1.4) los códigos binarios son justamente un caso especial de la codificación teórica posible en los sistemas PCM. En general una muestra cuantizada de una señal puede codificarse por medio de un grupo de m pulsos, cada uno de los cuales tiene n posibles niveles de amplitud. Estos m pulsos deben de ser transmitidos en el mismo intervalo original que esta dispuesto para la muestra cuantizada. Como la información transportada por estos m pulsos es equivalente a la que llevan los M niveles originales de amplitud, el número de posibles combinaciones de amplitudes de estos pulsos debe ser igual a M . Por lo tanto:

$$M = n^m$$

Si hay dos posibles niveles, $n=2$, se tiene justamente el código mencionado. Con $M=8$, se necesitan tres pulsos binarios; si $n=3$, resulta un código ternario. Obviamente, si $m=1$, $n=M$ y se tiene el número de muestras no codificadas pero cuantizadas. En tanto que el número n de niveles escogidos aumenta, m disminuye, así como el ancho de banda requerido para la transmisión.

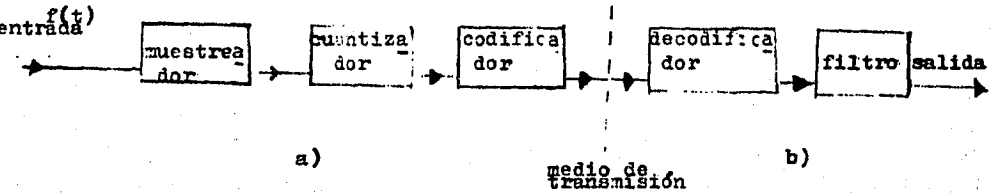
Esta posibilidad de codificar de diferentes formas es una de las razones del intenso uso que se hace de los sistemas PCM. Aunque la mayor parte de ello actualmente en uso utiliza la forma binaria, de manera que el ancho de banda requerido para la transmisión es mayor que el que se necesitaba antes de la codificación binaria.

Para que la información que llevan las señales binarias se represente, se necesita, usar la ausencia o presencia del pulso o su polaridad. Todo lo que tiene que hacer el receptor es reconocer, en consecuencia la ausencia o presencia del pulso o su polaridad, y a continuación decodificar la forma cuantizada original para reconstruir la señal. La forma del pulso o su amplitud exacta no tiene mayor significación. Al transmitirse pulsos binarios de amplitud suficientemente elevada, puede asegurarse la detección correcta del pulso en presencia del ruido con una relación de error tan baja como se desee.

Como ya se ha notado, los sistemas PCM binarios se presentan muy bien para que la señal se reconforme a intervalos periódicos durante la transmisión, lo cual es una práctica común en las comunicaciones telefónicas sobre circuitos telefónicos. La reconfiguración de las señales de estos repetidores intermedios mejora las decisiones sobre las propias señales cuando esta llegan finalmente a los receptores.

Ahora es interesante reunir los procesos de cuantización

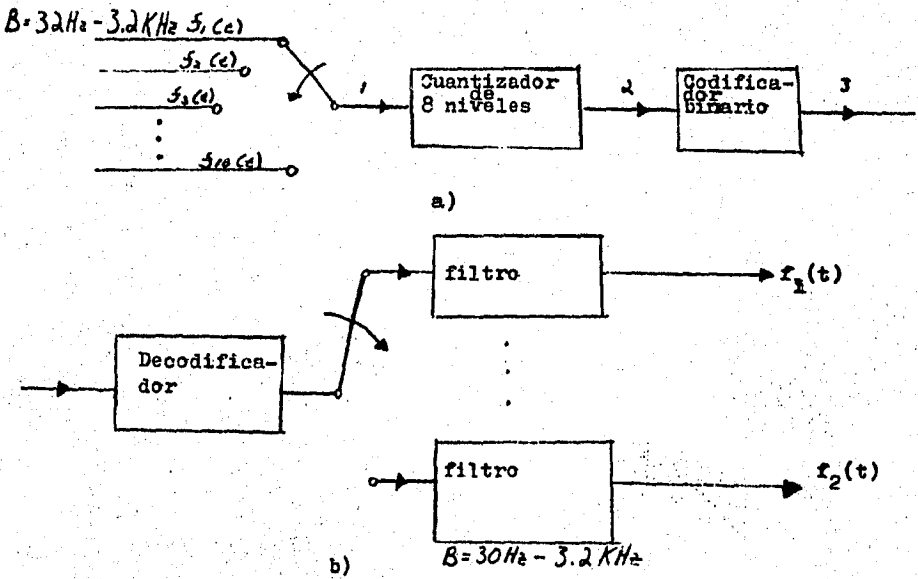
muestreo y codificación binaria, así como los procesos inversos correspondientes que se lleven a cabo en el receptor.



a) medio de transmisión b)

Fig. 2.2.1.1 Sistema PCM de un solo canal: a) transmisor. b)Receptor.

En la siguiente figura se tiene un sistema PCM de 10 canales y se muestran los anchos y los tiempos requeridos a medida que se avanza en el sistema.



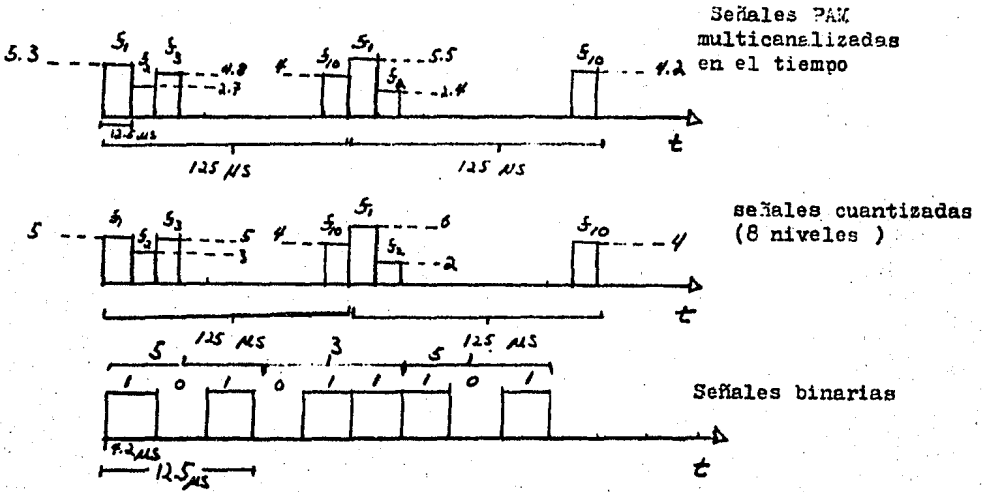


fig.2.2.1.2 Sistema PCM de 10 canales
 a) Transmisor. b) Receptor
 c) Formas de las señales.

Si los canales originales tienen cada uno 3.2 kHz de ancho de banda, como se muestra, y se usa una velocidad de muestreo de 8000 muestras por segundo, los pulsos multicanalizados en el tiempo a la salida del muestreador aparecen a intervalos de 12.5 μs. Suponiendo que el cuantizador es de ocho niveles, con pasos enteros, por similitud, los pulsos cuantizados de salida también aparecen a intervalos de 12.5 μs, pero con amplitudes enteras solamente. Cada pulso cuantizado se codifica entonces en tres pulsos binarios, cada uno de los cuales ocupa un espacio de tiempo de 4.2 μs, según se muestra.

Si los pulsos de dispusieran de modo que ocuparan los intervalos de tiempo completos como se muestra en la figura ,los anchos de banda requeridos en estos tres puntos mostrados serian aproximadamente:

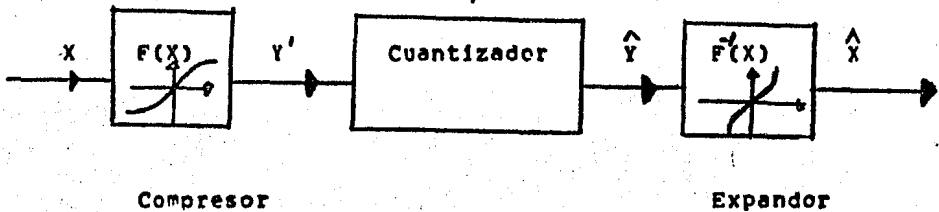
1. 80 kHz (1/12.5 μ s)
2. 80 Khz
3. 240 khz

(1) corresponde al ancho de banda PAM ;(3)es el ancho de banda de PCM binario.

2.2.2 PCM Logarítmica

Como se vio en la Cuantización no uniforme (2.1.1.3.3) para obtener una Relación Señal a Ruido (SNR) relativamente fija en un intervalo dinámico de señales más amplio es necesario introducir una desuniformización de los niveles de cuantización. Una opción, consiste en la práctica comprimir la señal en forma no lineal y después aplicar el espaciamento no uniforme de los niveles de la señal ya comprimida. En el receptor es expandida entonces con una característica inversa también no lineal. Es obvio que esto equivale al espaciamento no uniforme de los niveles.

Para realizar la PCM logarítmica supondríamos entonces el modelo siguiente:



Como se vio en la Cuantización Robusta una forma especialmente común que se utiliza en la práctica para la telefonía de voz es la compansión de la ley μ . Que tiene la siguiente forma:

$$y(x') = \frac{\ln(1 + \mu x' / V)}{\ln(1 + \mu)} \quad 0 < x' < V$$

el cual tiene cual tiene simetría impar con respecto al punto $x'=0$. El parametro puede variarse con el objeto de obtener una familia de curvas características. En los Estados Unidos, bell System ha adoptado una ley de compresión $\mu = 255$ para sus

sistemas de portadora digitales.

La siguiente figura nos muestra esta característica

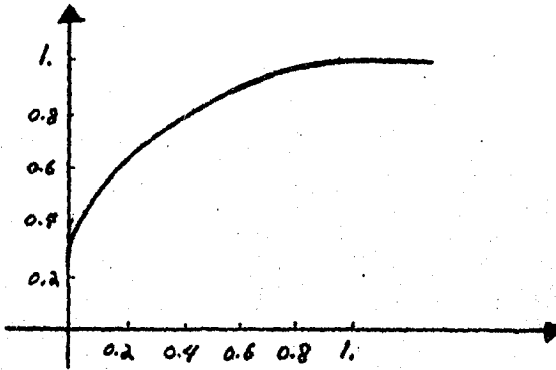


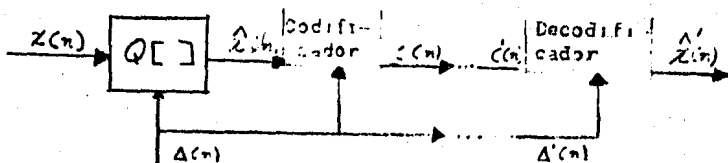
Fig. 2.2.2.1 Característica del compresor
compresión de la ley μ , $\mu=255$ (se muestra
solo el cuadrante positivo).

2.2.3 APCM Cuantización Adaptiva.

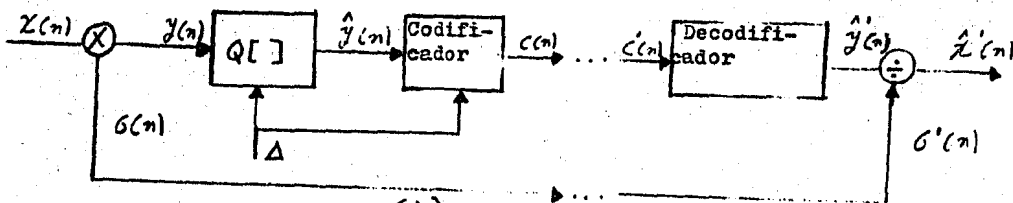
En la cuantización de señales de voz se desea escoger el tamaño del nivel de cuantización lo suficientemente grande para poder acomodar el rango pico a pico de la señal ,por una parte y por otra parte se desea hacer el nivel de cuantización tan pequeño tanto que minimice el ruido de cuantización.

Esta mezcla es por la naturaleza no estacionaria de la señal de voz y el proceso de comunicación de la voz. La amplitud de la señal de voz puede variar sobre un rango muy amplio dependiendo de la persona que esta hablando ,el medio de comunicación y en un extremo ,por los segmentos de vocal o consonante. Como se ha visto anteriormente una forma de acomodar estas fluctuaciones de amplitud es el uso de cuantizadores no uniformes. Otra alternativa es la de adaptar las propiedades del cuantizador a el nivel de la señal de entrada.

Cuando la cuantización adaptiva es utilizada directamente sobre las muestras de entrada del sistema es llamada PCM adaptiva o simplemente ,APCM. La idea básica de la cuantización adaptiva es dejar que el tamaño de los niveles de cuantización y los rangos varien de tal forma que sigan la variancia de la señal de entrada. Esto es mostrado en el siguiente esquema:



(a)



(b)

FIG. 2.2.3.1 Representación en diagrama de bloques de la cuantización adaptiva: a) representación del tamaño del escalón variable; b) representación de ganancia variable

Un punto de vista alternativo descrito en la figura anterior es el considerar un cuantizador fijo precedido por una ganancia variante con el tiempo el cual tiende a mantener la variancia constante. En el primer caso el tamaño del escalón tendra a crecer o decrecer con el incremento o decremento de la variancia de la señal de entrada. En el caso de un cuantizador no uniforme, esto implicaría que los niveles de cuantización y los rangos fueran escalados linealmente para seguir la variancia de la señal. En el segundo punto de vista el cual aplica sin modificación a los cuantizadores uniformes y no uniformes, la ganancia cambia inversamente con los cambios en la variancia de entrada tal que mantenga la variancia de la entrada al cuantizador relativamente constante. En ambos casos es necesario obtener una estimación de las propiedades de la

amplitud variando en el tiempo de la señal de entrada.

En las propiedades de la señal de voz variante en el tiempo ,es conveniente considerar la escala del tiempo en la cual los cambios ocurren. En el caso de los cambios de amplitud ,nos referiremos a los cambios que ocurren de muestra en muestra o a cambios rápidos en pocas muestras como cambios instantáneos. En las propiedades de amplitud ,por ejemplo la amplitud pico de un intervalo de consonantes o de vocales , esencialmente se mantienen sin cambio para un intervalo relativamente grande de tiempo. Tales variaciones lentas son referidas como variaciones silábicas ,implican que ellas ocurren a una razón comparable a la razón de sílabas en la voz

Sería conveniente clasificar a los esquemas de cuantización de acuerdo si ellos son adaptadores lentos a rápidos; por ejemplos silábicos o instantáneos. En una clase de los esquemas ,la amplitud o la variancia de la entrada de la entrada por si misma. Tales esquemas son llamados cuantizadores adaptivos con realimentación adelantada. En el otro caso de cuantizadores adaptivos ,el tamaño del escalón es adaptado base a la salida del cuantizador , $x(n)$,o equivalentemente ,en base a las palabras codificadas de salida , $c(n)$. Estos cuantizadores son llamados cuantizadores con retroalimentación. En general ,el tiempo de adaptación de las dos clases de cuantizadores puede ser ambas silábicos o instantáneos.

2.2.3.1 Adaptación con realimentación adelantada -

La siguiente figura describe una representación general de los cuantizadores con realimentación adelantada

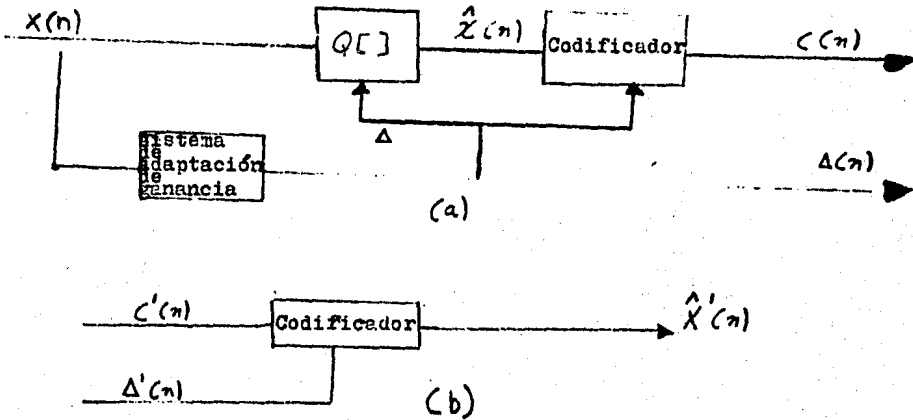


Fig. 2.2.3.1.1. Representación general de cuantizadores con realimentación adelantada; a) codificador b) decodificador

Asumimos por conveniencia que el cuantizador es uniforme así que es suficiente con variar el parámetro del tamaño del escalón. Esto se puede generalizar para el caso de cuantizadores no uniformes. El tamaño del escalón $\Delta(n)$, utilizado para cuantizar la muestra $x(n)$ en la figura anterior, debe estar disponible en el receptor en la misma figura. Por esto, las palabras codificadas $c(n)$ y el tamaño del escalón juntos representan la muestra $x(n)$. Si $c'(n)=c(n)$ u $\Delta'(n)=\Delta(n)$, entonces $\hat{x}'(n)=\hat{x}(n)$; sin embargo, si $c'(n) \neq c(n)$ o $\Delta'(n) \neq \Delta(n)$, por ejemplo si hay errores en la transmisión, entonces $\hat{x}'(n) \neq \hat{x}(n)$. Los efectos de los errores dependerán de los detalles de el esquema de adaptación.

La siguiente figura nos muestra el cuantizador adaptivo general con realimentación adelantada representado en términos

de la ganancia variando en el tiempo.

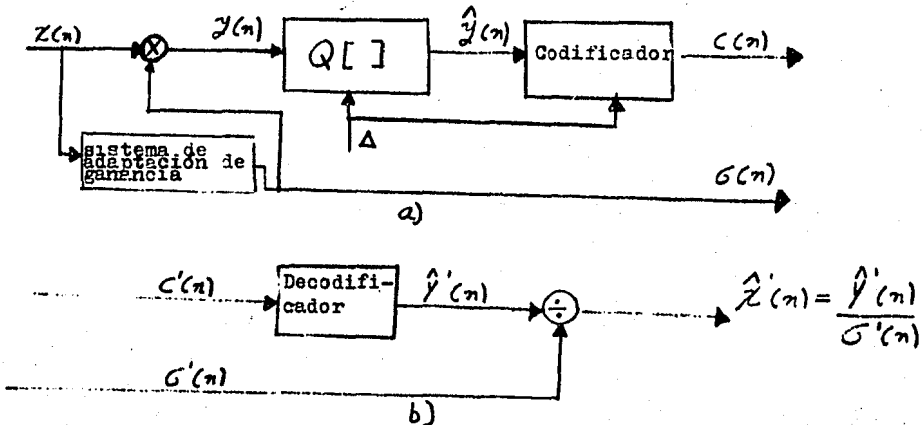


Fig. 2.2.3.1.2. Cuantizador adaptivo general con realimentación adelantada con ganancia variando en el tiempo; a)codificador; b)decodificador

En este caso, las palabras codificadas y la ganancia $G(n)$ juntas representan las muestras cuantizadas.

Para ver como trabaja este esquema es útil considerar algunos ejemplos. La mayoría de los sistemas de este tipo tratan de obtener una estimación de la variancia que varia con el tiempo. Entonces, el tamaño del escalón o los niveles de cuantización son hechos proporcionales a la desviación estandar, o la ganancia aplicada a la entrada puede ser hecha inversamente proporcional a la desviación estandar.

Una aproximación común es la de asumir que la variancia es proporcional a la energía en un intervalo de tiempo corto, la cual, es definida como la salida de un filtro paso baja con entrada, $x(n)$.

esto es;

$$\bar{V}^2(n) = \sum_{m=-\infty}^{n-1} x^2(m) h^2(n-m) \quad (2.3.1.1)$$

donde $h(n)$ es la respuesta impulso del filtro paso bajas.

Un ejemplo simple es;

$$h(n) = \begin{cases} \alpha & , n \geq 1 \\ 0 & , \text{ caso contrario} \end{cases} \quad (2.3.1.2)$$

Utilizando esta ecuación en (2.3.1.1) da:

$$\bar{V}^2(n) = \sum_{m=-\infty}^{n-1} x^2(m) \alpha^{n-m-1} \quad (2.3.1.3)$$

(n) en la ecuación (2.3.1.3) también satisface la ecuación de diferencia:

$$\bar{V}^2(n) = \alpha \bar{V}^2(n-1) + x^2(n-1) \quad (2.3.1.4)$$

El tamaño del escalón en la figura (2.3.1.1) será:

$$\Delta(n) = \Delta_0 \bar{V}(n) \quad (2.3.1.5)$$

o la ganancia variando con el tiempo en la figura (2.3.1.2) sería:

$$g(n) = \frac{G_0}{\sqrt{V(n)}} \quad (2.3.1.6)$$

La selección del parámetro α controla el intervalo efectivo que contribuye a la estimación de la variancia.

2.2.3.2 Adaptación retroalimentada. -

La segunda clase de cuantizadores adaptivos es descrito en las siguientes figuras, donde es notado que la variancia de la entrada es estimada de la salida del cuantizador o equivalentemente de los códigos de las palabras.

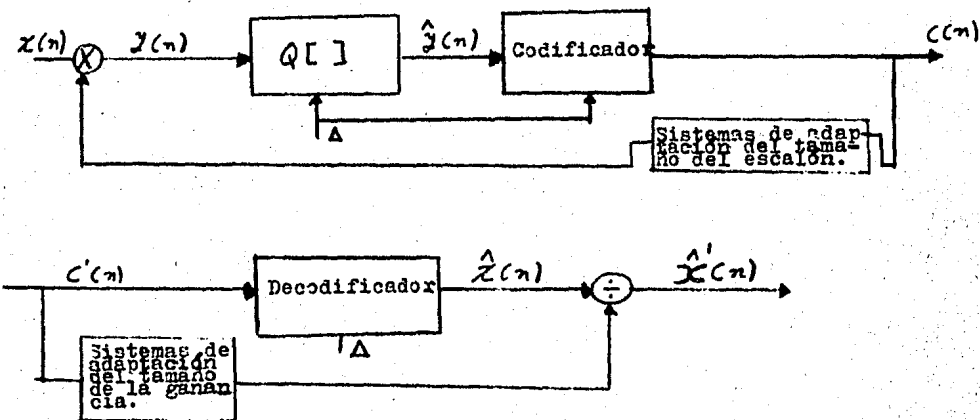


Fig. 2.2.3.2.1 Adaptación por retroalimentación general de las ganancias variando con el tiempo: a)codificador; b)decodificador

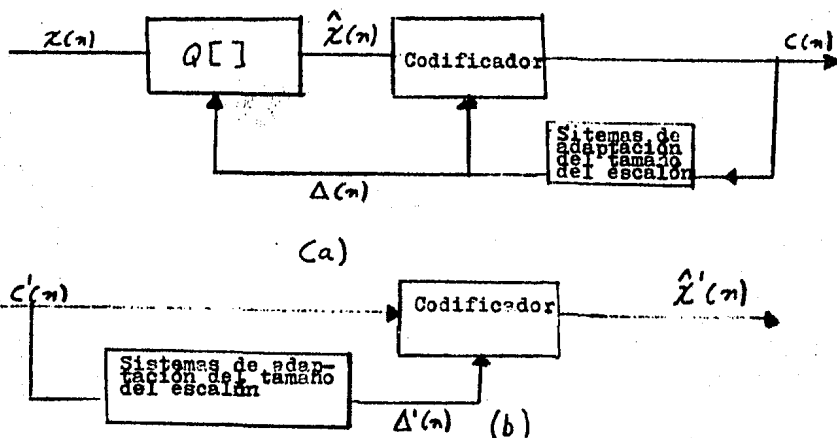


Fig. 2.2.3.2 Adaptación por retroalimentación general de el tamaño del escalón; a)codificador; b)decodificador

Como en el caso de adaptación hacia adelante, el tamaño del escalón y ganancia son proporcionales e inversamente proporcionales respectivamente a un estimador de la desviación estandar de la entrada como en las ecuaciones (2.3.1.5) y (2.3.1.6).

Tal sistema tiene la ventaja que el tamaño del escalón o la ganancia no necesitan ser transmitidos o retenidos ya que ellos pueden ser derivados de la secuencia de las palabras codificadas. La desventaja de tales sistemas es el incremento de la sensibilidad a los errores en las palabras codificadas, ya que tales errores no implican solo un error en los niveles de cuantización sino también en el tamaño del escalón.

Una aproximación simple es aplicar la ecuación (2.4.2.1) directamente a la salida del cuantizador

$$\hat{\sigma}^2(n) = \sum_{m=-\infty}^{\infty} \hat{x}^2(m) h(n-m)$$

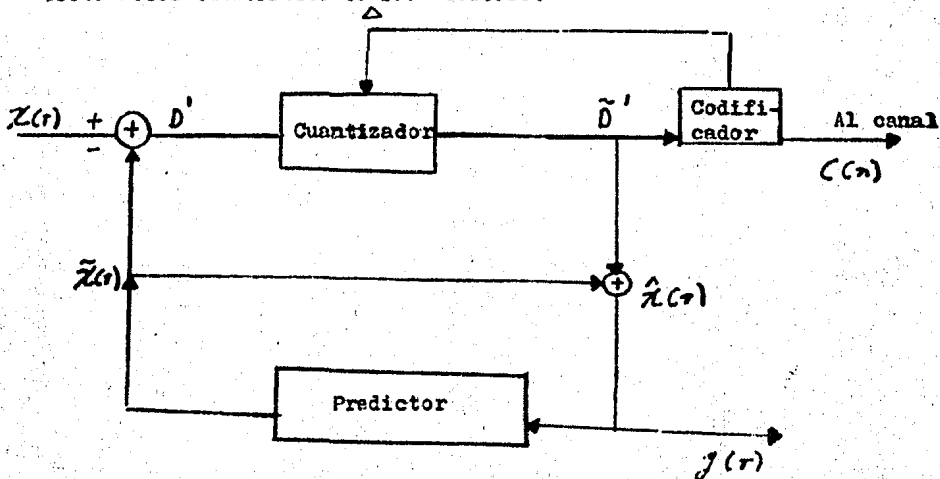
En este caso ,sin embargo no es posible utilizar buffers para implementar un filtro no causal. Esto es, el estimador de la variancia deberá estar basado solamente en los valores pasados de $x(n)$ ya que el presente valor no estará disponible hasta que la cuantización ha ocurrido , la cual en turno deberá ser después la variancia que ha sido estimada .

2.2.4 Modulación por Codificación de Pulsos Diferencial (DPCM)

Utilizando las redundancias que aparecen en la voz se pueden considerar técnicas mas inteligentes para codificar la voz. La redundancia, o la predictibilidad, pueden ser utilizadas tanto en el dominio del tiempo como en el dominio de la frecuencia.

Las amplitudes adyacentes en la forma de onda de la voz están altamente relacionadas. Esto significa que la variancia de la diferencia D entre las amplitudes de la voz $x(r)$ y $x(r-1)$ es mucho veces menor que la variancia de $x(r)$.

Esto inmediatamente sugiere un esquema donde la voz es representada no en términos de las amplitudes de la forma de onda si no en términos de la diferencia entre las amplitudes de la forma de onda. En otras palabras, en la codificación diferencial, se toman esas diferencias, se cuantizan y se recupera una aproximación de $x(r)$ esencialmente integrando las diferencias cuantizadas de las muestras.



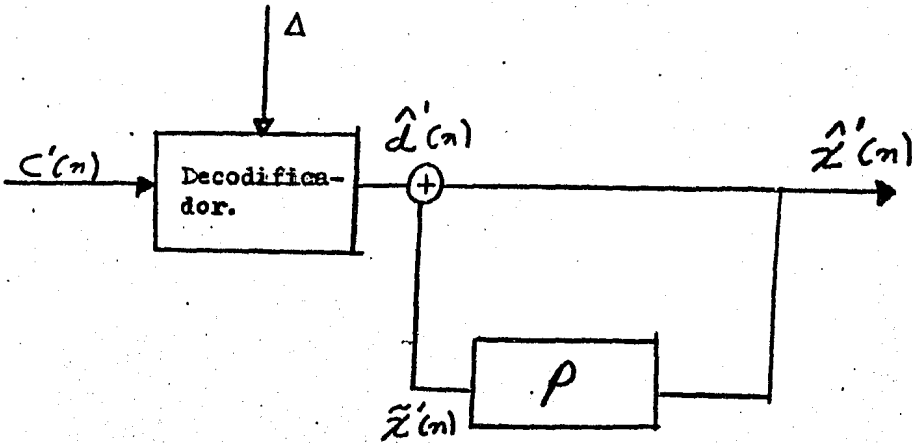


Fig. 2.2.4.1 Cuantizador DPCM

En este sistema la entrada al cuantizador es la señal

$$d(n) = x(n) - \tilde{x}(n)$$

la cual es la diferencia de la muestra sin cuantizar $x(n)$, y una estimación ,o predicción ,de la muestra de entrada la cual es denotada como $\tilde{x}(n)$. Este valor predicho es al salida de un sistema predictor P ,cuya entrada es una versión cuantizada de la versión de la señal de entrada $x(n)$. La señal de diferencia puede ser llamada la señal de error de predicción ,ya que es la cantidad por medio de la cual el predictor falla para predecir exactamente la entrada.

El sistema para reconstruir la entrada cuantizada del código de las palabras esta implícito en la figura anterior.

Este sistema involucra a un decodificador para reconstruir la señal de diferencia cuantizada de la cual el cuantizador de entrada es reconstruida utilizando el mismo predictor que en la parte superior de la figura. Es obvio que si $c'(n)$ es idéntica a $c(n)$ entonces $\hat{x}'(n) = \hat{x}(n)$ la cual difiere de $x(n)$ por el error de cuantización incurrido en la cuantización $d(n)$.

La relación señal a ruido del sistema anterior es ,por definición:

$$\text{SNR} = \frac{E [x (n)]^2}{E [e (n)]^2} = \frac{\sigma_x^2}{\sigma_e^2}$$

la cual puede ser escrita como:

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_d^2} \cdot \frac{\sigma_d^2}{\sigma_e^2} = G_D \text{SNR}_Q$$

donde:

$$\text{SNR}_Q = \frac{\sigma_d^2}{\sigma_e^2}$$

es la razón señal a ruido de la señal cuantizada , y la cantidad:

$$G_D = \frac{\sigma_x^2}{\sigma_d^2}$$

es definida como la ganancia debido a la configuración diferencial.

La cantidad SNR_Q es dependiente del cuantizador particular que es utilizado, y, dando un conocimiento de las propiedades de $d(n)$, SNR_Q puede ser maximizado. La cantidad G_p , si es más grande que la unidad, representa la ganancia que es debida en SNR al esquema diferencial. El objetivo es el de maximizar G_p por medio de una apropiada selección del sistema predictor P . Para una señal dada, \hat{V}_x^2 es una cantidad fija así que G_p puede maximizarse minimizando el denominador, minimizando la variancia del error prededido.

Para especificar $x(n)$ haremos uso de un predictor lineal. Esto es, $x(n)$ es una combinación lineal de los valores cuantizados anteriores:

$$x(n) = \sum_{k=1}^p \alpha_k \hat{x}(n-k)$$

El valor prededido es entonces la salida a un filtro de respuesta a impulso finita cuya función del sistema es:

$$P(z) = \sum_{k=1}^p \alpha_k z^{-k}$$

y cuya entrada es la señal reconstruida $\hat{x}(n)$. La señal reconstruida es la salida del sistema cuya función del sistema es:

$$H(z) = \frac{1}{1 - \sum_{k=1}^p \alpha_k z^{-k}}$$

y cuya entrada es la diferencia de la señal cuantizada. La variancia del error de predicción en la figura 2.4.1 es:

$$\sigma_d^2 = E\{d^2(n)\} = E\{(x(n) - \hat{x}(n))^2\}$$

$$= E\left\{ \left[x(n) - \sum_{k=1}^p \alpha_k \hat{x}(n-k) \right]^2 \right\}$$

$$= E\left\{ \left[x(n) - \sum_{k=1}^p \alpha_k x(n-k) - \sum_{k=1}^p \alpha_k e(n-k) \right]^2 \right\}$$

Para escoger el conjunto de los coeficientes del predictor $\{\alpha_j\}$, $1 \leq j \leq p$, que minimice σ_d^2 , debemos diferenciar σ_d^2 con respecto a cada parametro y poner las derivadas igual a cero, de ese modo obtener un conjunto de p ecuaciones:

$$\frac{\partial \sigma_d^2}{\partial \alpha_j} = -2E\left[\left[x(n) - \sum_{k=1}^p \alpha_k (x(n-k) + e(n-k)) \right] [x(n-j) + e(n-j)] \right]$$

$$= 0, \quad 1 \leq j \leq p$$

La cual se puede escribir en una forma mas compacta:

$$E\{(x(n) - \hat{x}(n))\hat{x}(n-j)\} = E\{d(n)\hat{x}(n-j)\} = 0$$

$$1 < j < p$$

de esta ecuación hacemos una importante observación que si los coeficientes del predictor son tales que minimicen σ_d^2 , entonces la señal de diferencia (error de prediccion) no esta correlacionado con los valores anteriores de la entrada al predictor $\hat{x}(n-j)$, $1 < j < p$

La ecuación anterior puede ser expandida en un conjunto de

p ecuaciones ;

$$\begin{aligned}
 E\{x(n-j)x(n)\} + E\{e(n-j)x(n)\} &= \sum_{k=1}^p \alpha_k E\{x(n-j)x(n-k)\} \\
 &+ \sum_{k=1}^p \alpha_k E\{e(n-j)x(n-k)\} \\
 &+ \sum_{k=1}^p \alpha_k E\{x(n-j)e(n-k)\} \\
 &+ \sum_{k=1}^p \alpha_k E\{e(n-j)e(n-k)\}
 \end{aligned}$$

donde $1 \leq j \leq P$. Si asumimos que la cuantización es razonablemente fina, puede ser asumido que $e(n)$ no está correlacionado con $x(n)$ y que $e(n)$ es una secuencia de ruido blanco:

$$E\{x(n-j)e(n-k)\} = 0, \quad \text{para toda } n, j, \text{ y } k$$

$$E\{e(n-j)e(n-k)\} = \sigma_e^2 \delta(j-k)$$

Utilizando estas suposiciones la ecuación anterior puede ser simplificada a:

$$\phi(j) = \sum_{k=1}^p \alpha_k [\phi(j-k) + \sigma_e^2 \delta(j-k)] \quad 1 \leq j \leq p$$

donde $\phi(j)$ es la función de autocorrelación de $x(n)$. Si dividimos ambos lados de la ecuación de arriba por σ_x^2 y definimos la autocorrelación normalizada como:

$$\rho(j) = \frac{\phi(j)}{\sigma_x^2}$$

entonces:

$$\rho = C \alpha$$

donde:

$$\rho = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \vdots \\ \vdots \\ \rho(p) \end{bmatrix}$$

y:

$$C = \begin{bmatrix} (1+1/\text{SNR}) & \rho(1) & \dots & \rho(p-1) \\ \rho(1) & (1+1/\text{SNR}) & \dots & \rho(p-2) \\ \vdots & \vdots & \ddots & \vdots \\ \vdots & \vdots & \vdots & \vdots \\ \rho(p-1) & \rho(p-2) & \dots & (1+1/\text{SNR}) \end{bmatrix}$$

y:

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_p \end{bmatrix}$$

y $SNR = \frac{\sigma_x^2}{\sigma_e^2}$. Así el vector de los coeficientes de predicción óptimo es obtenido la solución de la ecuación matricial:

$$\alpha = C^{-1}p$$

En general, la matriz C puede ser computada por una variedad de metodos numericos.

2.2.4.1 Modulación Delta -

La más simple aplicación de el concepto de cuantización diferencial es la modulación delta. En esta clase de sistemas, la razón de muestreo es escogida muchas veces mayor que la razón de Nyquist para la señal de entrada. Como resultado, las muestras adyacentes están altamente correlacionadas. En la modulación delta se tiene un cuantizador de un bit (2 niveles), el error de predicción es cuantizado a uno de los dos valores, uno positivo o uno negativo. Como se observa en la siguiente figura:

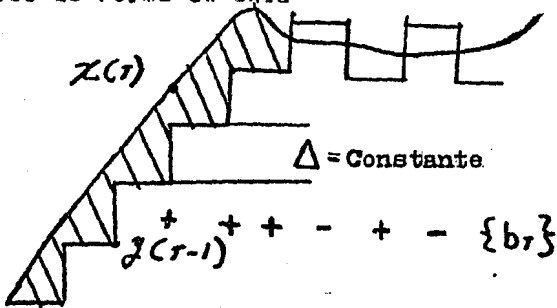


fig 2.2.4.1.1 Modulación Delta.

La operación de la modulación delta puede ser caracterizada muy simple. En cada muestra, se nota la la polaridad b , de la diferencia entre la muestra la muestra presente $x(r)$ y la última aproximación escalera de ella, $y(r-1)$; y entonces actualiza la función escalera en dirección de esta diferencia.

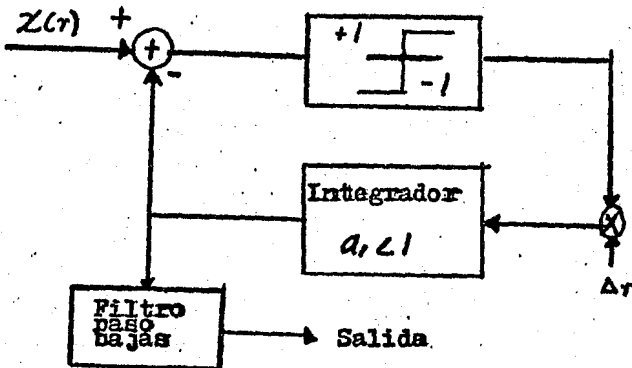


fig 2.2.4.1.2 Diagrama de bloques la modulación delta.

en este caso:

$$Y(r) = a_i Y(r-1) + \Delta r b_r$$

donde:

$$br = \text{sgn}(X(r) - Y(r-1))$$

Como en el caso de los sistemas PCM ,el ruido tambien se introduce cuando se emplea modulaci3n delta. A causa de que la se1al es cuantizada a dos niveles ,aparece un ruido granular o de cuantizaci3n ,similar al de PCM que se presenta en la salida del receptor . Al contrario de PCM que el ruido podria reducirse usando mas niveles de cuantizaci3n ,ahora solo puede reducirse muestreando con mayor frecuencia.

Adem1s del ruido de cuantizaci3n se presenta el ruido de sobrecarga ,y se presenta cuando los niveles de cuantizaci3n $\pm k'$ son muy peque1os para una se1al que var1a r1pidamente.

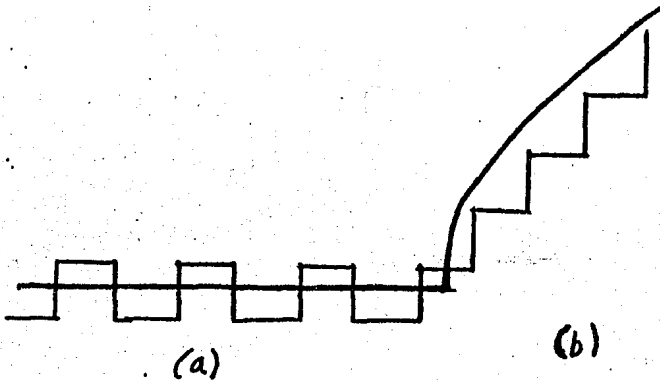


Fig. 2.2.4.3 Fuentes de ruido en la modulaci3n delta

a)ruido de cuantizaci3n; b)ruido de sobrecarga

2.2.5 Modulación por Codificación de Pulsos Diferencial Adaptiva

Convinando las técnicas de modulación adaptativa y modulación diferencial obtenemos una técnica que es llamada modulación por codificación de pulsos diferencial adaptiva.

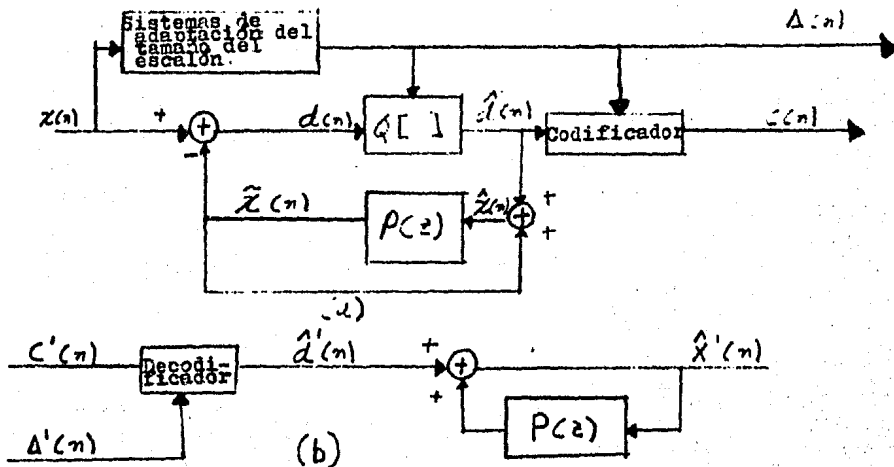


Fig. 2.2.4.1 ADPCM; a) Codificador; b) decodificador

Un caso especial de la modulación por pulsos diferencial adaptiva es la llamada modulación delta adaptiva.

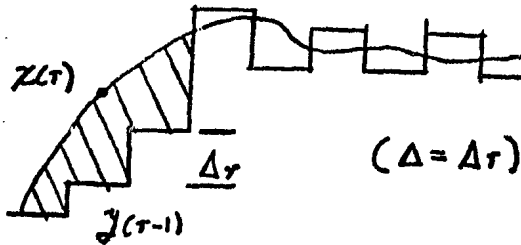


Fig. 2.2.5.2 Modulación Delta Adaptiva

como se vió en la modulación delta:

$$Y(r) = A1 \cdot Y(r-1) + \Delta r \text{ tr}$$

En este esquema el tamaño del escalón sigue los cambios de la pendiente de la señal de entrada. Las áreas que tienen rayas en la figura anterior representan el ruido de sobrecarga.

En un procedimiento muy simple con un bit de memoria ,se puede expandir el tamaño del escalón r por medio de un factor $P(>1)$ cuando hay sobrecarga ($br = br$), y reduce el tamaño del escalón por un factor $P (<1)$ cuando hay granular ($br = br$). Una regla de adaptación más general es de la forma:

$$\Delta r = f(\Delta r-1, br, br-1, br-2, \dots)$$

Una versión más robusta de la modulación delta adaptiva puede ser realizada en un procedimiento de adaptación que es caracterizado por una memoria mas grande M (típicamente $M=2$ o 3) para el incremento del escalón ,y una constante $par (.996)$ para

decrementar el tamaño del escalón:

$$r = 0.996 \Delta r_{-1} + \Delta_0 \cdot [\text{ADAPT}]r$$

$$[\text{ADAPT}]r = 1 \text{ si } \left| \sum_{s=1}^M b_{r-s-p} \right| = 4 \text{ para } p = 0 \text{ o } 1 \text{ o } 2$$

$[\text{ADAPT}]r = 0$ en caso contrario;

Δ_0 : positivo, factor de escala proporcional a la entrada
donde:

$$b_{\mu} = +1 \text{ (DM bits)}$$

2.3 Codificadores en el dominio de la frecuencia

En los algoritmos de codificación en el dominio del tiempo ,la voz es tratada como una señal de banda completa. Otra clase de algoritmos de codificación es aquella que divide la señal de voz en un número separado de componentes de frecuencia y codifica cada uno estas componentes por separado.

Estas técnicas en el dominio de la frecuencia tienen la ventaja adicional que le número de bits utilizado para codificar cada una de las componentes en frecuencia pueden variar dinámicamente y compartidos con otra bandas ,asi que la exactitud de la codificación es siempre puesta donde es necesaria en el dominio de la frecuencia. En efecto ,bandas con poca energía o sin energía pueden ser no codificadas.

Como en el dominio del tiempo ,existe una gran variedad de algoritmos en el dominio de la frecuencia ,de los más simples hasta los más complejos ,la diferencia principal entre estas técnicas son usualmente determinados por el grado de predicción que se utiliza en estas técnicas.

2.3.1 SBC (Codificación Sub-Banda)

En la codificación sub-banda la banda de voz es dividida entre típicamente entre cuatro a ocho sub-bandas por medio de un banco de filtros paso banda. Cada sub-banda es, en efecto, trasladada paso baja a la frecuencia cero por medio de un proceso de modulación equivalente a la modulación de banda lateral sencilla. Es entonces muestreada (o remuestreada) a su razón de Nyquist (el doble del ancho de la banda) y digitalmente codificada por medio de un codificador pulsos adaptivo (APCM). En este proceso cada banda puede ser codificada de acuerdo al criterio perceptual que son especificados a la banda. En la reconstrucción, las señales sub-banda son decodificadas y moduladas de regreso a sus bandas originales. Entonces ellas son sumadas para dar una réplica cercana a la señal original de voz.

La codificación por sub-bandas ofrece varias ventajas. Para cada banda se cuenta con cuantizadores adaptivos, así bandas con baja energía tendrán cuantizadores adaptivos pequeños que contribuirán con menos ruido de cuantización. Haciendo una colocación de bits apropiada en las diferentes bandas, el ruido por cuantización puede ser controlado en la frecuencia. En las bandas de baja frecuencia, donde el tono y la estructura del formant debe ser con exactitud preservada, un número grande de bits/muestra puede ser utilizado; y para las bandas de frecuencias superiores, donde sonidos fricativos y ruido pueden ocurrir en la voz, menos bits pueden ser utilizados.

La siguiente figura nos muestra la estructura básica del codificador sub-banda.

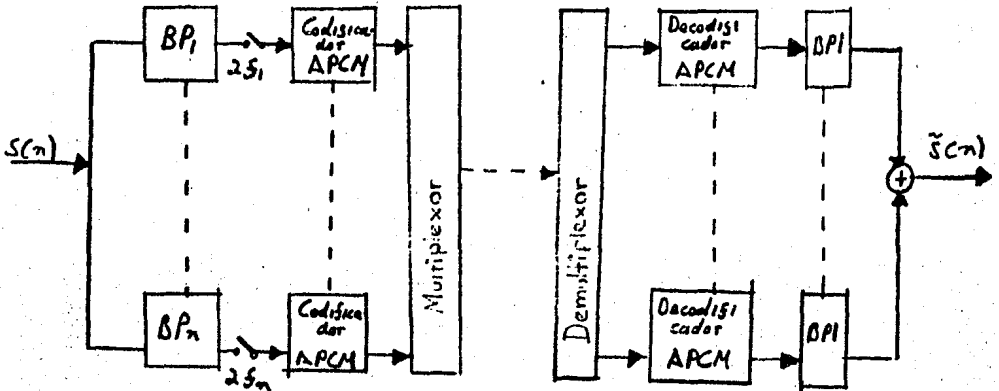


Fig 2.1.1.1 Diagrama en bloques del código sub-banda

Esencialmente consiste de un banco de filtros paso banda, codificadores APCM, y un multiplexor. La modulación es obtenida utilizando la técnica de muestreo de banda entero.

La parte más compleja del codificador es el banco de filtros, pero con las nuevas tecnologías tales como los filtros digitales esta complejidad está siendo rápidamente reducida a nivel chip.

CAPITULO 3

CODIFICADORES FUENTE (VOCODERS)

3.1 Introducción

Ya vimos los codificadores de forma de onda ,ahora veremos otra clase de codificadores llamada codificadores fuente (vocoders). Esta clase de codificadores dependen de una descripción frugal de la señal utilizando un conocimiento a priori de como la señal fue generada en la fuente. La idea es que algunas características físicas de la generación de la señal pueden ser cuantificadas ,y transformadas para obtener una eficiente descripción de la señal. Esto implica que la señal debe ser ajustada a un modelo específico y a una parametrización en conformidad.

El modelo tradicional de generación de voz es el modelo de sistema-fuente mostrado en la siguiente figura.

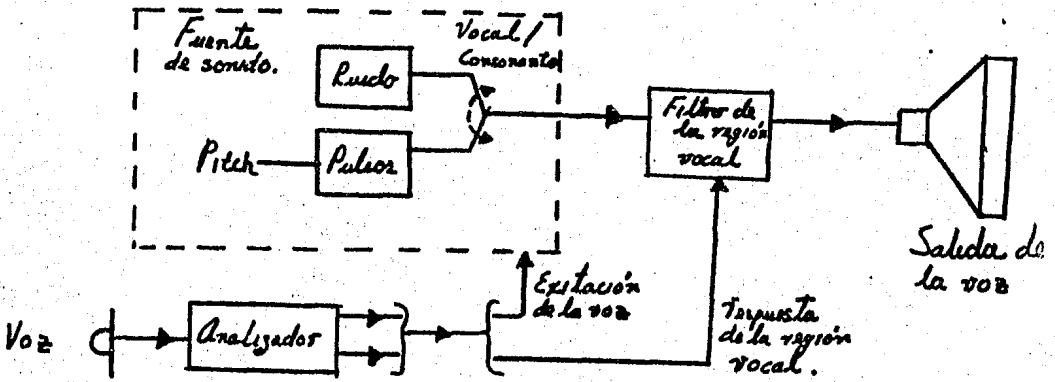


Fig 3.1.1 Modelo de generación de voz

El mecanismo de generación de sonido (la fuente) es asumido ser como linealmente separado de la modulación inteligente ,el filtro de la región vocal (el sistema). Otras asuncpciones son

las que los sonidos de la voz pueden ser separados como vocales o consonantes ,y que ellos son generados como un sonido de cuerda vocal cuasi-periódico ,o de un sonido aleatorio producido por un flujo de aire turbulento.

Los parametros de este modelo son tono de la voz ,los polos de frecuencia de los filtros de modulación ,y los parametros correspondientes de amplitud.

Los sonidos pueden ser generados por el aparato vocal de tres maneras

1.- Vocales. Las cuerdas vocales actuan como un oscilador aerodinámico, una válvula controla el flujo de aire hacia de los pulmones hacia la faringe. El flujo resultante es una secuencia cercana de pulsos de aire.

La impedancia acústica de la válvula ,relativa al punto de acción de la impedancia de la región ,es moderadamente alto ,asi que las cuerdas vocales pueden ser pensadas como una fuente de corriente la cual exita a una red lineal ,pasiva ,que varia poco en el tiempo.

2.- Fricativas. sonidos no vocales (consonantes) sostenidos son producidos por presión de sonidos aleatorios debidos a un flujo de aire turbulento en un punto estrecho en el sistema vocal. Desde un punto de vista eléctrico ,esta fuente aparece como una fuente de voltage en serie ,cuya impedancia interna es esencialmente dependiente del flujo ,es invariante con el tiempo y es típicamente de un valor grande.

3.-Altos. Transitorios ,sonidos consonantes son producidos de una presión de aire construida dentro de una completa oclusión ,y soltar abruptamente tal presión. El resultado es un transitorio de presión ,parecido a una función de paso y asi

teniendo un espectro cayendo a $1/f$, aplicado como una serie de voltage (presión) en el lugar de constreñimiento. La impedancia de la fuente es dictada por el tamaño del constreñimiento.

Para propósitos del vocoder, la fuente para sonidos vocales es representada como un generador de pulsos periódico. La fuente para sonidos no vocales (consonantes) es representado por un generador de ruido random. Las fuentes son normalmente consideradas mutuamente exclusivas con un parametro de la señal que indica la conmutación entre las fuentes de vocales y consonates.

La intensidad del sonido de excitación para cada fuente es también representado paramétricamente por medio de un amplitud o ganancia de la señal. En adición, la periodicidad o "tono" de la fuente de pulsos de las vocales debere ser especificada por el parametro de tono de la señal (pitch). La frecuencia del tono de las vocales es dependiente del parlante. Es típicamente a dos octavos de rango, cerca de 50-200 hz para los hombres y cerca de 100-400 para las mujeres.

En acordancia con el concepto de sistema lineal, la salida del sonido de la región vocal es la convolución (en el tiempo) de la forma de onda de excitación y la respuesta impulso del sistema vocal. En el dominio de la frecuencia, esto es equivalente a la multiplicación de la transformada de Fourier de estas cantidades. Consecuentemente, las resonancias acústicas de la región vocal modulan o modelan un espectro amplio de las fuentes. Diferentes sonidos de la voz corresponden únitamente a diferentes modelos espectrales. Las fuentes del sonido sirven como un portador acústico así que puede ser irradiado al aire.

Los vocoders dependen de una descripción paramétrica de la función de transferencia del área vocal. Esta descripción

paramétrica puede tomar una variedad de formas: por ejemplo, valores del espectro de amplitudes (short time) de la señal de voz evaluadas en frecuencias específicas (como en el caso de un Channel Vocoder), coeficientes de predicción lineal que describen la envolvente espectral (LPC vocoder), valores de frecuencias de las resonancias del espectro más grande (formant vocoder), muestras específicas de la función de autocorrelación (short time) de la señal de entrada (vocoder de autocorrelación), coeficientes de un conjunto de funciones ortonormales que aproximan la forma de onda de la señal (vocoder ortogonal).

Todos los modelos de vocoders, dependen del modelo expuesto en la figura 3.1.1.

3.2 Vocoders en el dominio de la frecuencia

La voz consiste de una sucesión de sonidos de voz o fonemas. Cada sonido de voz es caracterizado por la potencia de su espectro (el cuadrado de la magnitud de la transformada de Fourier de la forma de onda de la señal de voz).

El modelo espectral de cada sonido es determinado por medio de la configuración geométrica del mecanismo de producción de la voz humana, en otras palabras, la posición de la lengua, la abertura de los labios, etc.

El oído humano ha evolucionado de tal manera que analiza y distingue los diferentes espectros y así reconocer los sucesivos fonemas y reconstruir la voz

3.2.1 Channel Vocoder

El simple hecho a saber que la voz es codificada y analizada por medio de su espectro ,es explotado directamente en el así llamado Channel Vocoder.

En lugar de transmitir la señal de voz directamente (con un ancho de banda de 4000 Hz) ,uno puede transmitir solamente solo el espectro de cada señal de voz ,los cuales pueden ser razonablemente bien descrito por medio de 16 valores a lo largo del eje de frecuencias ,tomados cada 20 ms. De acuerdo al teorema de muestreo ,cada espectro requiere un ancho de banda $1/(2 \times 20 \text{ ms}) = 25 \text{ hz}$. El ancho de banda total reducido para la transmisión de la información es igual a $16 \times 25 = 400 \text{ Hz}$ o cerca de un décimo del ancho de banda de la señal.

El siguiente diagrama nos muestra la configuración de un Channel vocoder.

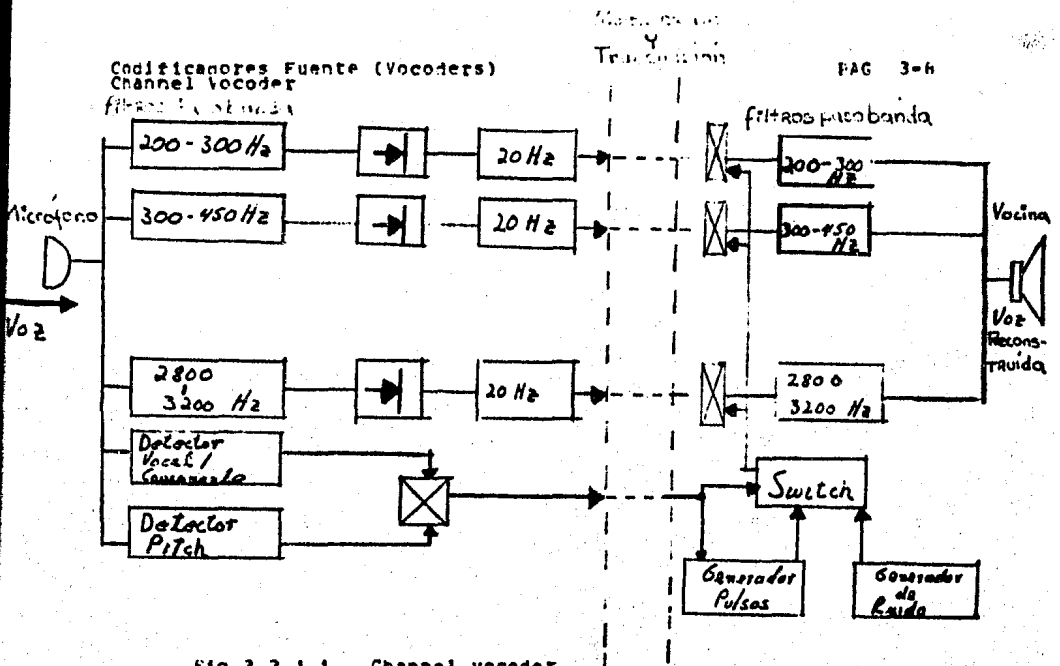


FIG 3.2.1.1. Channel vocoder

La voz es separada en bandas espectrales continuas. La salida de cada filtro es conectada a un rectificador y a un filtro paso bajas cuya salida representa el promedio de la amplitud de la señal variando en el tiempo para cada banda de frecuencia.

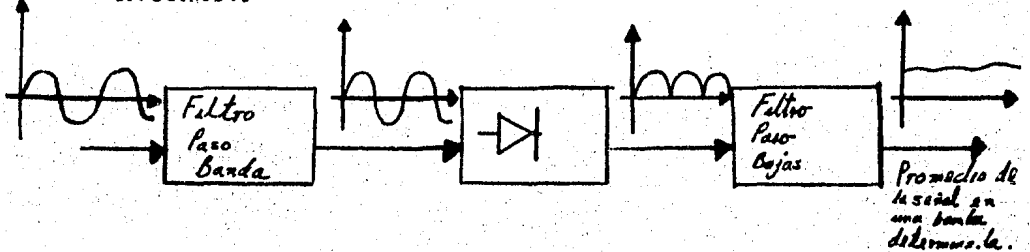


FIG 3.2.1.2 Analisis.

Como se mostró en la figura 3.2.1.1 hay un detector de vocales y consonantes y un detector de tono (pitch) los cuales determinan la estructura fina de la señal y producen un señal correspondiente de banda estrecha.

La extracción de la frecuencia natural de una señal de voz ,tradicionalmente llamada "deteccion del Pitch" ,es una de las tareas mas difíciles en el análisis de la voz .

En el sintetizador ,la señal del canal original es utilizada para para controlar la respuesta en frecuencia de un filtro variando en el tiempo (consistente de moduladores y filtros paso banda estrechos ,para corresponder a la medida de envolvente espectral en el analizador. La entrada de este filtro variante en el tiempo es proporcionada por una señal de excitación de espectro liso de la propia estructura espectral fina de la señal (pulsos cuasi-periódicos para las vocales ,o ruido blanco para las consonantes)

3.3 Vocoders en el dominio del tiempo.

En lugar de transmitir la información espectral, muestras de la función de autocorrelación pueden ser utilizadas para caracterizar la voz

El ancho de banda para tales muestras debería de estar espaciado $1/(2 \times 4000 \text{ Hz}) = 0.125$ aparte. Pero el número de muestras debería ser lo bastante grande para asegurar una resolución espectral suficiente en las bajas frecuencias. Así un vocoder por autocorrelación requeriría un ancho de banda mayor para la transmisión de voz que el channel vocoder.

Otro tipo de vocoder en el dominio del tiempo es el llamado Vocoder por expansión ortogonal, en este caso la potencia o el espectro en amplitud, o la función de correlación, o la forma de onda de la señal pueden ser expandidas en una serie de funciones ortogonales. Otro método de representar el espectro de un sonido de la voz es por medio de los coeficientes de predicción lineal, llamado LPC.

CAPITULO 4
EXPERIMENTACION

4.1 Experimentación

Para llevar a la práctica toda la teoría descrita en los capítulos anteriores se hizo un paquete de programas que abarcan algunas de las técnicas antes descritas.

Estos programas fueron realizados en una computadora PDP/1134. Los programas fueron realizados en su mayoría en lenguaje Fortran y algunos programas en Ensamblador.

La forma de poder trabajar con este paquete se describirá a continuación.

Todo el paquete de programas se encuentra en el disco del sistema y cada programa es traído a memoria solamente cuando esta siendo invocado por el usuario.

Para poder trabajar con este paquete lo primero que hay que hacer es instalar los programas en el sistema, esto se hace para que la minicomputadora tenga conocimiento de las tareas que serán ejecutadas. Para instalar el paquete se teclea "SISTEMA".

Para poder ejecutar cualquier comando del paquete es necesario que exista el archivo de datos que contenga la señal a procesar. Si no existe este archivo se puede introducir por medio de los convertidores analógico/digital con los que cuenta la máquina.

Se puede muestrear a diferentes frecuencias (1khz, 2khz, 6.4khz, etc) con 12 bits de entrada y 10 bits de salida.

El paquete maneja dos tipos de archivos uno con formato real que representan el voltaje de la señal al ser muestreada.

, rango 0-1, y el otro con formato entero que sirve para poder ser manejado por otro paquete llamado Interactive Laboratory System (ILS) rango 0-2048.

En este trabajo fueron llevadas acabo las siguientes técnicas:

- 1.- PCM con 1,2,4,9,10 bits.
- 2.- PCM Logarítmica (Ley $\mu=255$).
- 3.- DPCM con 1,2,4,9,10 bits.
- 4.- APCM con 1,2,4,9,10 bits.
- 5.- ADPCM 1,2,4,9,10 bits.
- 6.- SBC.
- 7.- Channel Vocoder.

A continuación se describe la forma de poder ser ejecutada cada técnica y los programas que contiene.

4.1.1 PCM

Para que esta técnica pueda ser ejecutada es necesario crear un archivo que contenga los intervalos de cuantización para un número de bits determinado (10,9,4,2,1). Para esto se corre el programa CNT, el cual pregunta el número de bits y la amplitud máxima. Ejemplo:

```

CNT
?
10
10

```

Solamente es necesario crear los intervalos de cuantización una sola vez cuando se genera el sistema o cuando se quiere modificar la amplitud de los intervalos de cuantización. Estos

archivos que contienen los intervalos de cuantización también se utilizan en las demás técnicas (APCM, DPCM, ADPCM, etc).

Estos archivos serán utilizados por el programa PCM, que hace un mapeo entre las muestras de entrada y su correspondiente salida. El archivo que contiene los intervalos de cuantización puede ser visto como una tabla donde para cada valor de entrada se tiene un valor de salida. La forma en que se hace la búsqueda de un determinado valor de entrada en la tabla es utilizando el algoritmo de búsqueda binaria.

A continuación se describe este algoritmo:

B1. [Inicialización] $l \leftarrow 1$, $u \leftarrow$ Número de Intervalos

B2. [Se obtiene el punto medio] Si $u < 1$, el algoritmo termina en caso contrario, $i \leftarrow [(l+u) / 2]$ el punto medio aproximado del área relevante de la tabla.

B3. [Comparación] Si $K < K_i$, go to B4; si $K > K_i$ go to B5 y si $K=K_i$, el algoritmo termina.

B4. [Ajuste de U] $u \leftarrow i-1$ y regresa a B2

B5. [Ajuste de l] $l \leftarrow i+1$ y regresa a B2

donde K es el valor de la muestra de entrada y K_i es el valor de la muestra de salida.

El PCM puede ser ejecutado con 1,2,4,9,10 bits. Una vez que se cuenta con el archivo de datos muestreado para ejecutar la técnica PCM se hace lo siguiente:

Se teclea PCM, y en seguida la minicomputadora preguntara por el nombre del archivo a cuantizar y el número de bits de cuantización. A continuación se muestra un ejemplo:

PCM

FILE NAME? E20/2239ND002.BAT

FILE FORMAT?

1.-119

2.-VOLTS

1

NUMBER OF QUANTIZATION BITS 10

FILE PCM.DAT1 CONTAINS QUANTIZED DATA

El archivo PCM.DAT contiene el resultado del proceso.

4.1.2 PCM Logarítmica (lev mu=255)

Así como en el PCM es necesario crear los archivos que contienen los intervalos de cuantización pero ahora con una característica logarítmica.

Para realizar esta función se cuenta con el programa LEY que crea estos archivos.

Este programa crea dos archivos ,uno que es utilizado cuando se comprime la señal y el otro cuando se expande. La característica de los archivos depende del grado de compresión (μ) y de el valor máximo que puedan tomar las muestras. A continuación se muestra un ejemplo de como corre este programa:

```

LEW
POR EL GRADO DE COMPRESION -> 255
POR EL VOLTAJE MAXIMO -> 1.

```

Este programa solamente es necesario correrlo cuando se

genera el sistema o cuando se quiere otro grado de compresión o tamaño de la amplitud

Se puede realizar un PCM logarítmica con 10 bits y cualquier grado de compresión (típico $\mu=255$). PCM logarítmica cuenta con dos programas para ser ejecutada CMP y EXP. El primer programa comprime la señal y la codifica, el segundo expande la señal y la decodifica.

Al igual que en PCM la búsqueda se hace por medio del algoritmo de búsqueda binaria. A continuación se muestra un ejemplo de como realizar esta tecnica:

Primero se comprime y se codifica

CMP

FILE NAME? [60,62]SNDD02.DAT
FILE FORMAT?
1.-ILS
2.-VOLTS

1

FILE [60,62]COMPMU.DAT#1 CONTAINS COMPRES DATA

2

Se expande y se decodifica

>EXP

FILE NAME? [60,62]COMPMU.DAT
FILE FORMAT?
1.-ILS
2.-VOLTS

1

FILE [60,62]EXPAMU.DAT#1 CONTAINS EXPAN DATA

2

El archivo EXPAMU.DAT contiene el resultado del proceso.

4.1.3 APCM

Para realizar esta técnica se requiere que el archivo de datos a procesar este en formato real (volts). Si el archivo se encuentra en formato entero (ILS) se debe de correr otro programa que convierta el archivo de formato entero a formato real. Este programa se llama VOL.

APCM hace uso también de los archivos que son creados por CNT (creación de los int. de cuantización).

Se puede realizar un APCM con 1,2,4,8,10 bits. Esta técnica cuenta con dos programas para ser ejecutada APC y APR. El primer programa hace la función de codificador y el segundo hace la función del decodificador. A continuación se muestra un ejemplo:

Primero se codifica:

```
APC
FILE NAME? E1,613CONVOL.DAT
NUMBER OF QUANTIZATION BITS 4
FILE APCM.DAT1 CONTAINS APCM DATA
```

Se decodifica:

```
APR
FILE NAME? E1,613APCM.DAT
FILE APCMREF.DAT1 CONTAINS ORIGINAL DATA
```

El archivo APCMREF.DAT contiene el resultado del proceso.

4.1.4 DPCM

Al igual que APCM para realizar esta técnica se requiere que el archivo de datos a procesar este en formato real (volts).

DPCM hace uso también de los archivos que son creados por CNT (creación de los int. de cuantización).

Se puede realizar un DPCM con 1,2,4,9,10 bits. Esta técnica cuenta con dos programas para ser ejecutada DPC y DPR. El primer programa hace la función de codificador y el segundo hace la función del decodificador. A continuación se muestra un ejemplo:

Primero se codifica:

>DPC

FILE NAME? [1,61]CONVOL.DAT

NUMBER OF QUANTIZATION BITS 4

FILE [60,62]DPCM.DAT#1 CONTAINS DPCM DATA

Se decodifica:

>DPR

FILE NAME? [60,62]DPCM.DAT

FILE [60,62]DPCMRE.DAT#1 CONTAINS DPCM DATA

El archivo DPCMRE.DAT contiene el resultado del proceso.

Cuando se trata de la modulación delta solo es necesario correr el programa que codifica. A continuación se muestra un ejemplo:

>DPC

```
FILE NAME? E1,611CONVOL.DAT
NUMBER OF QUANTIZATION BITS 1
FILE [60,62]DPCM.DAT#1 CONTAINS DPCM DATA
```

El archivo DPCM.DAT contiene el resultado del proceso.

4.1.5 ADPCM

Al igual que APCM para realizar esta técnica se requiere que el archivo de datos a procesar este en formato real (volts).

ADPCM hace uso también de los archivos que son creados por CNT (creación de los int. de cuantización).

Se puede realizar un ADPCM con 1,2,4,9,10 bits. Esta técnica cuenta con dos programas para ser ejecutada ADP y ADR. El primer programa hace la función de codificador y el segundo hace la función del decodificador. A continuación se muestra un ejemplo:

Primeramente se codifica:

>ADP

```
FILE NAME? E1,611CONVOL.DAT
NUMBER OF QUANTIZATION BITS 1
FILE [60,62]ADPCM.DAT#1 CONTAINS DPCM DATA
```

Se decodifica:

```
>ADP
```

```
FILE NAME? C607621ADPCM.DAT
```

```
FILE C607621ADPCMRE.DAT#1 CONTAINS DPCM DATA
```

El archivo ADPCMRE.AT contiene el resultado del proceso.

Cuando se trata de la modulación delta adaptiva solo es necesario correr el programa que codifica. A continuación se muestra un ejemplo:

```
>ADP
```

```
FILE NAME? E17413CONVOL.DAT
```

```
NUMBER OF QUANTIZATION BITS 1
```

```
FILE C607621ADPCM.DAT#1 CONTAINS DPCM DATA
```

El archivo ADPCM.DAT contiene el resultado del proceso.

4.1.6 SBC

Para poder ejecutar esta técnica es necesario crear los diferentes filtros paso banda por los cuales la señal será filtrada. Para poder realizar esta operación se recurre al paquete de ILS que crea los filtros. Se cuenta con un archivo de comandos indirectos que realiza esta operación.

No es necesario crear estos filtros cada vez que se ejecuta SBC, solamente basta hacerlo una vez. La creación es automática, el usuario no necesita saber nada del ILS. Para ejecutarlo se tecléa @FIPABA, preguntará solamente por la frecuencia a la que fue muestreado el archivo de entrada.

Ya que se cuenta con los filtros se puede ejecutar SBC. También se hace uso del paquete ILS para filtrar la señal. La técnica SBC cuenta con dos archivos de comandos indirectos, el

primer archivo SBC realiza el filtrado ,la decimación (por medio del programa Consb),la codificación (por medio del programa APCMSB), y el segundo archivo realiza la decodificación (APCMSBRE),interpola (programa Ceros),realiza el filtrado y suma todos los archivos (SUMASCB). Para ejecutar el primer archivo de comandos indirectos solo se tecléa @SBC preguntará por el nombre del archivo y por el número de bloques con los que cuenta. Para ejecutar la segunda parte se tecléa @SBCRE y solamente preguntará por el número de bloques. El archivo SBCRE.DAT contiene el resultado del proceso

4.1.7 Channel Vocoder

Al igual que el SBC antes de ejecutar el C.V es necesario crear los filtros que serán utilizados por esta técnica. Para poder realizar esto hay que correr el siguiente archivo de comandos indirectos @chafilt ,que solamente preguntará por la frecuencia a la que fue muestreado el archivo de entrada.

Una vez que se cuenta con los filtros se puede ejecutar C.V. , se cuenta con dos partes ,en la primera se simula la transmisión ,se hace el filtrado ,se rectifica la señal (programa rectif),se vuelve a filtrar ,se codifica (PCMCHA),utilizando comandos del paquete ILS se hace la detección de vocal/consonante y la detección del pitch. En la segunda parte se hace la recepción ,primero se modula las diferentes bandas con un generador de pulsos o con un generador de ruido dependiendo del detector vocal/consonante. Después se filtra cada banda y al último se suman todas las bandas (Sumascha) ,para correr la primera parte se tecléa @CHA se pregunta por el nombre del archivo. Para correr la segunda parte se tecléa @CHR y pregunta solamente por el tamaño del archivo. El archivo CHANNEL.DAT contiene la señal procesada.

CAPITULO 5

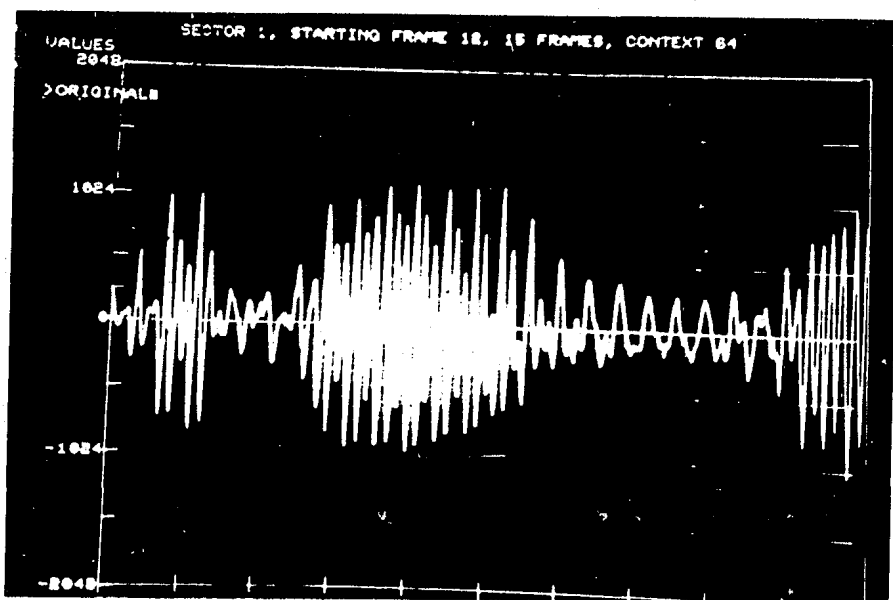
RESULTADOS

5.1 Resultados

Los experimentos antes descritos fueron realizados con una señal de voz muestreada a 1200 Hz de 100 bloques con 512 muestras por bloque.

Desafortunadamente las técnicas antes descritas no fue posible aplicar en los convertidores digital/análogos ya que en el momento que fueron llevados los experimentos a cabo estos no se encontraban disponibles en nuestra máquina. Por este motivo se presentan los resultados gráficamente. Estas gráficas fueron obtenidas de un graficador electrónico conectado a nuestra minicomputadora.

Continuación se muestran las gráficas que representan una parte de las señales obtenidas. Se muestra también la relación señal a ruido de cada uno de ellas con respecto a la original. Se graficaron 460 muestras de la señal empezando desde la muestra 700.



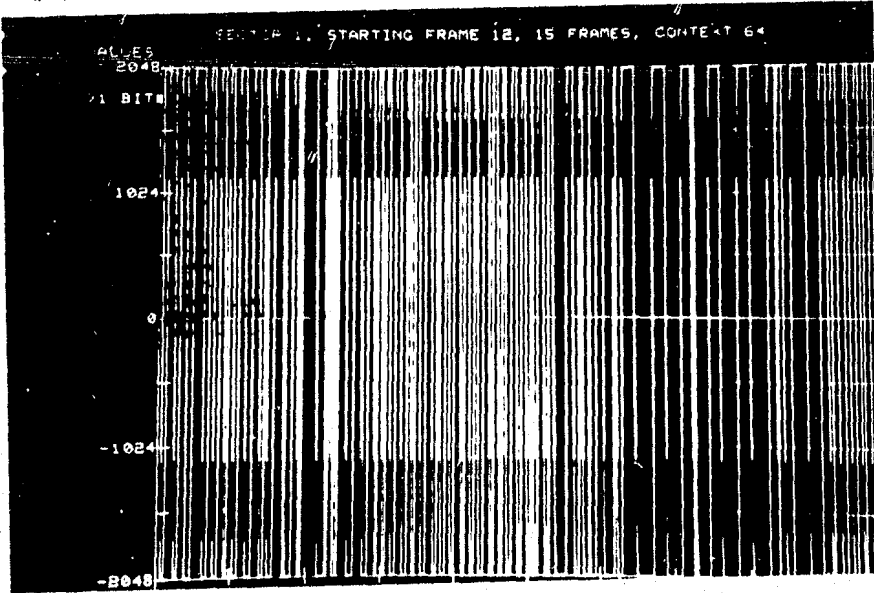


FIG 5.2 PCM 1 bit

SMP = -33.54685

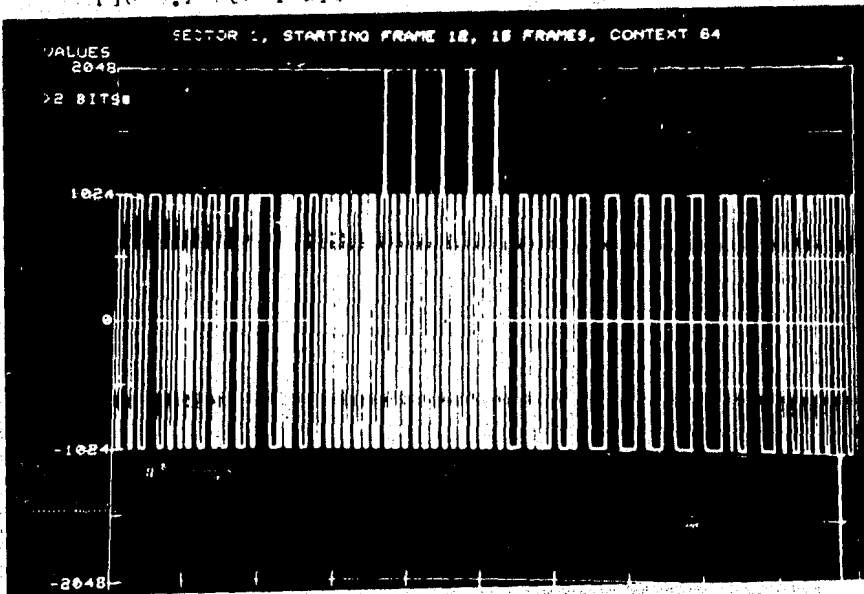


FIG 5.3 PCM 2 bits

SMP = -13.73211

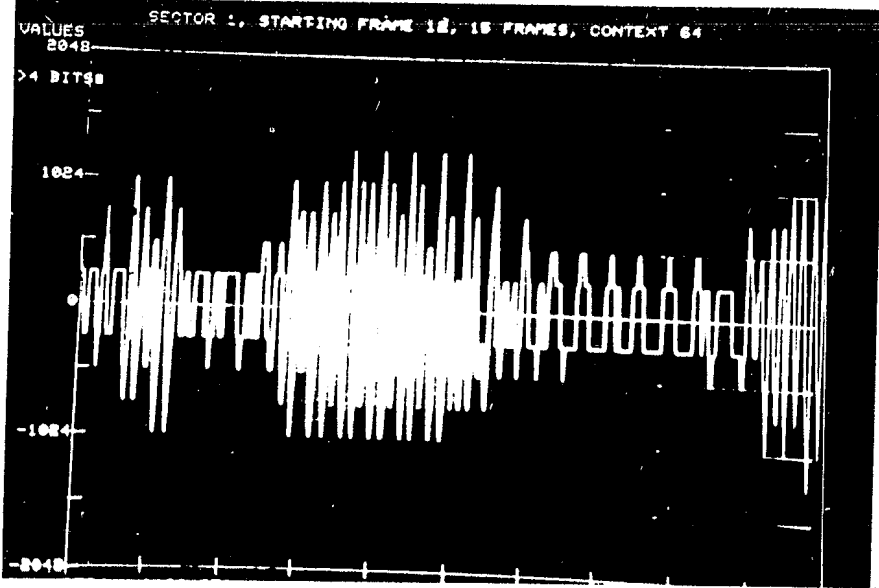


Fig 5.5 PC4 4 bits

SNR = 13.73211

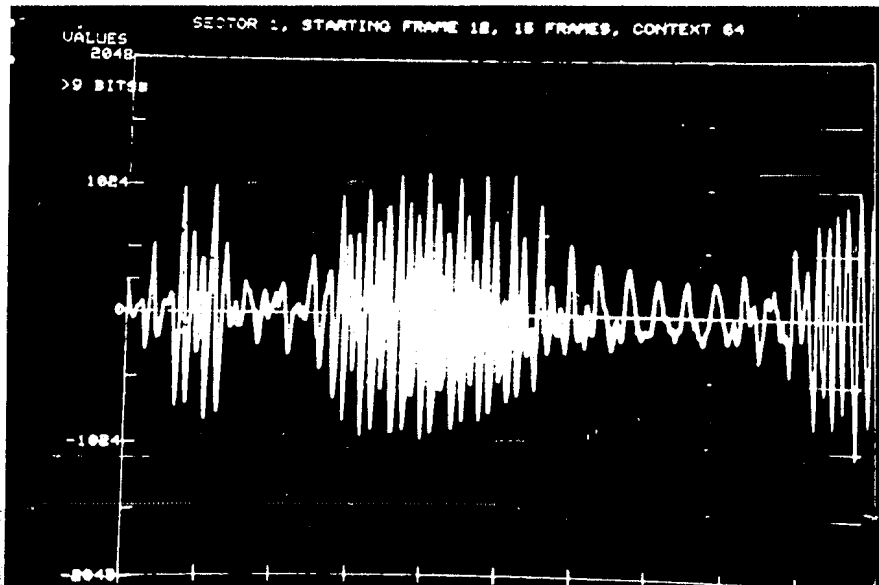


Fig 5.5 PC4 9 bits

SNR = 87.97646

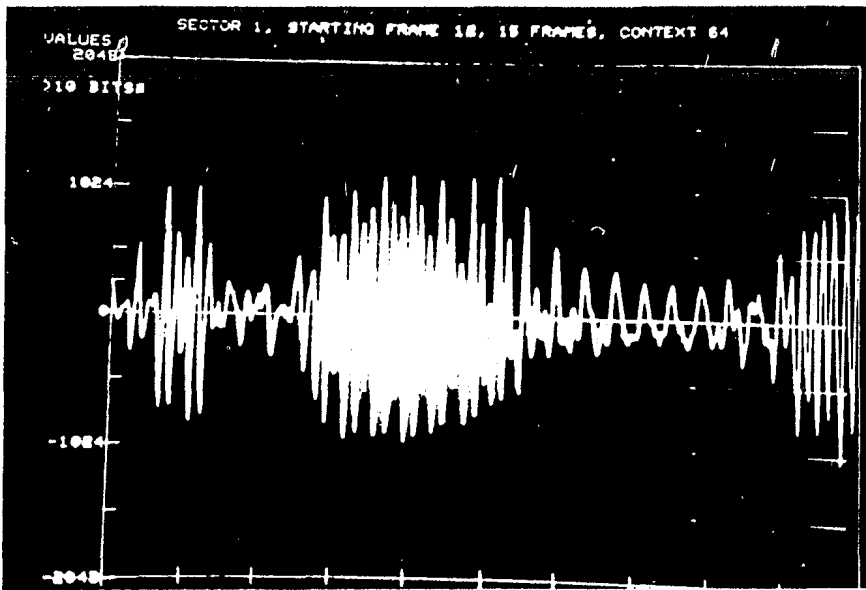


Fig 5.5 PC= 10 bits SNR =103.1892

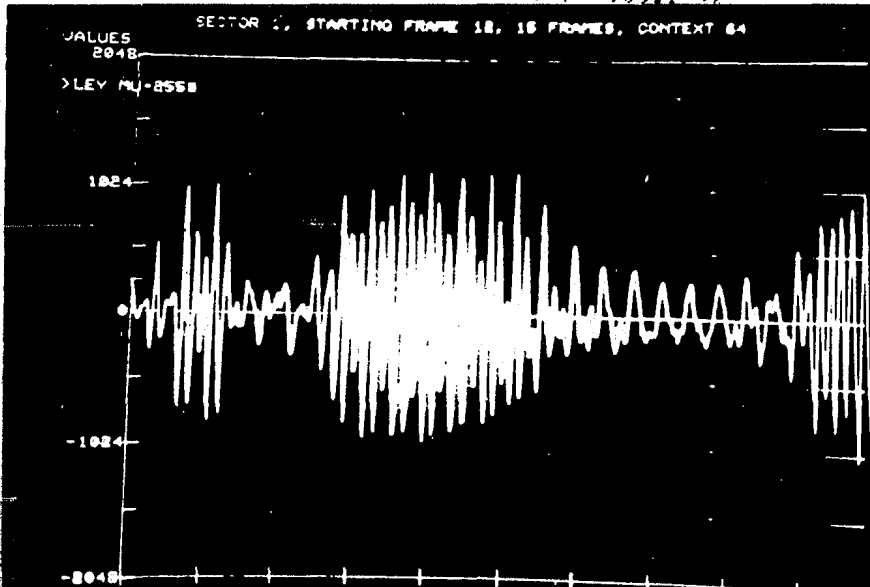


Fig 5.7 PC= Logarítmica lev mu=255 SNR =104.2153

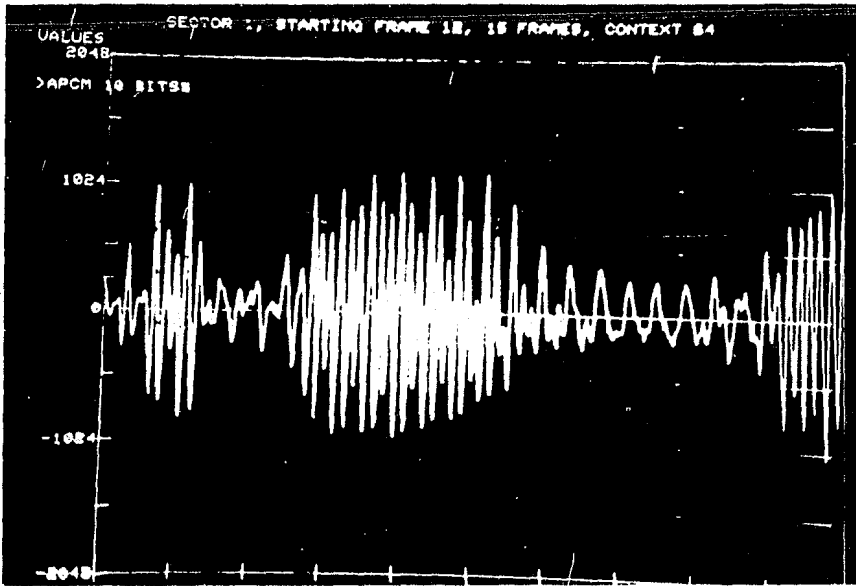


Fig 5.8 APCH 4 bits

SNR = 38.69196

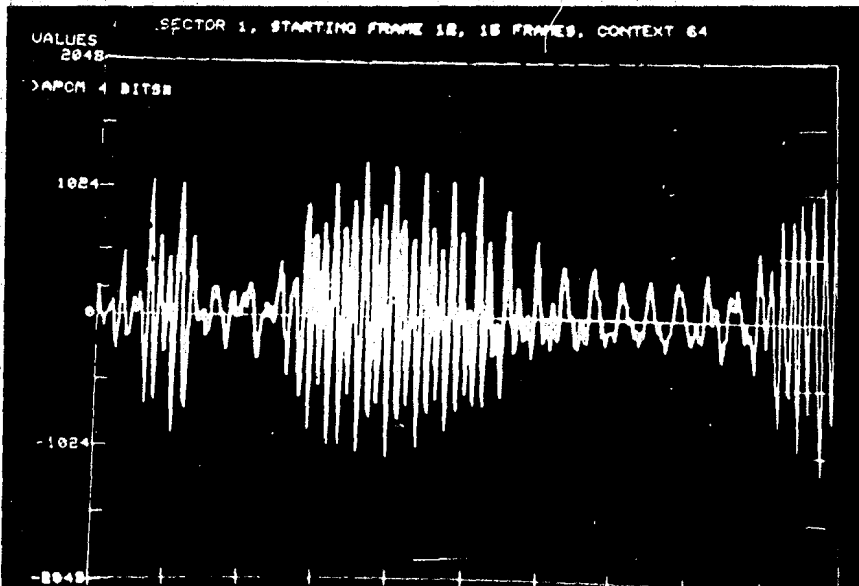


Fig 5.9 APCH 10 bits

SNR = 70.61601

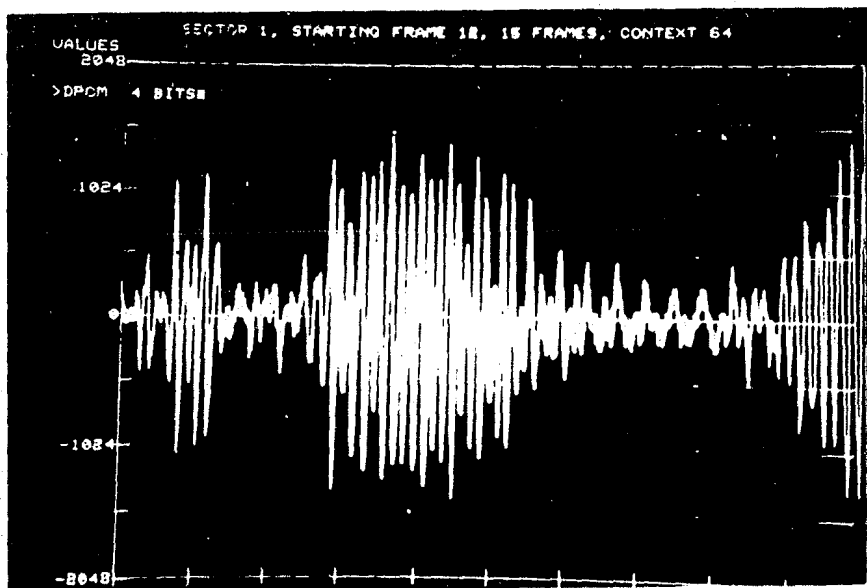


Fig 5.10 DPCM 4 bits

SNR = 16.76165

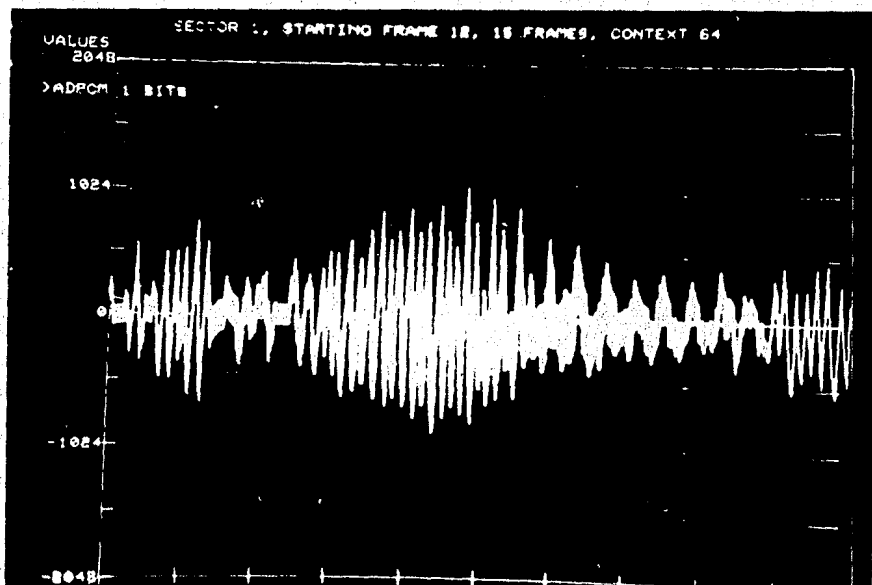


Fig 5.11 ADPCM 1 bit

SNR = 8.240221

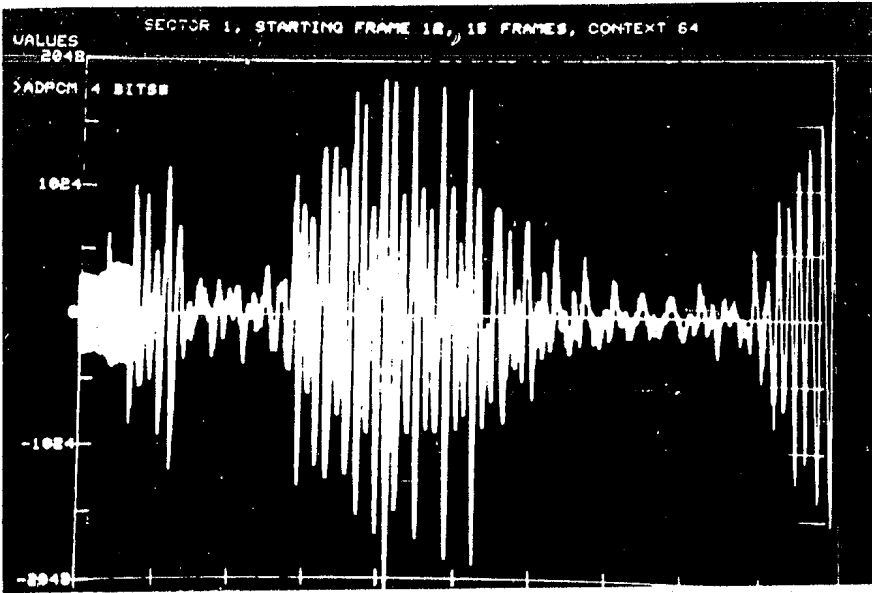


Fig 5.12 ADECY 4 bits

SNR = 3.69870

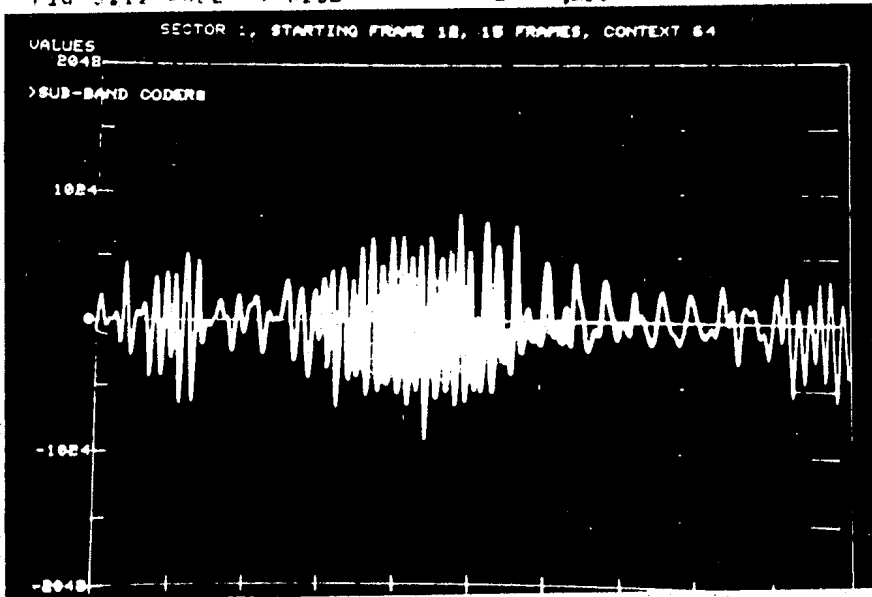


Fig 5.13 SRC

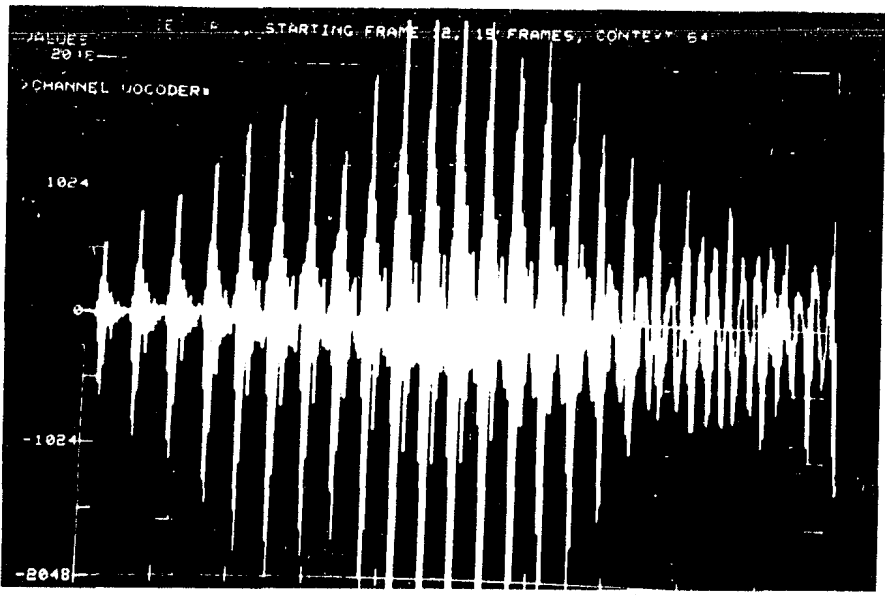
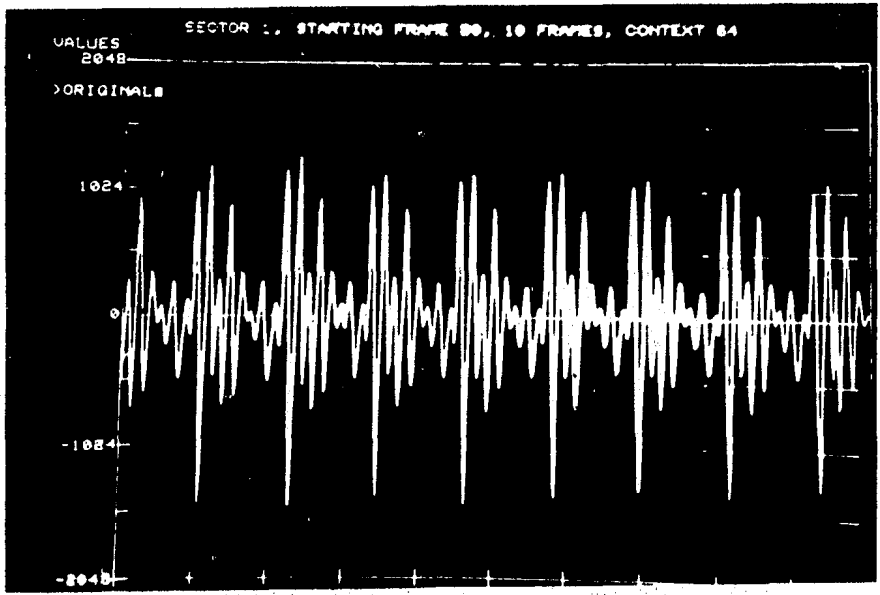


Fig 5.11 CHV

A continuación se muestra otra parte de la señal.



SEÑAL ORIGINAL

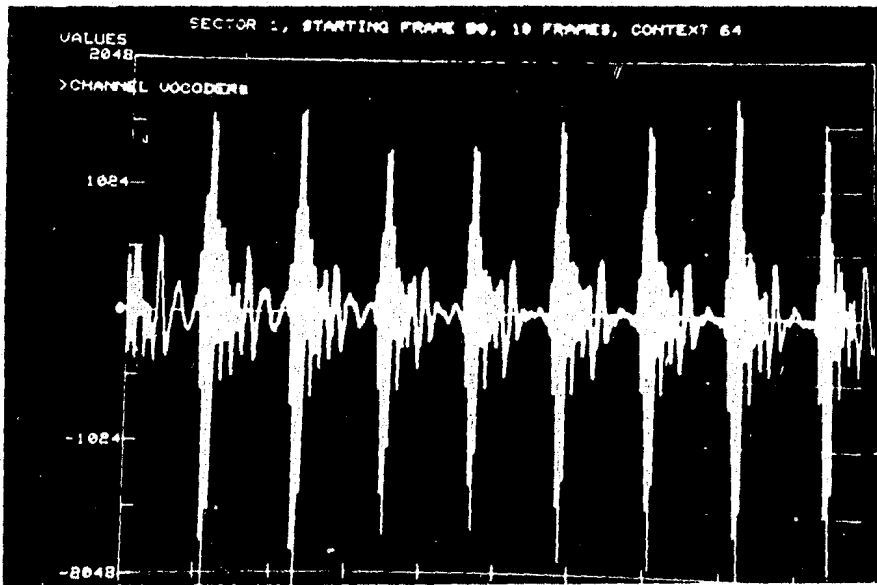


Fig 5.16 C-V

Para un PCM con 1 y 2 bits el reconocimiento de la señal es prácticamente imposible. A medida que se aumenta el número de bits la señal se va pareciendo cada vez más. Para 4 y 10 bits se tiene una aproximación muy buena.

Como se ve en los resultados gráficos la técnica que más se aproxima a la señal original y la que tiene una relación señal a ruido mayor es PCM logarítmica (lev mu=255).

Para tener un mejor resultado en la modulación delta adaptiva hay que muestrear con una frecuencia mayor que el doble del ancho de banda de la señal de entrada.

En el caso del SAC se podrían obtener mejores resultados si se tuvieran filtros digitales mas eficientes. Es el mismo caso que para el Channel Vocoder.

Indice Alfabético :

Búsqueda binaria	4-4
Channel Vocoder	3-7
Cod. en el dom. de la frecuencia.....	2-86
Codificación	2-47
Codificadores de forma de onda.....	2-1
Compresión	2-58
Conversión analógica-digital.....	2-2
Cuantización	2-15
Cuantización robusta	2-42
Cuantización uniforme	2-15
DPCM.....	2-70
Error de Cuantización	2-19
Fricativas	3-3
Filtro Paso Bajas.....	2-4
Muestreo	2-5
Modulación delta	2-79
Modulación delta adaptiva.....	2-82
PCM.....	2-52
PCM logarítmica	2-58
Pitch.....	3-9
Prefiltrado	2-3
Relación señal a ruido	2-20
RUIDO.....	2-19
Ruido de sobrecarga	2-80
SBC	2-87
Vocoders.....	3-1
Vocoders dom. de la frecuencia	3-6
Vocoders en el dom. del tiempo.....	3-10

Bibliografía

- [1] James Flanagan ,M. Schroeder . Speech Coding
IEEE Transactions on Communications ,
Vol.com-27,no 4 April 1979 pp 710-735
- [2] P.F. Panter ,w.Dite Quantization Distortion
in Pulse-Count Modulation with Nonuniform
Spacing of Levels.
Proceedings of the I.R.E ,39,44-48 (Jan.1951)
- [3] L.R Rabiner ,R.W Schafer Digital Processing of
Speech Signals.
- [4] R.A. Gabel ,R.A.Roberts Señales y Sistemas Lineales
Editorial Limusa Mexico 1975
- [5] M.Schwartz Transmisión de Información ,modulación
y ruido . Primero Edición McGraw-Hill de Mexico
1983.

Bibliografia

- [6] C.L.Liu ,Jane W.S.Liu Linear System Analysis
International Student Edition 1975 Mcgraw-Hill.**