



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

Escuela Nacional de Estudios Profesionales
"A R A G O N"



SINTESIS DE VOZ MEDIANTE TECNICA DE COM-
PRESION DIGITAL ADPCM (MODULACION POR
CODIFICACION DE PULSOS DIFERENCIAL
ADAPTIVO)

57372

T E S I S
QUE PARA OBTENER EL TITULO DE
INGENIERO MECANICO
E L E C T R I C I S T A
P R E S E N T A :
SILVIA VEGA MUYTOY

1987



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 esta 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.

El que yo sea un hombre
eso lo comparto con
otros hombres.

El que vea y oiga y
el que coma y beba
es lo que por igual hacen
todos los animales.

Pero el que yo sea yo, es
mío exclusivamente
y me pertenece
y a nadie más,
a ningún otro hombre,
ni a un ángel ni a Dios:
Excepto en cuanto
Yo soy uno con El.

M.E

AGRADECIMIENTOS

- * A la vida porque me ha permitido realizar actividades que llenan el espíritu y el cuerpo.

- * A mis padres porque me ayudaron a fortalecer mi espíritu.

- * A mis hermanos y familiares que forman mi gran universo.

- * A mis amigos por los momentos compartidos.

- * A ti que has llegado ha ser tan especial.

- * A mis profesores que con responsabilidad transmitieron sus conocimientos y ayudaron en mi formación profesional (Especialmente a mi asesor Ing. Raúl Ramírez Román).

- * A mi escuela por todo lo que significa.

POR MI RAZA HABLARA EL ESPIRITU

DEDICATORIA

°A los que el azar hizo que estuviéramos
conviviendo de una u otra forma en este
momento del tiempo y del espacio, para
tener logros y lazos que forman las ca-
denas invisibles que nos unen.

°A los que creyeron y confiaron en mi,
porque su fé vigorizó mi esfuerzo
y mi espíritu.

°A los que desconfiaron de mi, por-
que sus dudas constituyeron un reto
que salvar.

°A los inconformes, a los que quieren
transformar, a los que tienen volun-
tad de hacer y no poseer, a los que
tienen rebeldía frente a lo indesea-
ble porque son los que guiarán el
futuro.

RECONOCIMIENTOS

Por la valiosa colaboración que brindó en la elaboración del presente trabajo con su tiempo, sus comentarios y puntos de vista al Ing. Armando Mota Aguilar de doy las más sinceras gracias.

Agradezco profundamente al Ing. Marino Sánchez Parra por toda la ayuda que me brindó.

PREFACIO

El presente trabajo además de tener el objetivo de obtener el título de Ingeniero Mecánico Electricista, es una introducción a los sistemas de síntesis de voz así como a los sistemas de procesamiento digital de señales, para todos aquellos estudiantes que se interesen en el tema.

El trabajo sigue una secuencia para poder entender más fácilmente lo que significa el procesamiento digital de señales (capítulo 1). Y de aquí, se pasa a las clases de técnicas digitales para implementar síntesis de voz, así como una explicación de las categorías de cada una de ellas (capítulo 2 y 3).

Una técnica de compresión de señales actualmente desarrollándose ampliamente es el ADPCM (Modulación por Codificación de Pulsos Diferencial Adaptivo), por lo que su utilización viene siendo una realidad (capítulos 4 y 5) Así, que teniendo la facilidad de tener esta técnica en un chip puede hacerse realidad un sintetizador de voz a bajo costo (capítulo 6)

Para complementar todo lo visto en los seis capítulos constituyentes del trabajo se realizaron tres apéndices como herramientas auxiliares.

I N D I C E

Agradecimientos	
Dedicatoria	
Reconocimientos	
Prefacio	
Indice	
Introducción	2
Objetivo	13
<u>CAPITULO 1</u>	
Procesamiento Digital de Señales	16
Muestreo	17
Muestreo de Nyquist	18
Teorema de Muestreo en Forma General	22
Decimación e Interpolación de Formas de Onda Muestreadas	27
Reconstrucción o Demodulación de Señales Muestreadas	34
Cuantización	40
Ruido Cuantizado y Cuantización Uniforme	43
Proceso de Compresión y Expansión (Companding) de una Señal	48
Cuantización Logarítmica o Leyes Companding	50
Filtros Digitales	59
Sistemas FIR	62
SISTEMAS IIR	62
Ventajas y Desventajas de los Filtros FIR e IIR	63
Resumen	67
<u>CAPITULO 2</u>	
Técnicas Digitales para Implementar Síntesis de Voz	69
Tamaño de las Unidades de Voz	71
Propiedades Básicas de la Voz	72
Métodos y Técnicas de Síntesis	76

Codificación de Forma de Onda	77
Codificación de la Fuente (Vocoding) o Paramétricos	80
Codificación Híbrida	85
Sistemas de Texto Ilimitado (Texto a Voz) y Sistemas de Texto Limitado (Respuesta Sonora)	87
Calidad de la Voz	88
Resumen	90

CAPITULO 3

Técnicas de Compresión de Señales de Forma de Onda	92
Muestreo y Cuantización de Señales de Voz	92
Modulación por Codificación de Pulsos (PCM)	96
PCM Lineal	96
PCM Logarítmico	99
Modulación Delta (DM)	101
Modulación Delta Adaptiva (ADM)	108
Comparación de los Sistemas PCM Logarítmico, LDM y ADM	112
Modulación por Codificación de Pulsos Diferencial (DPCM)	113
Modulación por Codificación de Pulsos Diferencial Adaptivo (ADPCM)	116
Modulación Delta con Pendiente que Varía Continúamente (CVSD)	118
Codificación en Sub-Bandas (SBC)	120
Resumen	123

CAPITULO 4

Técnica de Compresión de Señales mediante ADPCM	126
Cuantización Adaptiva	126
Cuantización Adaptiva de Alimentación Adelantada	129
Cuantización Adaptiva Regenerativa	134
Cuantización Diferencial	137

Sistema ADPCM con Cuantización Adaptiva	141
Sistema ADPCM con Cuantización y Predicción Adaptiva	144
Conversión Directa de Código Digital	150
Resumen	151
▼ <u>CAPITULO 5</u>	
Estado Actual de la Tecnología para Implementar Síntesis de Voz	158
Compromiso entre Complejidad y Funcionamiento	162
Resumen	162
▼ <u>CAPITULO 6</u>	
Diseño de un Simulador para Evaluación de Voz mediante ADPCM	164
Sincronización	166
Limitación del Ancho de Banda	166
Digitalización de la Señal (Muestreo, Cuantización y Codificación)	168
Análisis y Síntesis	172
Almacenamiento y Reproducción	176
Resumen	183
▼ <u>Conclusiones</u>	185
▼ <u>Apéndice A</u>	192
▼ <u>Apéndice B</u>	206
▼ <u>Apéndice C</u>	213
▼ <u>Bibliografía</u>	247

SINTESIS DE VOZ
MEDIANTE TECNICA
DE COMPRESION
DIGITAL ADPCM
(MODULACION POR CODI
FICACION DE PULSOS
DIFERENCIAL ADAPTIVO)

I
N
T
R
O
D
U
C
T
I
O
N

INTRODUCCION

El avance tecnológico ha hecho que las computadoras en la sociedad moderna sean ampliamente utilizadas. El procesamiento digital de voz promete revolucionar los caminos de interacción del hombre con las máquinas, es decir, dotando a las máquinas con la habilidad para interpretar y actuar por medio de instrucciones habladas. Dado esto, el reconocimiento de voz, se está convirtiendo en una herramienta indispensable para poder lograr la automatización. La síntesis de voz, que da voz a un dato digital, está progresando rápidamente con lo que se liberará a los usuarios del uso de las pantallas y teclados. Por tanto, el desarrollo de las técnicas para producir voz han ido en aumento.

EL HABLA COMO MEDIO DE COMUNICACION

Los seres humanos utilizan principalmente el habla como medio de comunicación, siendo éste el único ser viviente que ha adquirido y usado el habla con éste fin, y para producir la voz ha utilizado el aparato fisiológico que tiene como propósito vital, comer y respirar. Esto ha sido posible debido a que se encuentra envuelto en una atmósfera de aire, con lo que se han tenido las condiciones para aprender a emitir señales producidas por vibraciones longitudinales en el aire, por tanto, la voz consiste en rápidas y determinadas fluctuaciones en la presión del aire. El proceso de comunicación por medio del habla se lleva a cabo cuando los sonidos y presiones del aire son generados y emitidos por el aparato vocal del orador y captados por el oído y transmitido al cerebro para su comprensión por el escucha, como se aprecia en la figura 1.

- 1o.-Formulación del Mensaje
- 2o.-Código del Lenguaje
- 3o.- Acción Neuro-Muscular
- 4o.- Cuerdas Vocales
- 5o.- Tracto Vocal

- 1e.-Oído
- 2e.- Transductor Neural
- 3e.- Código del Lenguaje
- 4e.- Comprensión del Mensaje

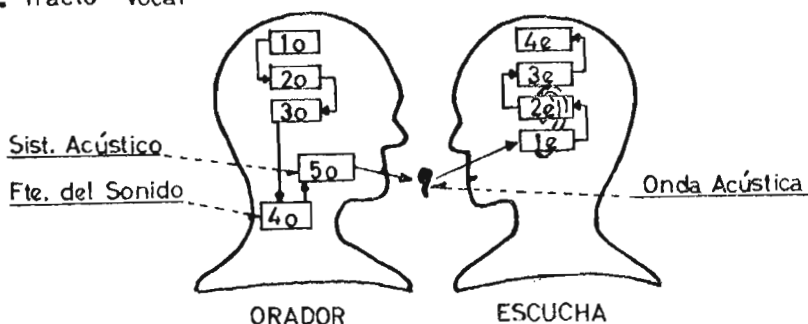


Fig. 1.- Proceso de comunicación por medio del habla.

Las ondas acústicas producidas por el hombre son las vibraciones longitudinales producidas cuando la cavidad torácica se expande y contrae para forzar el aire desde los pulmones hacia afuera a través de la traquea al glotis.

Si las cuerdas vocales están tensadas, el flujo de aire producirá vibración en las cuerdas produciendo pulsos cuasi-periódicos. Si las cuerdas vocales son extendidas separadamente, el aire pasa fluyendo a través del glotis produciendo un flujo de aire turbulento cuasi-aleatorio. El aire fluyendo a través de la faringe pasa al tracto vocal, dependiendo de la posición de los articuladores (labios, lengua, etc.) y del escotillón del paladar (paladar suave), el aire fluirá por la boca, por la nariz o por ambos y la voz será emitida.

En la figura 2 se muestra un diagrama esquemático de la producción de voz humana y el plano sagital de rayos X de la cabeza humana, donde es posible apreciar el tracto

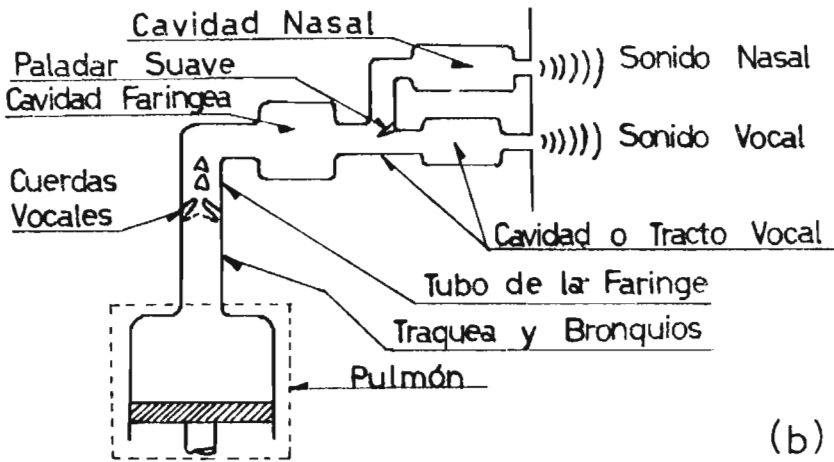
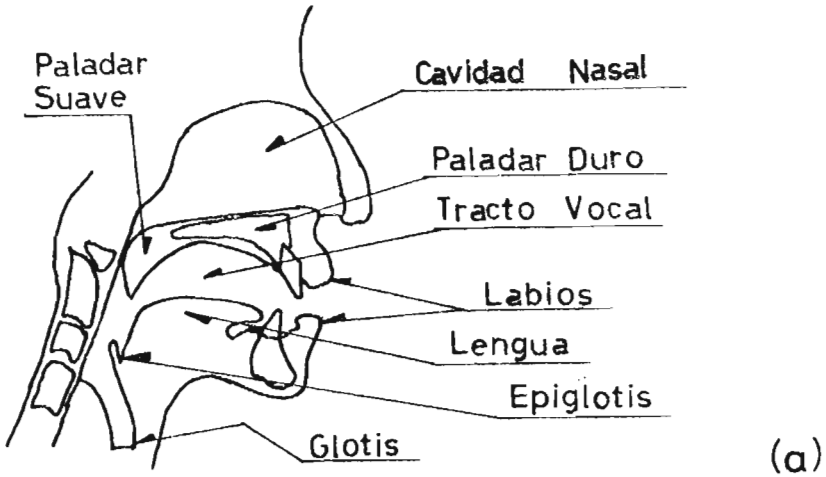


Fig. 2. - a) Plano sagital de rayos X de la cabeza humana.
b) Diagrama esquemático del mecanismo de producción de voz humana.

vocal que es un tubo acústico no uniforme que se extiende desde la glotis a los labios, con una longitud de alrededor de 17 cm en un adulto. Este es de una área de sección transversal no uniforme que depende de la posición de los articuladores y varía desde cero (completamente cerrada) a cerca de 20 cm^2 . El tracto nasal se extiende desde el paladar hasta las narices, teniendo una longitud aproximada de 12 cm y un volumen de 60 cm^3 .

El progreso y desarrollo del habla humana desde su forma primitiva como gruñidos y chillidos hasta el estado donde se encuentra como un código acústico especializado ha sido a través de un periodo largo de tiempo, pues se especula que los primeros hombres se comunicaban por medio de las señales de las manos, con lo que una de las hipótesis ampliamente difundidas es de Sir Richard Paget "el hombre manejó la invención del habla cuando sintió la necesidad de expresar sus pensamientos teniendo la dificultad de hablar con sus manos"

A través del paso de la historia, la producción del habla artificial ha sido de gran interés para el hombre, por lo que han surgido diversas formas de generarla, como puede apreciarse en las antiguas civilizaciones griega y romana en donde existieron estatuas que hablaban, voces milagrosas y oráculos en donde la voz era emitida por el encargado del templo o por el sacerdote y transmitida por medio de un tubo fonético oculto. Con el paso del tiempo, los cambios sociológicos y culturales dieron paso a mecanismos con fuelle. El intento de imitar voz humana no ha cesado y actualmente existen una gran cantidad de técnicas para poder realizarlo.

La comunicación es el proceso por medio del cual la

información se transfiere de la fuente al destino, por tanto existen sistemas de comunicación que son un conjunto de mecanismos que proporcionan el enlace para la información entre destino y fuente (transmisor, canal de transmisión, receptor). Para caracterizar la voz en la comunicación se pueden mencionar dos formas; La primera, es una idea introducida por Shannon, en donde la voz es representada en términos del contenido de un mensaje o información (por ejemplo, 10 fonemas por segundo) llamada TEORIA DE LA INFORMACION. La segunda forma es caracterizando la voz en términos de SEÑAL (onda acústica producida por voz o música) llevando la información del mensaje.

La información contenida en la voz es intrinsecamente de naturaleza discreta, es decir, está compuesta de una secuencia de sonidos, los cuales pueden ser clasificados fonemas los que son una concatenación de elementos. Estos sonidos y la transición entre ellos sorve como una representación de la información. El arreglo de dichos sonidos son gobernados por las reglas del lenguaje. El estudio de dichas reglas y sus implicaciones en la comunicación humana es del dominio de la lingüística y el estudio y clasificación de los sonidos de la voz se llama fonética.

El aspecto técnico del sistema de comunicaciones trae como consecuencia una variedad de representaciones de la señal de voz, sin embargo, existen dos características importantes que se deben mantener en cualquier sistema:

- 1.- Conservación del contenido del mensaje.
- 2.- La representación de la señal de voz debe ser en una forma conveniente para su almacenaje o transmisión.

PROCESAMIENTO DIGITAL DE SEÑALES

El problema general del procesamiento y manipulación de la información involucra dos aspectos importantes: La representación de la señal y la transformación de la señal en una forma conveniente. Por ejemplo, un sistema de identificación de una persona en forma automática de un conjunto de oradores, podría usarse una representación espectral en función del tiempo de la señal de voz. Una transformación posible de la señal puede ser el promedio espectral a través de una palabra o frase, la cual se comparará con el almacenado, y entonces basada sobre la medida espectral similar elige la identidad del interlocutor. Para la señal de voz, la representación se puede encontrar en cualquiera de sus tres formas:

- Forma de onda.
- Forma paramétrica o de la fuente.
- Forma híbrida.

Las técnicas digitales son ampliamente utilizadas para el procesamiento de señales dado que el desarrollo tecnológico, el costo y la flexibilidad de los sistemas y computadoras digitales sobrepasa a los sistemas analógicos. Para el procesamiento de voz existen muchas razones para la utilización de técnicas digitales, entre ellas, la voz en forma digital puede ser transmitida en un canal ruidoso sin degradaciones o ser transmitida simultáneamente con datos en una malla sin necesidad de distinción entre ellas, exceptuando en la decodificación.

PROCESAMIENTO DIGITAL DE LA VOZ

El desarrollo del área del procesamiento digital de señales se ha llevado a la realidad debido a la realiza-

ción del procesamiento digital de la voz. Los problemas generales en donde se ve involucrado el procesamiento digital de la voz puede dividirse en tres clases:

- Análisis de Voz.
- Síntesis de Voz.
- Análisis - Síntesis de Voz.

Cuando nos interesamos únicamente en el ANALISIS DE VOZ, que es simplemente el proceso de estimación de los parámetros (variables en el tiempo) del modelo para producir voz, desde la señal de voz que es asumida para ser la salida del modelo. (18) Los sistemas de análisis de voz sirven usualmente como la "entrada final" para sistemas automáticos de reconocimiento de voz y para sistemas de verificación o identificación automática de un orador.

Cuando nos interesa la SINTESIS DE VOZ, que es el proceso de producción de una señal acústica, controlando el modelo de producción de voz con un conjunto apropiado de parámetros. (18) Los sistemas de síntesis de voz son usualmente la "salida final" en forma de voz de una computadora o de una máquina lectora de un texto escrito.

La tercera clase involucra primero el análisis seguido de la síntesis de voz que es la más utilizada en las aplicaciones del procesamiento digital de la voz. La calidad de los datos extraídos en forma hablada dependerá de las técnicas utilizadas en el procesamiento, así como de la sofisticación de los codificadores y decodificadores, por tanto, si el modelo es lo suficientemente preciso y los parámetros son perfectamente determinados, la salida sonora será indistinguible de la voz natural.

Cuando se tiene determinada la necesidad y el uso del

sistema es cuando sabemos en cual de los tres problemas se está involucrado, así como que método de procesamiento digital de señales será utilizado para representar, procesar y almacenar o transmitir la voz.

En el siguiente cuadro se muestran algunas aplicaciones típicas de la utilización del procesamiento digital de la voz:

Aplicaciones del Procesamiento Digital de la Voz	{	- Codificación de Voz
		- Síntesis de Voz
		- Verificación e Identificación de un orador
		- Reconocimiento de Voz
		- Transmisión de Datos

CODIFICACION DE VOZ

Las técnicas de codificación de voz están siendo ampliamente estudiadas y se han ido realizando mejoras para lograr un uso más eficiente del ancho de banda del canal de transmisiones. También se desarrollan para lograr el almacenaje de más información en menos memoria.

SINTESIS DE VOZ

Los sistemas de síntesis de voz han sido ampliamente estimulados y han logrado rápidos avances debido a la necesidad de un almacenamiento de voz para las computadoras con sistemas de respuesta sonora, para la reproducción de un mensaje previamente codificado y almacenado que puede ser extraído por una persona desde una terminal o teclado o de un sistema telefónico, o bien de un juguete. Por ser un proceso digital tiene las siguientes ventajas sobre las

técnicas analógicas:

- a) La reproducción del mensaje es de una calidad equivalente al original.
- b) No existe desgaste mecánico de los componentes.
- c) El mensaje se reproduce siempre desde su inicio.
- d) Bajo costo.

VERIFICACION E IDENTIFICACION DE UN ORADOR

La verificación e identificación involucra obtener la autenticidad o identificación de un orador de un grupo grande de posibles oradores, así como en los sistemas de verificación debe decidir si el orador es la persona que demanda ser. Esto es potencialmente aplicable en requerimientos de control al acceso de información y en áreas restringidas al igual que en varias clases de transacciones crediticias automáticas.

RECONOCIMIENTO DE VOZ

El reconocimiento de voz es una área de naturaleza multidisciplinaria que es fuertemente dependiente de las restricciones establecidas sobre el orador. Este tipo de sistemas son muchos y muy variados tales como computadoras que responden a la comunicación sonora y máquinas de escribir operadas por voz.

AYUDA A IMPEDIDOS

Una variedad de técnicas de procesamiento de señales están siendo aplicadas para aplicaciones didácticas para facilitar el aprendizaje en general y para dar facilidades a personas impedidas como ciegos y minusválidos.

TRANSMISION DE DATOS

El crecimiento de la demanda para la transmisión de datos en general y principalmente utilizando la red telefónica y debido al incremento de la comunicación vía satélite, micro-ondas y sistemas de comunicación óptico, ha sido necesario que surgan sistemas que digitalizan la voz a bajas razones de bits tanto como sea posible.

O
B
J
E
T
I
V
O

OBJETIVO

El presente trabajo presenta los fundamentos matemáticos básicos para producir voz sintética en forma digital, así como una somera explicación de las técnicas digitales de forma de onda, y la explicación de la producción de voz (síntesis) utilizando la técnica ADPCM (Modulación por Codificación de Pulsos Diferencial Adaptivo), con lo que se obtiene una voz sintética de alta inteligibilidad, conservando la inflexión y entonación de la voz humana, con una baja relación de bits y bajo costo.

C
A
P
I
T
U
L
O

1.- PROCESAMIENTO DIGITAL DE SEÑALES

El desarrollo tecnológico ha hecho que el procesamiento digital de señales (PDS) en los sistemas de comunicaciones sea una alternativa práctica para el procesamiento de señales analógicas. La aplicación de la teoría del PDS como una alternativa viable es debida grandemente a la disponibilidad de circuitos MSI y LSI (integración a mediana y gran escala).

Los tres factores tecnológicos que han intervenido en el incremento de la utilización del PDS son:

- a) La existencia de circuitos integrados para ejecutar la aritmética, el almacenamiento, la sincronización y la selección de operaciones a alta velocidad, bajo consumo de potencia y bajo costo.
- b) El procesador de señales digitales tiene la flexibilidad para ser programado.
- c) El hecho de que las señales siempre están disponibles en forma digital ayuda a los circuitos en la realización óptima y rápida de una operación.

Para la comprensión del procesamiento digital de señales existen dos conceptos básicos: El MUESTREO y la CUANTIZACION. El muestreo es la transformación de datos de una señal continua o una analógica, dando como resultado una serie de valores discretos en el tiempo, siendo esto una representación exacta de la señal original, donde el rango de muestreo F_s será por lo menos dos veces el ancho de banda de la señal, por lo que si existe alguna falla en dicho muestreo resultará una distorsión llamada traslapamiento (aliasing). El procesamiento digital del dato muestreado no puede ocurrir hasta que dicho dato sea cuantizado o cuantifica

do. En el proceso de cuantización, a las muestras de amplitud constante se les asigna a cada una un número binario finito. En la cuantización uniforme los niveles cuánticos o tamaño de la etapa de cuantización son de igual tamaño, por lo que si existen algunas partes de la señal muestreada más grandes, existirá el efecto de recorte (clipping).

La señal muestreada y cuantizada resultante es discreta en tiempo (en virtud del muestreo) y en amplitud (en virtud de la cuantización), por tanto el número de series en tiempo que resulta de este doble proceso pueden ser ahora procesadas usando las capacidades disponibles en la computadora digital. El procesamiento en tiempo real de señales digitales puede ser realizado por una computadora.

1.1.- MUESTREO

Para el uso de métodos de procesamiento digital de señales sobre una señal analógica con ancho de banda limitada, es necesario representar la señal por una secuencia de números o una sucesión de pequeñas muestras.

Conceptualmente un periodo de muestra o muestro instantáneo es: $t = \tau$ que es el valor instantáneo $f(\tau)$ de la función $f(t)$, el cual en sí mismo tiene extensión en tiempo. El muestreo -ideal no es un proceso físicamente realizable u una definición más acorde con la operación práctica puede ser la siguiente: "Una MUESTRA es una medida de la amplitud de una señal, evaluada sobre un periodo de tiempo corto de duración, en el cual la señal cambia solamente en una cantidad insignificante".(17)

En la práctica una muestra tiene una duración pequeña pero no es infinitesimal, por tanto la señal muestreada con

siste de un tren de pulsos, cada uno con un momento igual al valor contemporáneo del original.

En la teoría básica de señales muestreadas hay dos aproximaciones; La primera en donde iniciamos con las propiedades fundamentales de la señal y mostrar que una señal con banda limitada queda definida completamente por un número de valores muestreados, esta aproximación ha sido adoptada por los teóricos de la información y por matemáticos. La segunda es una alternativa más directa desde el punto de vista de un ingeniero de diseño, donde se principia con un tren de pulsos a la frecuencia de muestreo y se considera esto como una clase de onda portadora sobre la cual la señal es impresa como modulación en amplitud.

1.1.1.- MUESTREO DE NYQUIST

Consideremos una señal de variación continua $f(t)$ que será convertida en una forma digital. Primero muestreando $f(t)$ periódicamente a un rango f_c muestras por segundo. El proceso de muestreo se puede visualizar con un interruptor o conmutador como en la figura 1.1, donde a la salida del conmutador obtenemos la señal muestreada. Y donde el interruptor permanece cerrado τ segundos de $f(t)$ y conmutando al rango $f_c = 1/T$, donde T es el intervalo de muestreo y $\tau \ll T$.

La señal muestreada $f_s(t)$ contiene toda la información acerca de $f(t)$, además $f(t)$ únicamente puede ser extraída de $f_s(t)$.

Asumiendo primero que la señal $f(t)$ tiene una banda limitada a B Hz, entonces es fácil mostrar que la señal muestreada no destruye cualquier contenido de información siem

pre que el rango de muestreo $f_c \geq 2B$. El rango de muestreo mínimo de $2B$ veces por segundo es llamado el rango de muestreo de Nyquist y $1/2B$ el intervalo de muestreo de Nyquist (10)

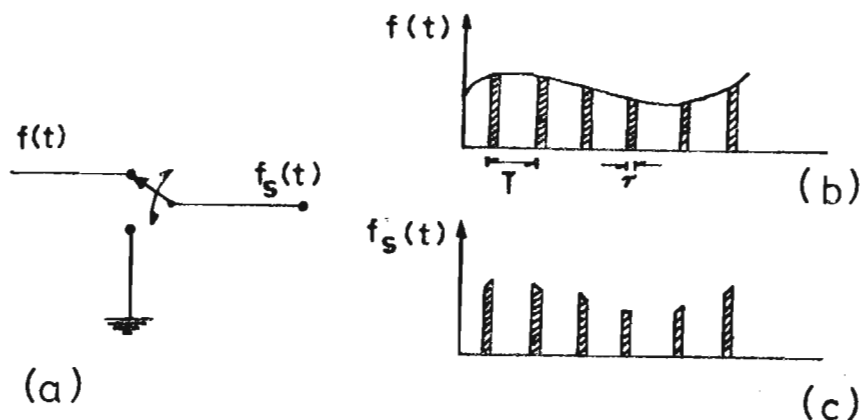


Fig. 1.1.- a) Muestreador de conmutación. b) Señal de entrada $f(t)$. c) Salida muestreada $f_s(t)$.

Lo anterior se puede demostrar por medio del análisis de Fourier

$$f_s(t) = f(t) s(t) \quad (1.1)$$

donde $f_s(t)$ está representada por el producto de $f(t)$ por $s(t)$ que es una serie de pulsos periódicos de amplitud unitaria, ancho τ y periodo $T = 1/f_c$ llamada función de conmutación o función gating, figura 1.2.

Extendiendo $s(t)$ en series de Fourier, la señal muestreada $f_s(t)$ puede ser escrita de la siguiente forma:

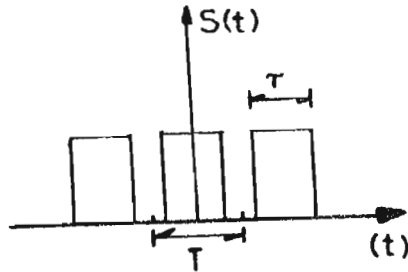


Fig. 1.2.- Función de Conmutación.

$$f_s(t) = df(t) \left[1 + 2 \sum \frac{\text{sen } \pi n \tau / T}{\pi n \tau / T} \cos 2\pi n f_c t \right] \quad (1.2)$$

Por el teorema de transferencia, la transformada de Fourier \$F_c(W)\$ representa \$F(W)\$, la transformada de \$f(t)\$ cambia positivamente y negativamente \$nW_c\$ con \$W_c = 2\pi f_c\$, entonces:

$$F_c(W) = \frac{1}{2} F(W - nW_c) + \frac{1}{2} F(W + nW_c) \quad (1.3)$$

Si \$F(W)\$ está centrada en cero, \$F_c(W)\$ estará centrada a \$\pm nW_c\$, del teorema de transferencia de frecuencias el producto \$f(t)\$ por \$\cos nW_c t\$ (donde \$f(t)\$ es la señal modulada y \$\cos nW_c t\$ la señal portadora) en la Ec. (1.2) da como resultado una nueva señal centrada alrededor de la portadora.

Si repetimos el proceso para cada \$n\$, encontrando los valores para los factores apropiados en la Ec. (1.2) y superponiendo todos los términos, la transformada de Fourier de \$f_s(t)\$ será:

$$F_s(W) = dF(W) + d \sum_{\substack{n=-\infty \\ n \neq 0}}^{\infty} \frac{\text{sen } \pi n d}{\pi n d} F(W - nW_c) \quad (1.4)$$

donde $d = \tau / T$ es el ciclo obligatorio.

La suma de las transformadas de Fourier individuales estan cada una centradas a un múltiplo de la frecuencia de muestreo, figura 1.3. Donde se puede ver que la amplitud de cada componente sucesiva decrece en:

$$(\text{sen } n \pi d) / n \pi d$$

Como se muestra en la figura 1.3, el efecto de muestrear $f(t)$ ha sido para transferir el espectro sobre todas las armónicas de la frecuencia de muestreo. Alternativamente, el efecto de multiplicar la función de muestreo periódico $s(t)$ por la no periódica $f(t)$ ha sido para abrir el espectro lineal discreto dentro de un espectro continuo simétricamente situado alrededor de las frecuencias lineales originales.

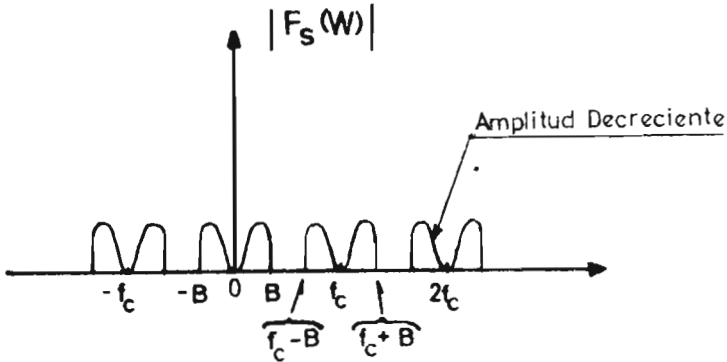


Fig. 1.3.- Amplitud espectral de la señal de entrada muestreada.

Una forma alternativa para llegar a la Ec. (1.4) desde

la Ec. (1.1) es por medio del teorema de convolución.

$$F_s(W) = \frac{1}{2\pi} F(W) * S(W) \quad (1.5)$$

De la figura 1.3, la frecuencia de limitación a la cual $F(W)$ y $F(W - W_c)$ se unirán, está dada por

$$f_c - B = B \quad f_c = 2B \quad (1.6)$$

el cual es el rango de muestreo de Nyquist.

La transmisión vía telefónica de la voz es filtrada generalmente a $B = 3.3$ KHz. El rango de Nyquist será así $2B = 6.6$ KHz pero generalmente es muestreada a un rango de 8 KHz para asegurar la separación del espectro en frecuencia y para simplificar el problema de filtración para recuperar la voz (función $f(t)$).

1.1.2.- TEOREMA DE MUESTREO EN FORMA GENERAL

El proceso de muestreo es particularmente importante en sistemas de tiempo discreto y en los de comunicación digital, en los cuales el proceso es inherente a la operación del sistema. Existen diversos tipos de operaciones de muestreo de importancia práctica, que son : (11)

- a) MUESTREO PERIODICO - donde los instantes de muestreo están equidistantemente espaciados (nT donde $n = 0, 1, 2, \dots$).
- b) MUESTREO DE ORDEN MULTIPLE - donde el esquema de los instantes de muestreo se repiten periódicamente.

- c) MUESTREO DE RITMO MULTIPLE - donde concuerdan simultáneamente dos operaciones de muestreo en $nT = pT_1$ y qT_2 , con T_1 y T_2 son constantes y con p y q son enteros.
- d) MUESTREO AL AZAR O ALEATORIO - donde los instantes de muestreo son causales, es decir, nT es aleatoria.

El teorema de muestreo de Shannon (desarrollado a partir del de Nyquist) es particularmente importante en la teoría moderna de comunicaciones. Además tiene una profunda significación en el concepto moderno de la teoría de la información, donde algunas medidas del contenido de la información de una señal específica debe ser relacionada al número de cantidades independientes necesarias para describir la señal completa. Si el número está escrito en unidades binarias el contenido de la información es medido en bits. El teorema de muestreo uniforme o periódico se puede enunciar de la siguiente manera:

"Si una señal limitada en banda que no contiene componentes espectrales mayores que la frecuencia de muestreo $|f| \leq B$ se describe en forma completa por medio de valores muestra instantáneos τ espaciados de manera uniforme en el tiempo, con periodo $T \leq 1/2B$. Si una señal ha sido muestreada a la velocidad de Nyquist o a una mayor $f_c \geq 2B$ y los valores muestra se les representa como una serie de pulsos o como impulsos ponderados, la señal puede ser reconstruida en forma exacta a partir de sus muestras por medio de un filtro pasobajas ideal de ancho de banda B ".

Este enunciado contesta a las preguntas; ¿Cuál es el rango al que se debe muestrear? ¿Hay algún límite para el rango de muestreo? ¿Se pierde información valiosa en el proceso de muestreo?.

Del teorema es evidente que el rango de muestreo mínimo de $2B$ muestras independientes por segundo es muy significativo, pues es suficiente para representar la señal $f(t)$, así como que es la apropiada para no perder información y para ser capaz de reconstruir eventualmente la señal de entrada continua. Por tanto el teorema de muestreo periódico surge a partir de la Ec. (1.6) y se puede generalizar más de la forma siguiente:

"Cualquier $2B$ muestras independientes por segundo caracterizan completamente una señal de banda limitada. Alternativamente cualquier fragmento único (independient) $2BT'$ de información es necesario para especificar completamente una señal sobre un intervalo T' segundos de longitud".(10)

Una función obedeciendo las anteriores condiciones puede ser expresada y completamente especificada en cualquier punto en términos del valor de una función y sus sucesivas derivadas a otro punto, por lo que si posee un número finito de derivadas no cero, tan sólo un número finito de fragmentos de información es necesario para describir dichas funciones como en las series de Taylor.

Utilizando las series de Fourier se puede probar el teorema de muestreo en su forma más generalizada, donde cualquier fragmento $2BT'$ de información es necesario para caracterizar una señal sobre un intervalo T' segundos. Por lo que si tenemos $f(t)$ con banda limitada y nos interesamos so

bre un intervalo T' segundos de longitud, entonces expandemos $f(t)$ en series de Fourier con T' como periodo base y con banda limitada a B Hz, obtenemos solamente un número finito de términos

$$f(t) = \frac{C_0}{T'} + \frac{2}{T'} \sum_{n=1}^{BT'} C_n \cos(W_n t + \theta_n) \quad (1.7)$$

$$f(t) = \frac{C_0}{T'} + \frac{2}{T'} \sum_{n=1}^{BT'} (a_n \cos W_n t + b_n \sin W_n t) \quad (1.8)$$

$$W_n = \frac{2\pi n}{T'} ; |C_n| = \sqrt{a_n^2 + b_n^2} ; \theta_n = -\tan^{-1} \left(\frac{b_n}{a_n} \right)$$

C_0 es el término DC y sirve para trasladar de nivel a $f(t)$ y no proporciona información nueva. Si B es el componente de frecuencia máxima de $f(t)$, entonces W_n tiene un valor máximo de $2\pi n/T' = 2\pi B$, por tanto, el valor máximo de n será BT' . Por consiguiente hay $2BT'$ coeficientes independientes de Fourier (en la Ec. (1.7) la suma de los C_n 's y θ_n 's y en la Ec (1.8) la suma de a_n 's y b_n 's) y cualquier $2BT'$ muestras independientes de $f(t)$ son necesarias para especificarla sobre el intervalo de T' segundos.

Además aquí es donde se encuentra el punto de conexión entre el teorema de muestreo y la capacidad de un canal, donde su ecuación es:

$$C = 1/T' \log_2 n \quad \text{si} \quad T' = 1/B$$

$$\text{entonces} \quad C = B \log_2 n \text{ bits/seg} \quad (1.9)$$

donde B es el ancho de banda del canal en Hz. En T' segundos, el canal permitirá una transmisión de $BT' \log_2 n$ bits.

Recordando la ley de Hartly que enuncia: "La información que puede ser transmitida sobre un sistema de banda li

mitada es proporcional al producto de tiempo del ancho de banda al tiempo de transmisión".

Para la ley de Hartly la información llevada por la señal debe ser proporcional a BT' y para el teorema de muestreo si $\tau = 1/2B$, entonces el canal transmitirá $2BT' \log_2 n$ por lo que en esencia la ley de Hartly y el teorema de muestreo es lo mismo.

También se debe de tener en cuenta que existen errores e irregularidades en el muestreo de una señal a intervalos uniformes T . Casi todos los sistemas prácticos de muestreo usan muestreo uniforme, por lo menos son diseñados sobre estas bases y cualquier desviación en su mayoría es pequeña y sin intención de que exista ésta y ya que es un retraso constante en el filtro de recobro de la señal, usualmente no es considerado como una distorsión.

Aunque también existen sistemas imperfectos que toman muestras a periodos desplazados y luego lo aplican al filtro de recobro con periodos desplazados, el error que se produce en el tren de muestras es llamado desplazamiento. Los desplazamientos pueden ser sistemáticos o aleatorios. El más sencillo desplazamiento sistemático de significancia práctica es el de repetición periódica en cada muestra alterada por una cantidad constante. El desplazamiento aleatorio de un periodo de muestreo es llamado efecto de temblor (jitter), el cual es como un ruido espectral en lugar de una simple frecuencia falsa. (17)

Otra distorsión o error en el muestreo que debe ser tomado en cuenta es el fenómeno de traslapamiento (aliasing), que se produce cuando no se ha tomado muy en cuenta la frecuencia de muestreo, por lo que para evitarla, la frecuen-

cia de muestreo f_c debe ser por lo menos dos veces la frecuencia de Nyquist (Ec. (1.6)), ver figura 1.4.

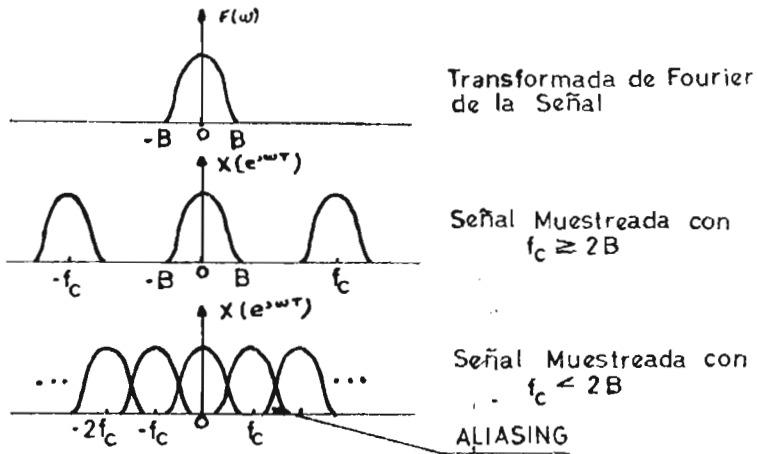


Fig. 1.4.- Muestra grfica del traslapamiento (aliasing)

As, para asegurar que las condiciones del teorema de muestreo sean satisfechas, la seal para ser muestreada es generalmente filtrada primero por un filtro pasobajas cuya frecuencia de corte es menor o igual que la mitad de la frecuencia de muestreo. Dicho filtro es frecuentemente llamado "prefiltro anti-aliasing" porque su propsito es garantizar la no ocurrencia del traslapamiento debido al muestro.

1.1.3.- DECIMACION E INTERPOLACION DE FORMAS DE ONDA MUESTREADAS

En ocasiones existe la necesidad de cambiar el rango de muestreo de una seal de tiempo discreto, el proceso de reduccin y el proceso de incremento del rango de muestreo es llamado DECIMACION e INTERPOLACION respectivamente.

Para ambos casos, tendremos en cuenta que una secuencia de muestras $x(n) = f_s(t) = f(nT)$, donde de la representación de la Ec. (1.1), $f(t)$ es una función de banda limitada con transformada de Fourier $F(\omega) = F(j\Omega) = 0$ para $|\Omega| > 2\pi B$. Entonces si $1/T = f_c = 2B$, la transformada de Fourier de $x(n)$ satisficará su proporcionalidad con la transformada de Fourier de $f(t)$ en banda base, la que se representa de la siguiente forma:

$$X(e^{jn\pi}) = \frac{1}{T} F(j\Omega) \quad |\Omega| < \frac{\pi}{T} \quad (1.10)$$

a) DECIMACION

Para describir el proceso de reducción del rango de muestreo podemos basarnos sobre un ejemplo, en el cual la voz es muestreada a un alto rango como en la Modulación Delta y se desea convertirla a un rango de muestreo menor, es decir, una representación multi-bit PCM.

Si deseamos reducir el rango de muestreo en un factor M , es decir, deseamos computar a una nueva secuencia que corresponda a las muestras de $f(t)$ tomadas a un periodo $T' = MT$, entonces

$$y(n) = f_s(t_1) = f(nTM) \quad (1.11)$$

donde $t_1 = nT'$. Y también es fácil ver que si $x(n) = f_s(t) = f(nT)$, entonces:

$$y(n) = x(Mn) \quad -\infty < n < \infty \quad (1.12)$$

De la ecuación anterior, se ve que $y(n)$ se obtiene re-

teniendo simplemente una de cada M muestras periódicamente. Del teorema de muestreo si $1/T' > 2B$, entonces las muestras $y(n)$ también serán las adecuadas para representar $f(t)$. La transformada de Fourier de $x(n)$ está relacionada con la de $y(n)$ de la siguiente manera:

$$Y(e^{j\Omega T'}) = \frac{1}{M} X(e^{j\Omega T'/M}) = \frac{1}{M} \sum X(e^{j\Omega T}) \quad (1.13a)$$

por la propiedad de escalamiento de la transformada de Fourier:

$$Y(e^{j\Omega T'}) = \frac{1}{M} \frac{1}{T} F(j\Omega) \quad (1.13b)$$

$$= \frac{1}{T} F(j\Omega) \quad -\frac{\pi}{T'} < \Omega < \frac{\pi}{T'} \quad (1.13c)$$

Si $T' = 3T$ ocurrirá el traslapamiento puesto que $1/T < 2B$. Pero esto se puede evitar si $x(n)$ es filtrada con un filtro pasobajas ideal con frecuencia de corte $\pi/T' = \pi/3T$, esta filtración produce una secuencia $w(n)$. El aliasing no ocurre cuando el rango de muestreo de la señal filtrada se reduce por el factor 3. Todo lo anterior se muestra gráficamente en la figura 1.5.

b) INTERPOLACION

Al igual que en el proceso de decimación, para describir el de interpolación lo ilustraremos con un ejemplo. Si la señal de voz es muestreada a un rango bajo para una codificación eficiente, pero para reconstruirla será necesario un rango de muestreo más alto.

Si tenemos una secuencia de muestras $x(n) = f_s(t) = f(nT)$ y deseamos incrementar el rango de muestreo por un

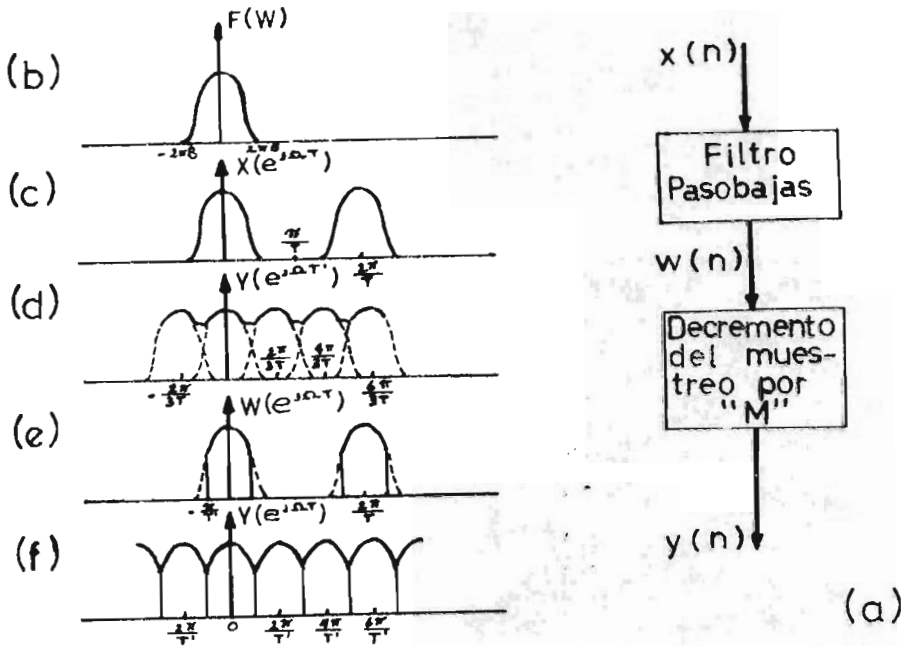


Fig. 1.5.- a) Diagrama a bloques de la representación de la decimación. b) Transformada de Fourier de $f(t)$ $F(W)$. c) Transformada de Fourier de $f_s(t) = x(n)$, $X(e^{j\Omega T})$. d) La muestra de la reducción del rango de muestreo por un factor 3. e) Transformada de Fourier de la salida del filtro $W(e^{j\Omega T})$. f) El proceso de decimación ya realizado.

factor L , es decir, debemos computar una nueva secuencia de muestras para $f(t)$ que será con periodo $T' = T/L$ entonces tendremos

$$y(n) = f_s(t_n) = f(nT/L) \quad (1.14)$$

con $t_1 = nT'$, donde podemos ver que $y(n) = x(n/L)$ para $n = 0, \pm L, \pm 2L, \dots$, pero hay que llenar las muestras no conocidas para los otros valores de n por medio del proceso de interpolación, que se puede realizar usando un filtro digital. Considerando una secuencia

$$\begin{aligned} v(n) &= x(n/L) & n &= 0, \pm L, \pm 2L, \dots \\ &= 0 & & \text{de otra forma} \end{aligned} \quad (1.15)$$

La transformada de Fourier de $v(n)$ será

$$\begin{aligned} V(e^{j\Omega T'}) &= X(e^{j\Omega T' L}) \\ &= X(e^{j\Omega T}) \end{aligned} \quad (1.16)$$

donde $V(e^{j\Omega T'})$ es periódica con periodo $2\pi/T = 2\pi/LT'$. El proceso de interpolación puede visualizarse en la figura 1.6, en donde en la figura 1.6a se representa el proceso por medio de un diagrama a bloques, en la figura 1.6b se representa la Ec. (1.16) para el caso de $T' = T/3$. Por lo que para obtener la secuencia de la Ec. (1.14) de la secuencia $v(n)$ tenemos que:

$$V(e^{j\Omega T'}) = \frac{1}{T'} F(j\Omega) \quad -\frac{\pi}{T'} \leq \Omega \leq \frac{\pi}{T'} \quad (1.17)$$

Y asumiendo lo de la Ec. (1.10), además de la figura 1.6b, que se requiere que las imágenes de $F(j\Omega)$ en $V(e^{j\Omega T'})$ que están centradas en $\Omega = 2\pi/T$ y $\Omega = 4\pi/T$, sean filtradas por el filtro digital pasobajas que rechace todas las componentes en frecuencia en el rango $\pi/T \leq \Omega \leq \pi/T'$. Por lo que para asegurar que la amplitud es la correcta para el intervalo de muestreo T' , la ganancia del

filtro deberá ser $L = T/T'$, por tanto de la figura 1.6a, tenemos:

$$Y(e^{j\Omega T'}) = H(e^{j\Omega T'}) V(e^{j\Omega T'}) \quad (1.18a)$$

donde $H(e^{j\Omega T'})$ es la función de transferencia del filtro

$$Y(e^{j\Omega T'}) = H(e^{j\Omega T'}) X(e^{j\Omega T}) \quad (1.13b)$$

$$= H(e^{j\Omega T'}) \frac{1}{T} F(j\Omega) \quad (1.18c)$$

por lo que para asegurar resulte la Ec. (1.17), entonces:

$$\begin{aligned} H(e^{j\Omega T'}) &= L & |\Omega| &\leq \pi/T \\ &= 0 & \text{de otra forma} & \end{aligned} \quad (1.19)$$

En la figura 1.6c, se muestra el proceso de interpolación realizado.

c) CAMBIO DEL RANGO DE MUESTREO NO ENTERO

Con el proceso de la combinación de la interpolación y el de decimación podemos tener un periodo de muestreo de $T' = MT/L$, con un factor L en la interpolación seguido por la decimación con un factor M , es posible aproximar arbitrariamente cualquier proporción deseada de rango de muestreo con una elección apropiada de L y M .

En la figura 1.7 vemos la combinación de procesos, en donde es posible apreciar que con un filtro pasobajas es suficiente para ambos.

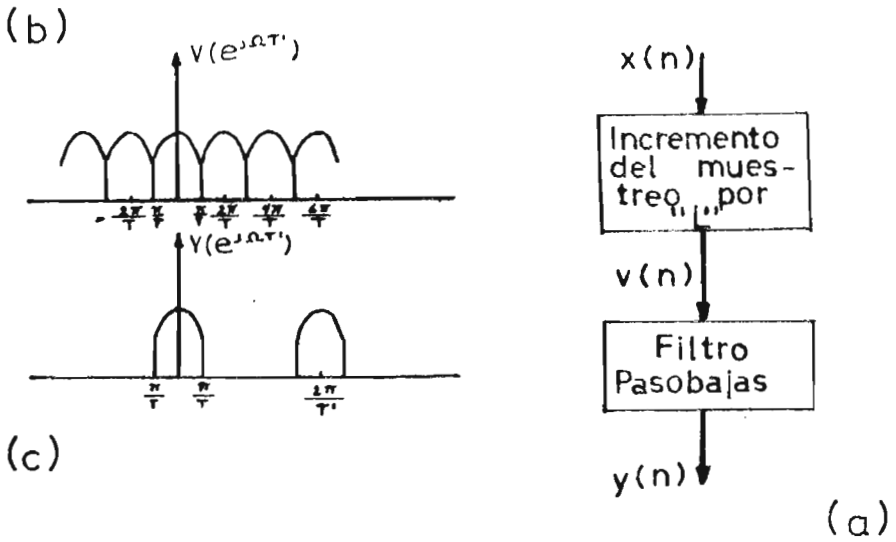


Fig. 1.6.- a) Diagrama a bloques del proceso de interpolación. b) Transformada de Fourier de la secuencia $v(n)$ Ec. (1.16). c) Proceso de interpolación ya realizado, Ec. (1.17).



Fig. 1.7.- Representación en diagrama a bloques del cambio de muestreo por un factor L/M .

1.1.4.- RECONSTRUCCION O DEMODULACION DE SEÑALES MUESTREADAS

Para reconstruir la señal muestreada es simplemente necesario generar desde cada muestra un impulso proporcional y pasar estas series de impulsos espaciados regularmente a través de un filtro pasobajas de frecuencia de corte a B Hz. La salida del filtro será entonces idéntico a la señal de entrada (excepto por un tiempo de retraso de cresta a cresta). Idealmente entonces realizaremos una reproducción perfecta de la señal si transmitimos la información con las amplitudes instantáneas exactas de la señal a intervalos espaciados a $1/2B$.

Sí muestreamos exactamente al rango de Nyquist, $f_c = 2B$, el filtro requerido debe ser de características de conte infinito, el cual sería un filtro ideal, en la práctica imposible. Un filtro pasobajas práctico con características de corte repentino tiene un resultado complejo en la filtración y alguna distorsión residual (parte de la banda lateral más baja que se encuentra cerca de f_c es transmitida). Esta situación puede ser disminuida realizando un muestreo a un rango más alto, es decir, tendremos disponible una banda de guarda, y los requerimientos del filtro son menos severos como vemos en la figura 1.8

Sí tenemos una señal $f(t)$ con banda limitada a B Hz que ha sido muestreada a intervalos de $1/2B$ segundos, la salida muestreada es de la forma mostrada en la Ec. (1.2), que nos servirá para la demodulación de muestras periódicas del filtro pasobajas.

Del teorema de muestreo sabemos que de la secuencia de muestras es posible reconstruir $f(t)$.

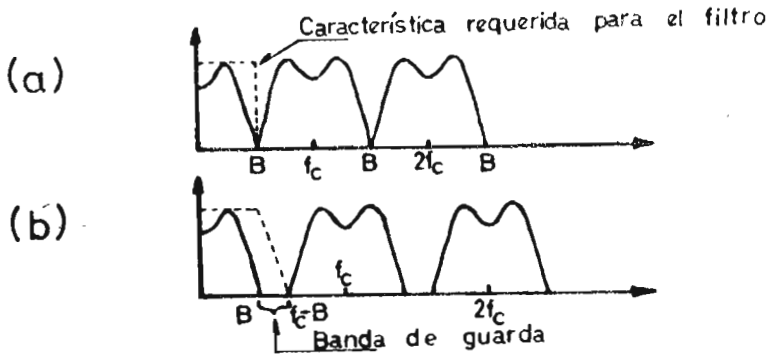


Fig. 1.8.- a) Demodulación de la señal muestreada utilizando un filtro pasobajas con $f_c = 2B$.
 b) Sí $f_c > 2B$ el filtro corta entre B y $f_c - B$ la atenuación a $f_c - B$ es una cantidad prescrita con respecto a la banda de paso.

Para demostrar que las muestras dadas son suficientes para reproducir $f(t)$, teniendo la transformada de Fourier de $f(t)$:

$$F(W) = \int_{-\infty}^{\infty} f(t) e^{-jWt} dt \quad |W| > 2\pi B \quad (1.20)$$

y debido a que tenemos una $f(t)$ con banda limitada $F(W)$ la podemos hacer periódica con un periodo $4\pi B$, entonces la expandemos en series de Fourier de periodo $4\pi B$ para usar la dentro del intervalo $|W| \leq 2\pi B$, por lo que ahora tenemos:

$$\begin{aligned} F(W) &= \frac{1}{4\pi B} \sum_{n=-\infty}^{\infty} C_n e^{j(2\pi n/4\pi n)} \\ &= \frac{1}{4\pi B} \sum_{n=-\infty}^{\infty} C_n e^{j\pi n/2B} \\ &= 0 \quad |W| > 2\pi B \end{aligned} \quad (1.21)$$

definiendo a C_n por

$$C_n = \int_{-2\pi B}^{2\pi B} F(W) e^{-jW/2B} dW \quad (1.22)$$

y definiendo $f(t)$ de $F(W)$

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} F(W) e^{jWt} dW \\ &= \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} F(W) e^{jWt} dW \end{aligned} \quad (1.23)$$

Si $t = -n/2B$, $f(t)$ para éste caso especial de t nos resultará:

$$\begin{aligned} f(-n/2B) &= \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} F(W) e^{-jWn/2B} dW \\ &= \frac{C_n}{2\pi} \end{aligned} \quad (1.24)$$

De la Ec. (1.23), vemos que podemos encontrar $f(t)$ para todos los tiempos posibles, al igual que podremos encontrar el coeficiente C_n para diversos intervalos de muestreo de $f(t)$ y al conocer dicho coeficiente podremos encontrar $F(W)$ de la Ec. (1.21), pero sí ésta expansión de series de Fourier para $F(W)$ la sustituimos dentro de la Ec. (1.23), la representación de la integral de Fourier de $f(t)$ será:

$$\begin{aligned} f(t) &= \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} F(W) e^{jWt} dW \\ &= \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} \frac{1}{4\pi B} \left[\sum_{n=-\infty}^{\infty} C_n e^{jWn/2B} \right] e^{jWt} dW \end{aligned} \quad (1.25)$$

Para que la integral sea finita y no surgan dificultades de integrales impropias intercambiaremos el orden de la

integración y la sumatoria, además C_n se aproximará a cero para cuando n sea lo bastante grande, por lo que la Ec. 1.25) nos quedará:

$$\begin{aligned} f(t) &= \sum_n \frac{C_n}{2\pi} \frac{1}{4\pi B} \int_{-2\pi B}^{2\pi B} e^{j\omega(t+n/2B)} d\omega \\ &= \sum_n \frac{C_n}{2\pi} \frac{\text{sen } 2\pi B(t+n/2B)}{2\pi B(t+n/2B)} \end{aligned} \quad (1.26)$$

pero de la Ec. (1.24) para $f(-n/2B) = C_n/2$, entonces:

$$\begin{aligned} f(t) &= \sum_{n=-\infty}^{\infty} f(-n/2B) \frac{\text{sen } 2\pi B(t+n/2B)}{2\pi B(t+n/2B)} \\ &= \sum_{n=-\infty}^{\infty} f(n/2B) \frac{\text{sen } 2\pi B(t-n/2B)}{2\pi B(t-n/2B)} \end{aligned} \quad (1.27)$$

De esta última ecuación, indicamos matemáticamente que tomamos cada muestra, la multiplicamos por el factor de peso $(\text{sen } x)/x$, entrado en el tiempo de muestras de ocurrencia y sumando los términos resultantes. Esto es lo que se realiza cuando se pasa las muestras a través de un filtro pasobajas ideal cortando a B Hz.

Si asumimos que $f(t)$ está limitada en banda a B Hz, y muestreada en una longitud de tiempo de τ que es el tiempo de muestreo periódico a intervalos de $1/2B$ segundos, podemos demostrar lo dicho anteriormente.

$f(t)$ puede asumirse como una constante cercana al tiempo de muestreo si $\tau \ll 1/2B$, como se muestra en la figura 1.9, entonces la transformada de Fourier para la muestra $f(n/2B)$ está dada por:

$$\begin{aligned} F_n(\omega) &= \int_{-\infty}^{\infty} f(t) e^{-j\omega t} dt \\ &= \tau f(n/2B) e^{-j\omega n/2B} \end{aligned} \quad (1.28)$$

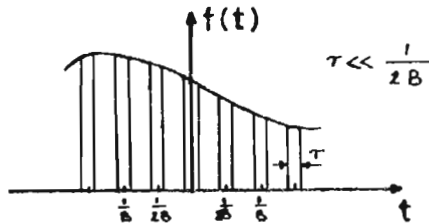


Fig. 1.9.- Función continua $f(t)$ muestreada.

pues $f(t)$ se asumió constante en τ segundos y cero en otra parte. Y si tenemos en cuenta el teorema de transferencia y que la transformada de Fourier de una función impulso es:

$$F(W) = \int_{-\infty}^{\infty} \delta(t) e^{-jWt} dt = 1 \quad (1.29)$$

veremos que la Ec. (1.28) es la transformada de Fourier de la función impulso de amplitud $\tau f(n/2B)$ y localizada en el tiempo a $t = n/2B$.

El factor amplitud $\tau f(n/2B)$ representa el área bajo la curva del muestreo individual y puesto que τ es muy pequeño, tendremos que aproximar la muestra en un impulso de la misma área. Ahora, este impulso se pasa a través de un filtro pasobajas ideal de ancho de banda B Hz, y asumiendo características como cambio de fase cero y amplitud unitaria (asumidas para simplificación), la respuesta del filtro a la salida de la muestra $f(n/2B)$ aplicada a la entrada será:

$$\begin{aligned} g_n(t) &= \frac{1}{2\pi} \int_{-2\pi B}^{2\pi B} F(W) e^{jWt} dW \\ &= f(n/2B) \frac{\tau}{2\pi} \int_{-2\pi B}^{2\pi B} e^{jW(t-n/2B)} dW \end{aligned} \quad (1.30)$$

Integrando, tenemos:

$$g_n(t) = 2B\tau f(n/2B) \frac{\text{sen } 2\pi B(t-n/2B)}{2\pi B(t-n/2B)} \quad (1.31)$$

lo cual se muestra en la figura 1.10, teniendo un valor máximo al punto de muestreo dado $t = n/2B$, y cero a todos los otros puntos y el pico de cada término $1/2B$ ocurre a un punto de muestreo donde todas las otras salidas son cero, y la salida a cada uno de los puntos muestreados es exactamente proporcional a la magnitud de la entrada muestreada en dichos puntos.

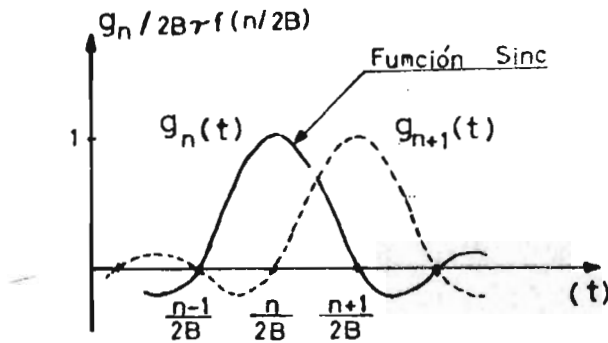


Fig. 1.10.- Respuesta del filtro a la señal muestreada.

Ya que tenemos un filtro lineal ideal, la salida completa es la superposición de las salidas muestreadas individuales:

$$\begin{aligned} g(t) &= \sum_{n=-\infty}^{\infty} g_n(t) \\ &= 2B\tau \sum_{n=-\infty}^{\infty} f(n/2B) \frac{\text{sen } 2\pi B(t-n/2B)}{2\pi B(t-n/2B)} \\ &= 2B\tau f(t) \end{aligned} \quad (1.32)$$

La salida del filtro pasobajas $g(t)$, es así, idénticamente proporcional a la señal original $f(t)$ en todos los instantes de tiempo y no solamente a los puntos muestreados.

La entrada original $f(t)$ puede ser entonces reproducida de las muestras pasadas a través de un filtro pasobajas ideal de ancho de banda de B Hz. El factor proporcional $2B\tau$ representa la proporción de los anchos de banda del filtro y pulsos muestreados. Ya que $\tau \ll 1/2B$ y que el ancho de banda efectivo del pulso es mucho más grande que el del filtro. La amplitud de la salida es reducida por la proporción de anchos de banda, es decir, por $2B\tau$

De la Ec. (1.2), mostrada aquí de nuevo;

$$f_s(t) = df(t) \left[1 + 2 \sum_{n=1}^{\infty} \frac{\text{sen } \pi n d}{\pi n d} \cos \frac{2 \pi n t}{T} \right]$$

un filtro ideal-pasobajas de amplitud unitaria con ancho de banda B , produce $g(t) = df(t)$ a la salida.

El filtro ideal es realmente un artificio con utilidad matemática y que naturalmente sólo puede buscarse en la práctica una aproximación y la más cercana a lo ideal sólo se puede alcanzar con un filtro digital.

1.2.- CUANTIZACION

En la tecnología de la comunicación moderna es necesario que las señales analógicas además de ser muestreadas sean digitalizadas antes de la transmisión.

Naturalmente es imposible transmitir la amplitud exacta de una muestra. Por lo que la amplitud de una muestra

es transmitida como la amplitud de un pulso o como la posición en tiempo de un pulso, sin embargo, es posible permitir ciertos niveles discretos en éstos. Entonces, cuando la señal es muestreada, el nivel más cercano al nivel de la señal verdadera es transmitida. Cuando ésto es recibido y amplificado, tendrá un nivel un poco diferente de los niveles especificados, debido al ruido y a la distorsión que causa errores en la información recobrada. Entonces la señal reconstruida se dice tendrá casi el nivel originalmente enviado.

Representar la señal permitiendo ciertos niveles discretos solamente, o sea realizar el proceso de digitalización de la señal analógica original es llamado el proceso de CUANTIZACION.

A diferencia del proceso de muestreo, en el proceso de cuantización resulta una pérdida irrecuperable de información, ya que es imposible reconstruir la señal analógica original desde la versión cuantizada. Sin embargo, desde el concepto de información, no hay necesidad de transmitir todas las amplitudes posibles de la señal. Debido a que el circuito detector o demodulador no es capaz de distinguir variaciones finas o pequeñas en amplitud de la señal, por lo que el último en recibir la información ya sea como sonido o como imágen está limitado con respecto a la fina graduación de la señal que puede distinguir, es decir, nuestros oídos o nuestros ojos.

Debido a ésta última limitación se hace posible la cuantización. En un sistema, los pulsos muestreados pueden ser después cuantizados o pueden ser simultáneamente realizados ambos procesos.

El proceso de cuantización introduce algún error en la

reproducción eventual de la señal, la señal demodulada diferirá algo de la señal original. El efecto total es como si ruido adicional haya sido introducido dentro del sistema. Esta discrepancia entre una forma de onda y su aproximación cuantizada es llamada comunmente RUIDO CUANTIZADO. Esta no es del todo la mejor elección de como llamarlo, ya que el daño está cercanamente relacionado con la señal y no tiene todas las características aleatorias normalmente asociadas con el ruido y algunas veces referida como error (En el caso de la transmisión de sonido ésto se manifiesta como un cacareo de fondo. En el caso de la transmisión de imágen, la graduación continua de grises desde el negro al blanco es remplazado por un número discreto de grises y la imágen también parece algo ruidosa).

Sí la cuantización de las muestras de la señal fueran transmitidas directamente como pulsos de alta variación, el sistema resultante será simplemente un PAM (Modulación por Amplitud de Pulsos) cuantizado. Pero con niveles de voltajes discretos o ennumerados, cada nivel puede ser codificado en alguna forma arbitraria antes de la transmisión.

El pulso de la señal muestreada y cuantizada es comunmente codificada dentro de un grupo equivalente o paquetes de pulsos binario a amplitud fija. Lo que finalmente nos proporciona la señal binaria. La codificación de los niveles de amplitud dentro de la forma binaria puede ser realizada por diversos caminos.

El número de dígitos usados depende sobre el número de niveles de cuantización utilizados, por ejemplo, 128 niveles requieren 7 dígitos binarios o bits para representación de cada nivel. Ya que estos dígitos binarios pueden ser transmitidos en el intervalo de muestreo originalmente asignado

para una muestra cuantizada, el ancho del pulso será más estrecho y el ancho de banda de transmisión subirá proporcionalmente al número de pulsos binarios necesarios para la representación de la señal.

El código binario es un caso especial de la codificación teóricamente posible en un sistema PCM. En general cualquier muestra de señal cuantizada puede ser codificada dentro de un grupo de m PULSOS, cada uno con n NIVELES DE AMPLITUD posibles. Debido a que la información llevada por esos m pulsos es equivalente a la información llevada por los NIVELES DE AMPLITUD ORIGINAL M, el número de posibles combinaciones de amplitud de los m pulsos deben ser igual a M, entonces:

$$M = n^m \quad (1,33)$$

Del ejemplo mencionado antes, $m = 7$; $n = 2$; $M = 128$.

La forma de transmisión binaria suministra por otro lado, una mayor inmunidad al ruido y ésto es por tanto el esquema de comunicación digital.

1.2.1.- RUIDO CUANTIZADO Y CUANTIZACION UNIFORME

Como se dijo ante, una señal analógica para ser codificada dentro de una forma digital, primero debe ser cuantizada dentro de grados de amplitud discreta. Una vez cuantizada, los valores instantáneos de la señal nunca pueden ser realmacenados exactamente. Esto por tanto, da surgimiento a errores de variación aleatoria llamado ruido cuantizado. Este ruido puede ser reducido a cualquier grado deseado eligiendo el grado de quantum o nivel de separación lo bastan-

te fina o pequeña,

Sin embargo, como se aprecia en la Fc. (1.33), si se incrementa el número de dígitos binarios será mayor el número de niveles y de aquí que el ancho de banda será necesariamente más ancha. Por lo que normalmente se permite una cuantización de pocos niveles tan consistente como sea posible con el objetivo de transmisión. Para transmisión de voz vía PCM, con 128 niveles o 7 dbits son suficientes para asegurar una transmisión de alta calidad.. Pero el número puede diferir para otras aplicaciones y a la calidad deseada,

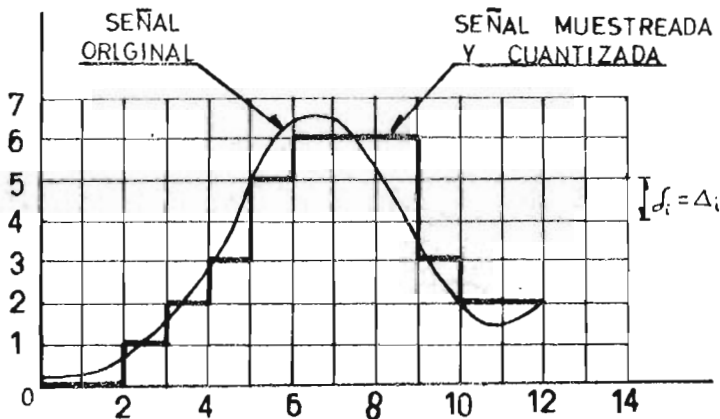


Fig. 1.11.- Señal muestreada y cuantizada, donde el cuantum Δ_i es de tamaño uniforme.

La cuantización uniforme es realizada cuando el cuantum es de tamaño uniforme como se aprecia en la figura 1.11, y es un caso especialmente importante.

Podemos calcular el voltaje de error medio cuadrático

debido a la cuantización. Para lo cual dividimos el rango de la señal dentro de N intervalos y dejamos los i -ésimos intervalos sean de ancho δ_i centrada sobre el valor X_i . Donde la amplitud de señal está situada en el rango:

$$X_i - \frac{1}{2} \delta_i \leq x \leq X_i + \frac{1}{2} \delta_i \quad (1.34)$$

donde X_i representa la amplitud cuantizada, siendo esto una suposición importante que simplificará grandemente el análisis. Así, el error cuadrático instantáneo es $(X - X_i)^2$. La distribución de probabilidad de la amplitud de la señal es $P(x) dx$. Si el quantum δ_i es pequeño en relación a la excursión de la señal, entonces $P(x)$ variará poco y puede ser aproximada a $P(X_i)$. Entonces el error cuadrático debido a las señales cayendo dentro del quantum es:

$$\begin{aligned} E_i^2 &= \int_{X_i - \frac{1}{2} \delta_i}^{X_i + \frac{1}{2} \delta_i} (X - X_i)^2 P(X) dx \\ &\approx P(X_i) \int_{-\frac{1}{2} \delta_i}^{\frac{1}{2} \delta_i} z^2 dz \\ &= \frac{1}{12} P(X_i) \delta_i^2 \end{aligned} \quad (1.35)$$

y sí:

$$P_i = \int_{X_i - \frac{1}{2} \delta_i}^{X_i + \frac{1}{2} \delta_i} P(X) dx \approx P(X_i) \delta_i \quad (1.36)$$

Sustituyendo esto en la Ec. (1.35) nos queda:

$$E_i^2 = \frac{1}{12} \delta_i^2 P_i \quad (1.37)$$

por lo que el error cuadrático total será:

$$E^2 = \sum E_i^2 = \frac{1}{12} \sum \delta_i^2 P_i \quad (1.38)$$

Por definición, la $\sum_i P_i$ es unitaria, en el caso de cuantización uniforme con cada quantum δ_i del mismo tamaño, entonces el error total será:

$$E^2 = \frac{1}{12} \int^2 \quad (1.39)$$

sobre la suposición que la distribución de la amplitud de la señal cae dentro del rango cubierto por la escala de cuantización. Pero si va más allá del rango con una probabilidad significativa, el error adicional deberá ser calculado separadamente.

La potencia total del ruido cuantizado que se calcule por medio de la Ec. (1.39), es una medida del ruido esperado en la banda base después de la demodulación. Usualmente, la banda base está restringida por filtros para una frecuencia algo menor que la mitad del rango de muestreo, pero en un sistema de transmisión normalmente eficiente la consecuente reducción en ruido es improbable que exceda 1 dB.

En la figura 1.12, se muestra una onda senoidal sometida a cuantización uniforme. El error instantáneo está limitado por $\pm 1/2 \delta$ y la forma de onda es, excepto cercana a los puntos de giro, la aproximación de la forma de onda triangular de longitud variada.

Ahora introduciremos la idea de relación o proporción señal a ruido (SNR). Existen dos tipos de ruido introducido por un sistema (PDS). Uno de ellos es el ruido cuantizado, y el otro es un pulso falso de ruido causado por la incorrecta interpretación de la amplitud señalada de un pulso por el receptor o por cualquier repetidor. Este ruido puede surgir en cualquier lugar a lo largo del sistema, y es acumulativo, sin embargo, este ruido decrece tan rápidamente

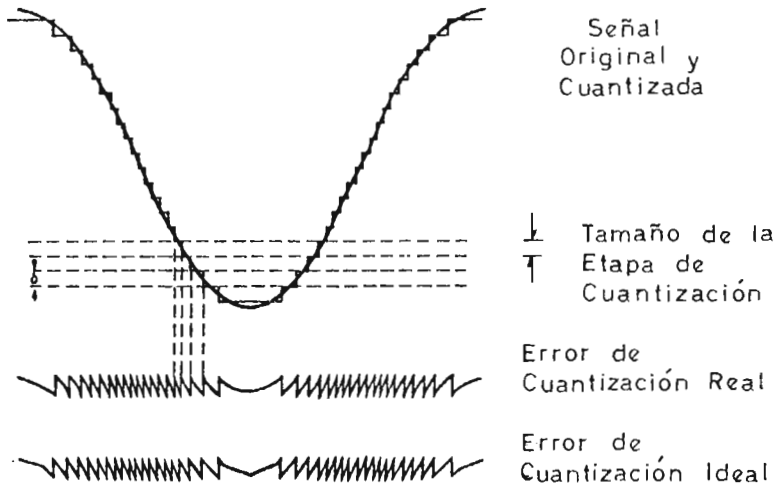


Fig. 1.12.- Cuantización de una onda senoidal

como la potencia de la señal es incrementada antes del umbral y en la práctica, es despreciable. Por tanto, la relación señal a ruido en un sistema (PDS) está fijado únicamente por el ruido cuantizado.

Si la señal es grande comparada con una simple etapa de cuantización, el error máximo introducido en las muestras sucesivas por la cuantización es la mitad de una etapa cuantizada. Todos los valores de error sobre este valor máximo es igualmente probable. El error rms introducido es por tanto $1/2\sqrt{3}$ veces la altura de una etapa de cuantización simple. Cuando la señal es reconstruida de las muestras decodificadas conteniendo el error de cuantización, la relación señal a ruido rms será por tanto:

$$SNR = 2\sqrt{3} n^m \quad (1.40)$$

donde n^m es el número de niveles. Expresandola en dB tenemos

$$\begin{aligned} 20 \log_{10}(\text{SNR}) &= 20 \log_{10} 2\sqrt{3} + m (20 \log_{10} n) \\ &= 10.8 + m(20 \log_{10} n) \end{aligned} \quad (1.41)$$

para el sistema binario $n = 2$, entonces:

$$(\text{SNR})_{\text{dB}} \cong 10.8 + 6m \quad (1.42)$$

Si m es el número de dígitos y es un factor relacionado al ancho de banda total usado en la transmisión, y si el ancho de banda de la señal que es transmitida es $W = mB$, entonces sí tenemos un incremento en los niveles N , tenemos un incremento en el número de dígitos m y por consiguiente en el ancho de banda.

Ya que el ancho de banda es proporcional a m , el número de pulsos en el grupo código, la salida SNR se incrementa exponencialmente con el ancho de banda. Pero el SNR en dB varía linealmente con el número de dígitos por grupo de código y de aquí con el ancho de banda, Ec. (1.42), (15).

1.2.2.- PROCESO DE COMPRESION Y EXPANSION (COMPANDING) DE UNA SEÑAL

Sí nos concentramos en una transmisión telefónica, podremos tener un orador débil (bajo volumen), donde el cuantizador deberá interpretar los sonidos del nivel más bajo, es decir, debe tener un grano muy fino para conservar la inteligibilidad. Granos más ordinarios pueden ser soportados por los picos de voz del mismo orador o por otro más fuerte (alto volumen).

Los intervalos son clasificados en tamaño, fino a amplitudes pequeñas (al centro del rango), y progresivamente más ordinarios con el incremento en amplitud (hacia los extremos del rango). Esto puede ser realizado indiferentemente, por cuantización no uniforme desempeñada directamente en el codificador, o como en la práctica, con el uso de niveles de cuantización cónica (tapered), donde las señales son primeramente comprimidas no linealmente en amplitud para forzar a todas las señales a estar situadas dentro de un rango especificado. Esta característica de compresión es típicamente de forma logarítmica. Los niveles de cuantización uniforme son entonces aplicados para esta señal comprimida, el efecto es para proporcionar proporcionalmente más niveles de cuantización a los niveles de señal más pequeña, como si el espaciamiento de niveles de cuantización hayan sido reducidos a niveles de señal más bajos. En el decodificador la señal es expandida a la amplitud que tenía con una compensación a través de una forma logarítmica inversa.

La combinación del proceso de compresión y expansión es denominado companding. Para la voz, los efectos del companding o niveles de cuantización cónica no uniforme es proporcionar una proporción uniforme de señal a ruido cuantizado sobre un rango dinámico mucho más ancho de señales. Algunos tipos de cuantización no uniforme pueden también ser derivados por traslación digital después de la codificación de grano fino uniforme. En la figura 1.13 se muestran tres métodos en forma de diagrama a bloques para obtener cuantización no uniforme.

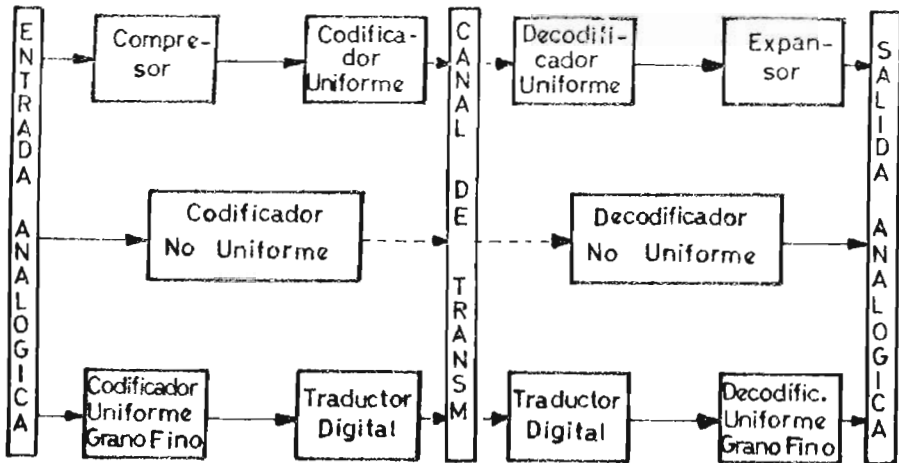


Fig. 1.13.- Métodos de obtención de cuantización no uniforme.

1.2.3.- CUANTIZACION LOGARITMICA O LEYES COMPANDING

La ley companding se realiza cuando introducimos los niveles de cuantización cónica, y como ya se explicó, en la transmisión de voz de diferentes oradores susnndo las mismas condiciones, el rango de intensidad de voz variará hasta 40 dB.

Una de las razones para la utilización del espaciamiento no uniforme de los niveles de cuantización, es debido a que con cuantización uniforme no resulta fácil cubrir efectivamente el rango dinámico producido por los diversos oradores con diferentes tonos suaves y tonos fuertes. La otra razón es que, la distribución de amplitud de señal es raramente uniforme. En voz, las pequeñas amplitudes son mucho

más probables que las grandes.

Con cuantización no uniforme en las Ecs. (1.36) a (1.38), la probabilidad P_i es muy diversa en valor y las contribuciones del ruido de los diversos intervalos difieren. Por lo que será conveniente que el ruido total de la Ec. (1.38) sea reducido, para ésto se incrementa el tamaño de los intervalos que son improbables de ser ocupados y reduciendo el tamaño de aquellos intervalos cuyos usos son altamente probables.

La no uniformidad de la distribución de amplitud tiene un efecto adicional. Una señal codificada transporta la máxima información solamente si todos los niveles son equiprobables. Así, la dispersión en P_i reduce el flujo de información y por tanto parece benéfico el uso de grandes intervalos en esas partes del rango donde la densidad de probabilidad es menor. La máxima información no coincide en general con el mínimo ruido.

En la práctica, una señal tiene una distribución de amplitud alineada y normalmente estará sujeta a una mezcla de cuantización uniforme y logarítmica. La potencia del ruido no puede ser menor que el debido a cada tipo de cuantización, pero tampoco puede exceder a la suma de los dos.

Podemos hablar de cuatro leyes companding que son:

- i) Ley A o ley logarítmica con segmentos tangenciales lineales.
- ii) Ley cuasi-logarítmica o ley μ .
- iii) Ley seno hiperbólico inverso.
- iv) Ley hiperbólica.

Para la cuantización de las señales de voz con calidad aproximadamente uniforme sobre un amplio rango de volumen se puede usar alternativamente las leyes companding. Cualquier ley companding debe recurrir a una relación lineal o cercana a lo lineal a niveles bajos.

i) La LEY A es definida por la siguiente ecuación:

$$\begin{aligned} y &= \frac{Ax}{1 + \log A} & 0 \leq x \leq 1/A \\ &= \frac{1 + \log Ax}{1 + \log A} & 1/A \leq x \leq 1 \end{aligned} \quad (1.43)$$

La pendiente inversa de la curva de compresión, la cual define los intervalos del quantum es:

$$\begin{aligned} \frac{dx}{dy} &= \frac{K_1}{A} & 0 \leq x \leq 1/A \\ &= K_1 x & 1/A \leq x \leq 1 \end{aligned} \quad (1.44)$$

donde $K_1 = 1 + \log A$.

Para una cuantización uniforme esta pendiente debe ser unitaria. Así, el cuanta sobre el segmento lineal central, el cual tiene un efecto dominante sobre señales pequeñas es disminuido por el factor para esta ley A/K_1 , que es la ventaja companding. Este factor tiene un valor típico de 25 dB.

Otra medida útil de cuantización logarítmica es el tamaño del cuanta, medida sobre una escala logarítmica (el logaritmo de la relación entre dos umbrales próximos).

Para calcular y evaluar el ruido cuantizado para cualquier señal deseada y ley companding, podemos partir de la Ec. (1.38), donde $\frac{1}{2} N d_i \approx (dx/dy)_{x=x_i}$ es la pendiente aproximada de una tangente de la curva companding y el error queda:

$$E^2 = \frac{1}{12} \sum P_i d_i^2 = \frac{1}{3 N^2} \sum P_i (dx/dy)_{x=x_i}^2 \quad (1.45)$$

para muchos niveles, la sumatoria se puede aproximar por una integral, con lo que la ecuación anterior será:

$$E^2 = \frac{1}{3 N^2} \int_{-1}^1 P(x) (dx/dy)^2 dx \quad (1.46)$$

El SNR es por tanto:

$$SNR = \frac{1}{3 N^2} \frac{\int P(x) (dx/dy)^2 dx}{\int P(x) x^2 dx} \quad (1.47)$$

el ruido debido a cuantización logarítmica ideal será:

$$SNR = K^2 / 3 N^2 \quad (1.48)$$

Para el caso de la ley A sustituimos (1.44) en la (1.46) y nos resulta un error cuadrático:

$$E^2 = \frac{K_1}{3 N^2} \left\{ \int_{-1}^1 P(x) x^2 dx + \int_{-1/A}^{1/A} P(x) (\Delta^{-2} - x^2) dx \right\} \quad (1.49)$$

donde la primera integral es la potencia de la señal en términos de la distribución de probabilidad (S^2).

La relación señal a ruido se obtiene dividiendo la Ec.(1.49) entre S^2

$$\frac{E^2}{S^2} = \frac{K_1^2}{3 N^2} \left\{ 1 + \frac{1}{S^2} \int_{-1/A}^{1/A} P(x)(A^2 - x^2) dx \right\} \quad (1.50)$$

donde el primer término es conocido como el ruido debido a una cuantización logarítmica ideal y el segundo término es debido a la ley logarítmica A.

ii) La LEY \mathcal{M} es definida por la ecuación:

$$y = \frac{\log(1 + \mathcal{M}x)}{\log(1 + \mathcal{M})} \quad (1.51)$$

lo cual no es estrictamente lineal ni estrictamente logarítmica, pero será aproximadamente lineal a niveles bajos ($x \ll \mathcal{M}^{-1}$) y aproximadamente logarítmica a niveles altos ($x \gg \mathcal{M}^{-1}$).

El tamaño de la etapa logarítmica no es constante, pero a niveles altos tiende asintóticamente a :

$$K_2 = \log(1 + \mathcal{M}) \quad (1.52)$$

Haciendo sustituciones como en el caso anterior la Ec. (1.46) en términos de la ley \mathcal{M} el error medio cuadrático será:

$$E^2 = \frac{K_2^2}{3 N^2} \int P(x)(x + \mathcal{M}^{-1})^2 dx \quad (1.53)$$

y la relación señal a ruido es:

$$\frac{E^2}{S^2} = \frac{K_2^2}{3 N^2} \left\{ 1 + \frac{1}{\mu^2 S^2} + \frac{2}{\mu S^2} \int P(x) dx \right\} \quad (1.54)$$

iii) La LEY LOGARITMICA SENO HIPERBOLICO INVERSO se define por:

$$y = \frac{\text{sen } h^{-1} cx}{\text{sen } h^{-1} c} \quad 0 \leq x \leq 1 \quad (1.55)$$

Donde el tamaño de la etapa logarítmica a niveles altos tiene de asintóticamente a :

$$K_3 = \text{sen } h^{-1} c \quad (1.56)$$

Con las sustituciones apropiadas en la Ec. (1.46) en términos de la ley $\text{sen } h^{-1}$ el error medio cuadrático será:

$$E^2 = \frac{K_3^2}{3 N^2} \int P(x) (x^2 + c^{-2}) dx \quad (1.57)$$

en donde la relación señal a ruido es:

$$\frac{E^2}{S^2} = \frac{K_3^2}{3 N^2} \left\{ 1 + \frac{1}{c^2 S^2} \right\} \quad (1.58)$$

iv) La LEY COMPANDING HIPERBOLICA es diferente de las tres anteriores porque no hace una aproximación asintótica, sin embargo, el quantum se incrementa uniformemente con el nivel de la señal, y puede ser formulada completamente en un codificador secuencial con una entrada

$x = \frac{y}{m+1 - my}$ por lo que la ley hiperbólica es definida por:

$$y = \frac{(1+m)x}{1+mx} \quad (1.59)$$

Donde el tamaño de la etapa logarítmica no tiene una asíntota horizontal como en las otras leyes, pero tiene un valor mínimo de:

$$K_4 = \frac{4m}{1+m} \quad (1.60)$$

El ruido cuantizado es:

$$E^2 = \frac{K_4^2}{3 N^2} \int P(x) \left[\frac{(1+mx)^2}{4m} \right]^2 dx \quad (1.61)$$

y la relación señal a ruido será:

$$\frac{E^2}{S^2} = \frac{K_4^2}{3 N^2} \left[\frac{(1+s)^2}{4ms} \right]^2 \quad (1.62)$$

En la tabla 1.1, se hace una comparación tangible y práctica de las propiedades de cada una de las leyes companding.

En la figura 1.13 se muestra la relación señal a ruido calculada para la distribución exponencial (como voz) con 25 dB de ventaja companding y 128 niveles de cuantización o 7 dígitos binarios.

La línea punteada del lado derecho (b), muestra la proporción señal a ruido debido a cuantización uniforme con el mismo número de niveles, esta línea tiene pendiente unitaria porque el ruido es constante. La línea punteada del lado

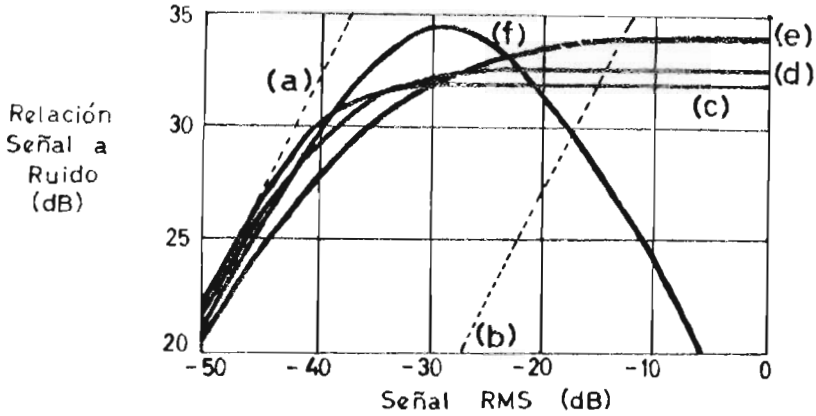


Fig. 1.13.- Relación señal a ruido para la distribución exponencial (como voz) con 128 niveles de cuantización para las cuatro leyes companding.

izquierdo (a) es el valor asintótico más bajo en todas las leyes companding. La diferencia entre a y b es la ventaja companding. La curva (c) define la ley A; la curva (d) define a la ley seno hiperbólico inverso; la curva (e) define a la ley μ ; y la curva (f) define a la ley hiperbólica. Dichas curvas exhiben un incremento progresivo con el incremento de la relación señal a ruido y las tres primeras (c, d, e) se mantienen aproximadamente constantes a partir del valor máximo, mientras que la cuarta cae.

La conclusión de lo anterior es que cualquiera de las leyes logarítmicas permiten una extensión del rango de voz. Además existen diversos métodos de incorporación de cuantización no uniforme en el diseño de un codificador y todas estas leyes companding pueden ser realizadas.

TABLA 1.1 PROPIEDADES DE LAS LEYES COMPANDING

TIPO DE LEY	ECUACION DE DEFINICION	$\frac{1}{x} \frac{dx}{dy} = \frac{1}{2} N$ veces el tamaño de la etapa logarit.		VENTAJA COMPANDING	
		EXPRESION GENERAL	VALOR MINIMO O ASINTOTICO		
Ley A o Ley Logarítmica con Segmentos Tangenciales Lineales	$y = \frac{Ax}{1 + \log A} \quad 0 \leq x \leq \frac{1}{A}$ $= \frac{1 + \log Ax}{1 + \log A} \quad \frac{1}{A} \leq x \leq 1$	$\frac{1 + \log A}{Ax}$ $0 \leq x \leq \frac{1}{A}$ $1 + \log A$ $\frac{1}{A} \leq x \leq 0$	$k_1 = 1 + \log A$ $Ax \geq 1$	$\frac{A}{1 + \log A}$	$\frac{A}{k_1}$
Ley μ o Ley Cuasi-Logarítmica	$y = \frac{\log(1 + \mu x)}{\log(1 + \mu)}$	$\frac{(1 + \mu x) \log(1 + \mu)}{\mu x}$	$k_2 = \log(1 + \mu)$ $\mu x \rightarrow \infty$	$\frac{\mu}{\log(1 + \mu)}$	$\frac{\mu}{k_2}$
Ley Seno Hiperbólico Inverso	$y = \frac{\text{senh}^{-1} Cx}{\text{senh}^{-1} C}$	$\frac{(\text{senh}^{-1} C) \sqrt{1 + C^2 x^2}}{Cx}$	$k_3 = \text{senh}^{-1} C$ $Cx \rightarrow \infty$	$\frac{C}{\text{senh}^{-1} C}$	$\frac{C}{k_3}$
Ley Hiperbólica	$y = \frac{x + mx}{1 + mx}$	$\frac{(1 + mx)^2}{(1 + m)x}$	$k_4 = \frac{4m}{1 + m}$ $mx = 1$	$1 + m$	

1.3.- FILTROS DIGITALES

El término filtro digital se refiere al proceso computacional o algoritmo por el cual una señal muestreada o secuencia de números actuando como una entrada, es transformada en una segunda secuencia de números denominada señal de salida. El algoritmo puede ser, filtración pasobajas, filtración pasobanda, interpolación, generación de derivadas, etc. El proceso se considera lineal y se utiliza el principio de superposición para analizar la relación entrada - salida.

Las clases de filtros digitales lineales son subdivididas en dos tipos:

- FILTROS FIR (RESPUESTA AL IMPULSO FINITA)
- FILTROS IIR (RESPUESTA AL IMPULSO INFINITA)

Ambos tipos de filtros pueden tener una realización recursiva y no recursiva.

En la realización recursiva la relación funcional entre la secuencia de entrada al filtro y la secuencia de salida puede ser descrita de la siguiente manera:

$$y(n) = F[y(n - 1), y(n - 2), \dots, x(n), x(n - 1)..] \quad (1.63)$$

donde la muestra de la salida actual $y(n)$ está en función de las salidas pasadas y también en función de las muestras de entrada actual y pasadas.

En la realización no recursiva, la relación entre las secuencias de entrada y las de salida está dada por:

$$y(n) = F[x(n), x(n - 1), \dots] \quad (1.64)$$

en donde las muestras de salida presentes $y(n)$ están en función solamente de las entradas pasadas y presentes.

El método de diseño para cada clase de filtro difieren marcadamente, por lo que son considerados separadamente. Cada clase de filtro tiene propiedades distintas.

El filtro FIR tiene memoria finita y puede tener características excelentes de fase, sin embargo, éste tiende a requerir de un gran número de términos para obtener un corte repentino.

Por otro lado, el filtro IIR tiene una memoria infinita y requiere de menos términos para obtener el corte repentino. Con el ahorro de términos existen características poco excelentes de fase, sin embargo, los filtros con corte repentino son mucho más fáciles para diseñar utilizando una estructura recursiva. (5)

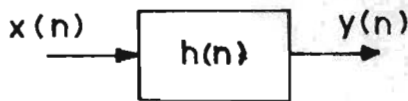


Fig. 1.15.- Diagrama a bloques del filtro digital.

En la figura 1.15 se muestra el diagrama a bloques de un filtro digital, donde $x(n)$ es la entrada, $y(n)$ es la salida y $h(n)$ es la respuesta al impulso.

Los filtros digitales lineales de transferencia invariante, relacionan la entrada y la salida con las expresio-

nes siguientes:

$$y(n) = \sum_{k=-\infty}^{\infty} h(k) x(n-k) = h(n) * x(n) \quad (1.65a)$$

$$y(n) = \sum_{k=-\infty}^{\infty} x(k) h(n-k) = x(n) * h(n) \quad (1.65b)$$

La relación correspondiente en términos de la transformada Z está dada por:

$$Y(Z) = H(Z) X(Z) \quad (1.66)$$

La transformada Z de la respuesta unitaria de la muestra $H(Z)$, es llamada la función del sistema.

La siguiente ecuación diferencial es una forma general que sirve para implementar filtros lineales de transferencia invariante.

$$y(n) - \sum_{k=1}^N a_k y(n-k) = \sum_{r=0}^M b_r x(n-r) \quad (1.67)$$

Evaluando la transformada Z en ambos lados de la ecuación anterior nos dá.

$$H(Z) = \frac{Y(Z)}{X(Z)} = \frac{\sum_{r=0}^M b_r Z^{-r}}{1 - \sum_{k=1}^N a_k Z^{-k}} \quad (1.68)$$

donde M y N son el orden del polinomio del numerador y del denominador respectivamente.

1.3.1.- SISTEMAS FIR

Los sistemas FIR son sistemas lineales de transferencia invariante con RESPUESTA al IMPULSO de duración FINITA.

Estos filtros no tienen polos, es decir, todos los coeficientes a_k son cero en la Ec. (1.68), por tanto la ecuación diferencial general para éste tipo de filtros nos queda:

$$y(n) = \sum_{r=0}^M b_r x(n-r) \quad (1.69)$$

la cual muestra que los filtros FIR son filtros "all - zero" donde los ceros son las raíces de $H(z)$.

Una de las propiedades importantes de los sistemas FIR es poder tener una fase lineal exacta, que se puede obtener a partir del polinomio dado por:

$$H(z) = \sum_{n=0}^M h(n) z^{-n} \quad (1.70)$$

La posibilidad de fase lineal exacta es útil en aplicaciones de procesamiento de voz, donde un alineamiento preciso de tiempo es esencial.

1.3.2.- SISTEMAS IIR

Los sistemas IIR son sistemas lineales de transferencia invariante con RESPUESTA al IMPULSO de duración INFINITA.

Estos filtros contienen tanto polos como ceros, por lo que la ecuación diferencial (1.67), puede ser escrita como:

$$y(n) = \sum_{k=1}^N a_k y(n-k) + \sum_{r=0}^M b_r x(n-r) \quad (1.71)$$

La transformada Z de esta ecuación será igual a la ecuación (1.68), la cual muestra la estructura polos - ceros de los filtros IIR.

Los polos se encuentran en las raíces del polinomio $a_k z^{-k}$, y los ceros en las raíces del polinomio $b_r z^{-r}$.

1.3.3.- VENTAJAS Y DESVENTAJAS DE FILTROS FIR E IIR (15)

Entre las ventajas y desventajas que tienen los filtros FIR se pueden mencionar:

- 1.- Es fácil y exacto el diseño de fase lineal.
- 2.- La respuesta a una frecuencia arbitraria puede ser fácilmente aproximada a una respuesta al impulso lo suficientemente largo.
- 3.- Los filtros son siempre estables
- 4.- Los filtros tienen buenas propiedades de cuantización.
- 5.- Técnicas de diseño excelentes están en disposición para la mayoría de clases de filtros.
- 6.- El número de coeficientes del filtro requerido para filtros de corte repentino es generalmente grande.

Por otro lado, las ventajas y desventajas que tienen los filtros IIR son:

- 1.- La aproximación es fácil de las características de magnitud arbitraria.
- 2.- Diseño muy eficiente, especialmente para filtros de cor

te repentino (con un número pequeño de polos y ceros).

- 3.- Técnicas de diseño clásico para diseño de filtros pasobajas y pasobanda están disponibles.
- 4.- Fácil prueba de estabilidad; si in filtro es inestable, modificaciones simples pueden hacerse para asegurar la estabilidad sin cambios en la magnitud de respuesta del filtro.
- 5.- Los efectos de la extensión finita de la palabra es bien entendidaa y puede ser fácilmente controlada.
- 6.- Puede haber alguna buena aproximación en el diseño de fase lineal.

Por lo anterior, para el diseño práctico de filtros, el diseñador tiene requerimientos específicos, ya sea para un filtrado en el dominio en frecuencia y/o en el dominio del tiempo. Es decir, será necesario instalár el rango de muestreo, el tipo en forma detallada de su estructura, el número de format, el dato y coeficiente de tiempo de la palabra y la técnica de cuantización.

La elección entre los filtros FIR, es llevado a cabo directamente en el dominio digital. Usualmente, la respuesta al impulso finita es derivada directamente, ya que en muchas implementaciones se hace uso directo de los coeficientes numéricos del filtro.

Aunque los filtros FIR pueden ser diseñados con fase arbitraria, la mayoría de las aplicaciones involucran la fase lineal exacta, debido a que ésto es fácilmente realizado. Para éste tipo de filtros, los ceros están distribuidos a través del plano Z y los polos tienen un valor cero.

Los métodos de diseño apropiado y códigos software han

sido programados y esos programas proveen métodos extremadamente versátiles para generar filtros pasobajas, pasoaltas, pasobanda, pasobanda múltiple y banda de paro o intercepción, Entre los métodos podemos mencionar a tres que son:

- Diseño Ventana (Window).
- 1 " Frecuencia de Muestreo.
- " Optimo.

La implementación de filtros FIR de fase lineal está frecuentemente dada en forma de recta como se muestra en la figura 1.16, aunque la cascada de secciones de segundo y cuarto orden son también usuales y eficientes.

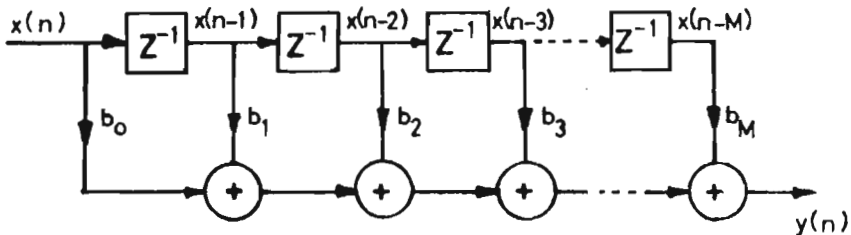


Fig. 1.16.- Representación en diagrama a bloques de un sistema FIR.

El diseño de filtros IIR pasobanda se realiza utilizando filtros analógicos con dominio en el plano S de polos y ceros, y usando transformaciones para obtener las singulari

dades en el dominio Z .

La magnitud del filtro IIR tiene frecuentemente menor orden y por lo tanto una realización hardware más eficiente, pero es realmente diseñado cuando la magnitud y las características de fase son requeridas (filtros all pass).

Cuando se requiere un diseño de magnitud no estandar, algunas técnicas de optimización son directamente aplicadas sobre las singularidades en el plano Z , para satisfacer los requerimientos necesarios. Una amplia variedad de métodos de implementación para éste tipo de filtros son:

- Diseño Butterworth.
- " Bessel.
- " Chebyshev.
- " Elípticos.

Estos diseños se basan sobre transformaciones de procedimientos de diseño clásico analógico. La implementación de filtros IIR ocurre frecuentemente como cascada de secciones de segundo orden, como se muestra en la figura 1.17.

Finalmente, con respecto a filtros digitales podemos hablar de filtro digital no lineal, como lo es el DPLL (Digital Phase Lock Loops), que se usa para sincronización y demodulación. Se puede decir, que DPLL es semejante a un PLL analógico, sin embargo, la mayoría de los DPLL usados en la práctica no lo son exactamente, pero se puede decir más bien que toman ventaja de la naturaleza peculiar del procesamiento digital de señales para realizar la misma función.

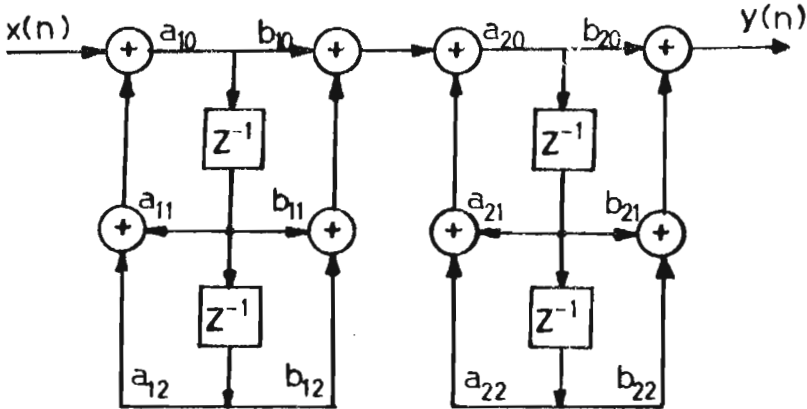


Fig. 1.17.- Representación de un sistema IIR en diagrama a bloques.

1.4.- RESUMEN

En este capítulo se han presentado los conceptos básicos para la comprensión del procesamiento digital de señales, así como algunas herramientas empleadas en su análisis.

El amplio uso, así como su comprensión de la teoría del procesamiento digital de señales se ha incrementado constantemente debido a los avances tecnológicos que han permitido un desarrollo más acelerado en el campo de las comunicaciones; dicho procesamiento se aplica en sistemas de comunicaciones de datos en banda de voz, sistemas de radar, etc.

C
A
P
I
T
U
L
O

2.- TECNICAS DIGITALES PARA IMPLEMENTAR SINTESIS DE VOZ

En años recientes, se ha desarrollado rápidamente la generación automática de formas de onda de voz originadas por datos digitales, que es llamada SINTESIS DE VOZ. La síntesis de voz involucra la conversión de una secuencia comando o texto de entrada (palabras y oraciones), dentro de formas de onda de voz, usando algoritmos y previamente codificando datos de voz. El texto puede ser introducido por teclado, reconocimiento óptico de caracteres, o desde una base de datos previamente almacenada. En la figura 2.1 se muestra el diagrama a bloques de un sintetizador de voz.

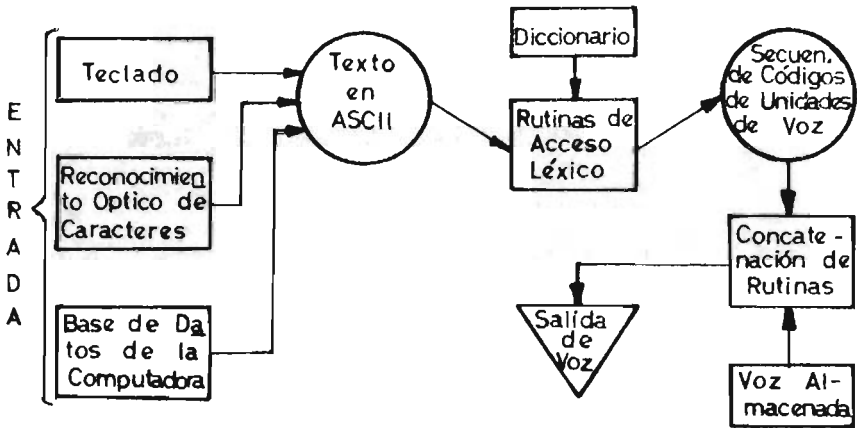


Fig. 2.1.- Diagrama a bloques de un sintetizador de voz.

La voz sintética fué generada primeramente en grandes computadoras, algunas veces interfascada con un modelo del tracto vocal analógico, antes que los chips de procesamiento

de señales de propósito especial fueran fabricados.

El campo de la voz digital se está volviendo rápidamente en un atractivo como una área de oportunidades para aplicaciones comerciales, y para la interacción hombre - máquina. La tecnología en esta área está siendo estimulada por adelantos en otros campos. Algoritmos nuevos han sido desarrollados para una codificación de voz eficiente en forma digital, gracias a la utilización de las ventajas de las propiedades de producción y percepción de la voz, todo esto se combina para el mejoramiento de la metodología de la síntesis de voz.

Se puede decir que la mayoría de los sintetizadores de voz, usan un ancho de banda de 3 KHz y 5KHz (aplicaciones telefónicas y para tener una calidad mayor, respectivamente) El uso de estas frecuencias se debe a que las vocales son percibidas muy bien con una frecuencia de 3 KHz, debido a que pueden ser especificadas por los tres primeros "formants" o resonancias del tracto vocal (la frecuencia "formant" es inversamente proporcional a la longitud del tracto vocal, entonces en un adulto masculino ocurrirá con un promedio de 1000 Hz, y en un niño la tercera resonancia ocurre a los 300 Hz). La percepción de algunas consonantes es dañada ligeramente, si no se conserva este ancho de banda.

Los sintetizadores de voz comunes presentan un compromiso (trade-off) entre las demandas de maximización de calidad de voz, aunadas a la minimización del espacio de memoria, complejidad del algoritmo y el tiempo de programación. Los sintetizadores de voz pueden ser caracterizados por el tamaño de las unidades de voz que se concatenan para producir la voz de salida, así como por el método usado para codificar, almacenar y sintetizar voz.

2.1.- TAMAÑO DE LAS UNIDADES DE VOZ

Podemos hablar de unidades largas o grandes y unidades pequeñas de voz. Las unidades largas, tales como frases u oraciones se concatenan en los sintetizadores más simples, en este tipo de unidades, se introduce alta calidad de voz a la salida (dependiendo del método de síntesis), pero necesita grandes cantidades de memoria para que todas las palabras o frases sean sintetizadas, después de que han sido habladas aisladamente.

Para una voz sintética con máxima naturalidad, cada palabra o frase debe ser originalmente pronunciada con la apropiada sincronización y entonación, para todas las oraciones en las cuales serán usadas. Así, si una palabra ocurriera en varias estructuras sintéticas, la pronunciación será recordada y almacenada. Esto se logra concatenando las palabras originalmente habladas en aislamiento, lo que conduce a una pérdida de inteligibilidad y naturalidad. También existe una degradación en la calidad de la voz cuando se reduce la capacidad de la memoria.

Cuando nos referimos a unidades pequeñas de voz, hablamos de la unidad mínima de las palabras, llamada FONEMA (el lenguaje Español tiene 29 vocales y consonantes). Y se puede generar la voz desde una secuencia de fonemas, el problema es que la pronunciación de un fonema en una palabra o frase es fuertemente dependiente del contexto en que se encuentra.

La necesidad de calcular una apropiada entonación, dan como resultado sintetizadores complejos y una salida de voz menos natural. Sintetizadores híbridos (usados en sistemas de investigación), concatenan inmediatamente el tamaño de

unidades de voz almacenadas, utilizando la igualación de parámetros espectrales en los límites entre unidades. (3)

2.2.- PROPIEDADES BASICAS DE LA VOZ

El proceso de producción de la voz, comunmente se modela como la respuesta de un filtro lineal en función del tiempo (el tracto vocal entre el glotis y los labios), como se muestra en la figura 2.2.,

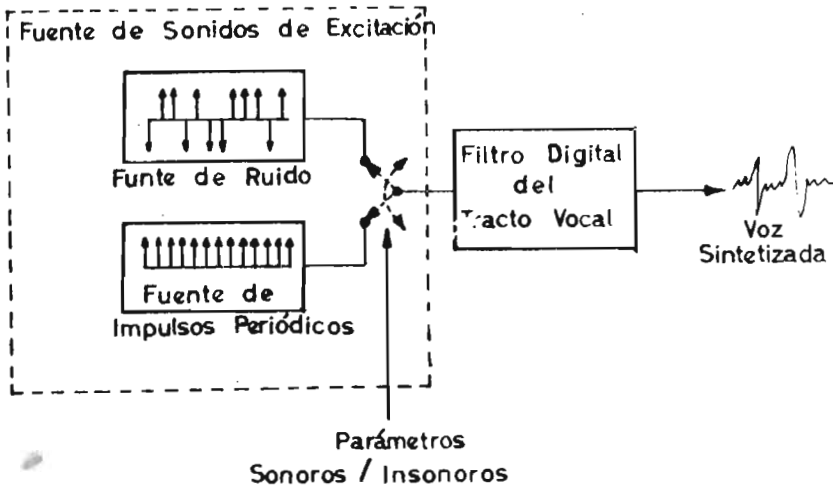


Fig. 2.2.- Representación en diagrama a bloques del modelo eléctrico básico de producción de voz.

En este modelo eléctrico, la voz es generada por una fuente de sonidos, la cual puede consistir de una excitación de una forma de onda de ruido, de una forma de onda de pulsos

periódicos, o bien de una combinación de los dos.

Como pulsos periódicos durante las regiones sonoras de voz, que es cuando el aire de los pulmones pasa entre las clausuras de las cuerdas vocales, es decir, cuando éstas es tan vibrando. Como ruido durante regiones insonoras de voz, a causa de las turbulencias en puntos de constricción en el tracto vocal. El filtro acústico es el que simula el tracto vocal, la excitación es transformada a parámetros espectrales.

Considerables redundancias muestran las señales de voz, debido a los mecanismos físicos del tracto vocal humano y de la estructura inherente del lenguaje. Así mismo, a causa de los mecanismos físicos del oído, y debido a nuestra habilidad para percibir sonidos de voz. Las propiedades de producción y percepción de voz están resumidas en la tabla 2.1.(4)

Para observar las propiedades de producción de voz, tomaremos el ejemplo que se muestra en la figura 2.3,

En la figura 2.3a, podemos observar periodos de tiempos largos y cortos de voz. En los tiempos largos vemos el comportamiento no estacionario, debido a la variación de la amplitud, y el comportamiento "sonoro/insonoro/silencioso", se ve variando con el tiempo (figura 2.3a'). En los tiempos pequeños (figura 2.3a'') se puede apreciar que es localmente estacionaria.

El modelo espectral de tiempo pequeño para voz sonora e insonora, se muestra en la figura 2.3b. Como se puede observar, en el segmento sonoro se forman formas de onda de estrutura pseudoperiódica, mientras en los insonoros, tiene

TABLA 2.1.- PROPIEDADES BASICAS DE PRODUCCION Y PERCEPCION DE VOZ

A.- PRODUCCION DE VOZ

1.- NO ESTACIONARIA EN TIEMPOS LARGOS

- Amplitud no estacionaria.
- Regiones * Sonoras/Insonoras/Silencios *
- Estructura del lenguaje y la fonética.

2.- ESTACIONARIA EN TIEMPOS CORTOS

- Estructura Formant (correlación en tiempos pequeños).
- Estructura de Tono "Pitch" (sonoros).
- Estructura de Ruidos (insonoros).

B.- PERCEPCION DE VOZ

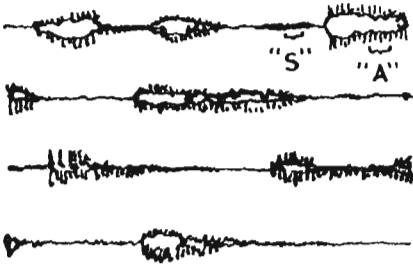
- Rango dinámico espectral local.
 - Enmascaramiento o disimulamiento auditivo.
-
-

una apariencia de ruido.

La envolvente del espectro, está determinada por la respuesta a la frecuencia del filtro o del tracto voca. Las resonancias del filtro, se observan como picos en la envolvente espectral, y es referida como formants.

En el proceso de percepción de la voz, el oído actúa como un analizador espectral de Fourier o como un banco de filtros, en donde cada canal del banco de filtros tiene su propio control automático de ganancia y sus limitaciones del rango dinámico.

Segmentos Largos de Voz (600 mseg)



Segmentos Pequeños de Voz (25 mseg)

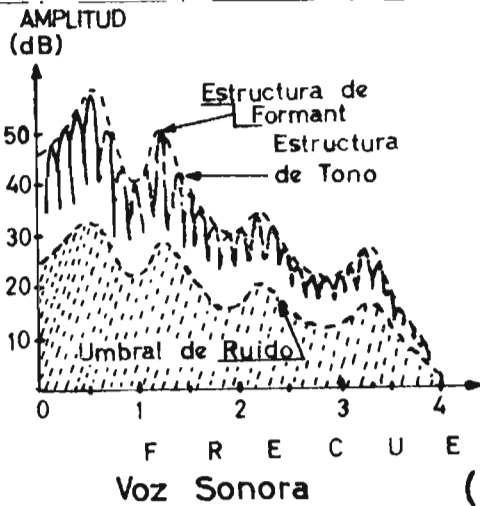


Segmento "A" (Sonora)



Segmento "S" (Insonora)

(a)



(b)

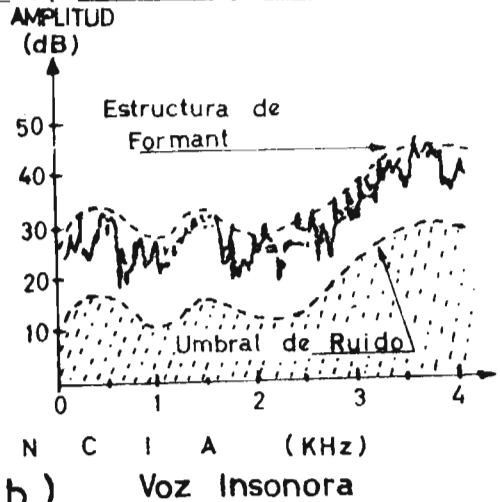


Fig. 2.3.- Segmentos de voz para visualizar las propiedades de la voz;

a) Segmentos de voz en tiempos largos y pequeños,

b) Modelo espectral para voz sonora e insonora,

Debido a la construcción del oído, se produce el fenómeno de enmascaramiento auditivo, con lo que con señales que su espectro en frecuencia se encuentra fuera de la gama de los 20 Hz a 20 KHz, no son audibles. El oído igualmente es insensitivo a desfases, a distorsiones y ruidos que están fuera de la banda audible (ultrasónica o infrasónica). El umbral del ruido enmascarado se puede apreciar en la figura 2.3b.

Tomando ventaja de las redundancias y restricciones en la producción y percepción de voz, es posible comprimir la razón de bits necesario para codificar y transmitir señales de voz. Las diversas técnicas de codificación hacen uso de estas ventajas.

2.3.- METODOS Y TECNICAS DE SINTESIS

Los sintetizadores pueden ser clasificados dentro de dos clases:

- SISTEMAS TEXTO A VOZ
- SISTEMAS DE RESPUESTA SONORA DE VOZ,

Los primeros sintetizan la voz constructivamente desde un texto, usando unidades pequeñas de voz, y procesando una lingüística extensiva. La segunda, produce la voz directamente de voz codificada previamente, utilizando técnicas de procesamiento de señales.

También pueden ser clasificados los sintetizadores, por como parametrizan la voz para su almacenamiento y síntesis, los cuales pueden quedar de la siguiente forma: (4)

- CODIFICADORES DE FORMA DE ONDA
- CODIFICADORES DE LA FUENTE (VOCODERS) O PARAMETRICOS
- CODIFICADORES HIBRIDOS

2.3.1.- CODIFICACION DE FORMA DE ONDA

La codificación de forma de onda, se caracteriza por que tratan de reproducir exactamente la forma de onda de la señal original, cualquiera que sea la naturaleza de ésta (voz, música, datos en banda de voz, etc.), y consiste básicamente en muestrear, cuantizar y codificar. La codificación de forma de onda, puede ser:

- A.- EN EL DOMINIO DEL TIEMPO
- B.- EN EL DOMINIO DE LA FRECUENCIA

A.- TECNICAS EN EL DOMINIO DEL TIEMPO

Las técnicas de codificación de forma de onda en el dominio del tiempo, se resume en la tabla 2.2. Este tipo de codificadores, muestrean periódicamente las formas de onda de las unidades de voz (el mínimo se considera la razón de Nyquist), y se almacenan las muestras utilizando las técnicas digitales de la tabla,

La mayoría de los codificadores en el dominio del tiempo, buscan minimizar el error o diferencia media cuadrática originada por el ruido cuantizado (debido a la representación de muestras de voz analógica, usando un número finito de niveles de amplitud), entre la voz original y la versión sintética, usando muestras codificadas.

En los esquemas de síntesis en el dominio del tiempo

TABLA 2.2.- TECNICAS DIGITALES DE CODIFICACION EN EL DOMINIO DEL TIEMPO

- PCM	Modulación por Codificación de Pulsos (Lineal y Logarítmica)
- DPCM	Modulación por Codificación de Pulsos Diferencial
- ADPCM	Modulación por Codificación de Pulsos Diferencial Adaptivo
- DM	Modulación Delta
- ADM	Modulación Delta Adaptiva
- CVSD	Modulación Delta con Pendiente que Varía Continúamente
- APC	Codificación por Predicción Adaptiva

las cantidades de voz almacenadas en la memoria, son representaciones comprimidas de una forma de onda como función del tiempo. La acción del sintetizador, es simplemente desempaquetar dichas formas de onda que fueron almacenada, para producir una señal de voz a la salida,

Los circuitos VLSI para síntesis en el dominio del tiempo son simple, debido a que el "hardware" en esta técnica, no es necesario convertir el dato almacenado del dominio de la frecuencia al dominio del tiempo. (8)

La síntesis de voz en el dominio del tiempo es una tecnología joven, sin embargo, se han hecho rápidos progresos en la calidad de voz y en la razón de bit. El aspecto revolucionario de la técnica en el dominio del tiempo no es la implementación "hardware", sino el análisis que permite que

las formas de onda de voz puedan ser almacenadas en una forma altamente comprimida.

B.- TECNICAS EN EL DOMINIO DE LA FRECUENCIA

Las técnicas de codificación en el dominio de la frecuencia, es la segunda categoría básica de codificación de forma de onda. Dentro de estas técnicas tenemos:

- SBC Codificación en Sub-Bandas
- ATC Codificación Adaptiva por Transformada

Ambos métodos proporcionan buena calidad de voz a una razón de bit de 16 Kbps.

Los esquemas de síntesis en el dominio de la frecuencia modelan la voz humana como una combinación de un pequeño número de fuentes de excitación (oscilaciones de las cuerdas vocales y turbulencia de aire), y una variedad más grande de estados de filtros de salida (representación de los estados de resonancia del tracto vocal). La compresión de datos se realiza, almacenando la excitación del tracto vocal y se filtran los parámetros, en lugar de la forma de onda original.

El "hardware" para síntesis en el dominio de la frecuencia, contiene circuitos integrados relativamente complejos (muchos polos analógicos o filtros digitales), que transforman los parámetros en el dominio de la frecuencia dentro de una señal de audio en el dominio del tiempo, con un espectro en frecuencia en función del tiempo, el cual aproxima el espectro de la forma de onda original.(8)

Las técnicas en el dominio del tiempo y en el dominio

de la frecuencia, serán tratados más ampliamente en el capítulo siguiente.

2.3.2.- CODIFICACION DE LA FUENTE (VOCODING) O PARAMETRICA

Los "vocoders" sintetizan la voz en base a un modelo que simula el aparato productor de voz humana, e intenta con seguir el mismo espectro de la voz original, en lugar de re crear la forma de onda muestra por muestra. Los canales "coder" están entre los tipos más primitivos históricamente de técnicas de compresión de voz.

Los codificadores de la fuente, almacenan un número li mitado de parámetros que son relevantes perceptualmente de las unidades de voz, para reducir el espacio de voz. La voz estará dividida dentro de pequeños marcos (10 a 30 mseg) y se determinan los parámetros describiendo el espectro y la periodicidad de la voz dentro del marco.

Se puede hacer un análisis tosco de la amplitud del es pectro, aprovechando que el oído es relativamente insensitivo a la distorsión de fase. En la elección del tamaño del marco, hay un compromiso entre almacenaje y calidad de voz, debido a que si los marcos son más pequeños, entonces requerirán más almacenaje.

En la tabla 2.3, se muestran las técnicas de codificación de la fuente.

El banco de filtros fue la primera construcción realizada para el analizador y el análisis espectral del banco de filtros se puede lograr utilizando la transformada discreta de Fourier.

TABLA 2.3.- TECNICAS DE CODIFICACION DE LA FUENTE

-
- Banco de Filtros
 - Vocoder de Canal
 - Modelo Homomórfico
 - Autocorrelación
 - Modelo de Formant
 - Código de Predicción Lineal (LPC)
-

Este banco de filtros, fué sustituido por el canal "vocoder" en el analizador. Los parámetros de excitación obtenidos por el analizador, son usados para generar una excitación consistente de pulsos espaciados por el periodo de tono (pitch) para la voz sonora y una secuencia de ruido para los sonidos insonoros.

En la función de autocorrelación de tiempo pequeño de la onda de la voz, si la longitud de la ventana o marco de autocorrelación contiene varios periodos de tono, la función de autocorrelación exhibirá picos en los múltiplos de los periodos de tono. Consecuentemente con estos picos, la frecuencia fundamental puede ser determinada. Por otro lado, la ausencia de picos sirve para identificar intervalos insonoros o de silencio.

La forma de onda de la voz, consiste de una convolución de la función de excitación y la respuesta al impulso del tracto vocal. Para la desconvolución de la voz se puede adaptar el filtrado homomórfico, que es la no linealidad general de las técnicas de filtración.

Para sistemas de análisis y síntesis de voz basados en

la filtración homomórfica, el valor del "cepstral" de tiempo pequeño es usado como parámetro para representar el tracto vocal o su equivalente, la envolvente espectral de la voz. Los parámetros de excitación son tomados de la porción de tiempos largos del "cepstrum". Para voz sonora, ocurren picos en el "cepstrum" en los múltiplos del tono periódico, y para voz insonora no ocurren tales picos.

En el sintetizador homomórfico, la entrada consiste de un valor "cepstral" de tiempo pequeño y los parámetros de excitación. En el sintetizador, los valores del "cepstral" de tiempo pequeño son convertidos a una aproximación de la respuesta al impulso del tracto voca. Los parámetros de excitación, son usados para generar una función de excitación, la cual es convolucionada con la aproximación de la respuesta al impulso, obtenida en el "cepstral", con lo que resulta la voz a la salida. (2)

- SINTETIZADORES DE FORMANT

El sintetizador paramétrico que se ha empleado más frecuentemente utiliza codificación de "formant" y anchos de banda para representar el espectro almacenado. Esta técnica, separa la forma de onda dentro de una componente de excitación y una representación del filtro del tracto vocal. La excitación es un tren de impulsos periódicos para voz sonora y ruido blanco para voz insonora.

El tracto vocal es usualmente modelado como una serie de resonadores digitales de segundo orden, cada uno modelando un "formant" o el efecto de la cuerda vocal. Esto aproxima bien el espectro de voz para vocales, y puede ser controlado con un parámetro de amplitud.

Los sonidos insonoros tienen la fuente de excitación dentro de la cavidad vocal, excitando primeramente el tracto vocal entre los labios y la fuente, con lo que el tubo acústico está recortado y la mayoría de la energía está a frecuencias muy altas. Para estos sonidos, es comunmente usado un banco de resonadore en paralelo, ya que cada filtro puede ser individualmente controlado por amplitud.

Los sonidos nasales generalmente necesitan un resonador extra, debido a que la trayectoria acústica que incluye la cavidad nasal es más grande que para las vocales por lo que se incrementa el número de resonancias en el ancho de banda de la voz. La carencia de variación en tiempo en los anchos de banda, es más perceptible en este tipo de sonidos, que es tán sujetos a grandes pérdidas en la cavidad nasal, lo que resulta ser "formants" más amplios.

En los sonidos nasales, son encontrados ceros espectrales a frecuencias donde la cavidad vocal presenta impedancia cero para la presión de las ondas saliendo de la faringe. Los ceros pueden directamente ser introducidos utilizando un resonador inverso de segundo orden, dentro del filtro que si mula el tracto vocal. El resonador extra y el resonado inverso, se pueden considerar como un par de polos-ceros, con efectos de cancelación, excepto cuando sonidos nasales son simulados. Aín embargo, la localización de polos y ceros automáticamente en voz natural es una labor difícil, ya que una modelación cero con baja calidad es perceptible en estos sonidos.

- CODIFICACION DE PREDICCION LINEAL

Otro sintetizador ampliamente usado es el LPC, que al igual que el de "formant", separa la forma de onda dentro

de una componente de excitación y una representación del tracto vocal. Sin embargo, estos sintetizadores son más simples conceptualmente, todos los efectos espectrales están incluidos en los parámetros LPC, los cuales son calculados automáticamente en el análisis de la voz original.

La mayoría de los sistemas LPC son representados por el filtro del tracto vocal *solamente polo* (all pole). A causa de que el modelo LPC, no puede separar los efectos de la radiación de los labios, ni los efectos de las cuerdas vocales desde la respuesta al tracto vocal, existen generalmente un máximo de cuatro "formants" excitados abajo de los 4 KHz, y cada uno está bien modelado por un resonador de segundo orden, es decir, un par de polos complejos en el plano espectral Z . Para describir la filtración entre las cuerdas vocales y los labios, son suficientes 8 polos.

El filtro digital solamente polo de síntesis predice una muestra de voz basada sobre una combinación lineal de N muestras previas. El problema de los sistemas LPC solamente polos, es que la excitación glotal introduce ceros dentro del espectro cuando se producen sonidos nasales, debido a que éstos son radiados desde más de una parte del tracto vocal o cuando existe ramificación entre la fuente de excitación y la salida de voz. Los ceros generalmente causan modificaciones en resonancias adyacentes, especialmente en sus anchos de banda.

Si el filtro LPC es excitado por impulsos periódicos para voz sonora, el tracto vocal es excitado una vez por el período de tono (pitch) en el instante específico de tiempo, y la duración de la excitación, no es mucho más grande que un período de muestra. Excitando el filtro de síntesis con un impulso solamente, causa una calidad menor de voz.

Para reducir la razón de bit en el almacenamiento, puede ser usada una versión residual de la señal, la que podemos obtener si la voz sonora es pasada a través de un filtro cuya respuesta es el inverso del filtro de síntesis LPC, la salida de éste es la señal residual, lo que generalmente contiene impulsos fuertes en los tiempos cuando las cuerdas vocales están cerradas. Sin embargo, la distribución de la energía residual es significativa en todas partes del periodo de tono.

Los sintetizadores LPC, son más populares que los otros "vocoders", debido a su algoritmo sencillo, produciendo una representación espectral apropiada para una codificación económica.

En los "vocoders", se requiere a priori conoce el mecanismo por medio del cual se genera la voz, debido a que éstos simulan la voz humana. Además, los conocimientos de la estructura fonética y del lenguaje son importantes para aplicarlos en los parámetros del "vocoder", y de esta forma conseguir una voz sintética más natural.

Todas las técnicas de codificación de la fuente se explicaron ligeramente, pero las más importantes comercialmente y más frecuentemente utilizadas son las LPC y la de "formants".

2.3.3.- CODIFICACION HIBRIDA

Esta tercera clase de codificación, como su nombre lo indica, combina las características de codificación de forma de onda y la de codificación de la fuente, en un intento de llenar el vacío que existe entre el funcionamiento de ambos métodos. Dentro de esta clase de codificación tenemos:

- Vocoder Excitado por Voz (VE-V)
- Escalamiento Armónico

La primera técnica, utiliza métodos de codificación de forma de onda en la primera porción de la banda de voz, de 0 - 800 Hz, y el resto de la banda de voz entre 800 - 4000 Hz, utiliza la codificación de la fuente. La excitación de la fuente para la parte de la codificación de la fuente, la banda de voz se obtiene por una distorsión no lineal de la banda base de la forma de onda codificada.

En la segunda técnica, la señal de voz es comprimida paramétricamente en el ancho de banda y el rango de muestreo. Esta señal comprimida puede ser codificada digitalmente por métodos de codificación de forma de onda para obtener una razón de bit reducida. La señal comprimida, puede utilizarse para "acelerar" o "retardar" la voz.

El método de escalamiento armónico se enfoca primeramente sobre redundancias de la voz, debido a la estructura de tono y a que es estacionaria localmente, por lo que trabaja mejor con técnicas de codificación que no lo hacen depender sobre estas propiedades. Este método se ha realizado en el dominio del tiempo y también en el dominio de la frecuencia.

El escalamiento armónico en el dominio de la frecuencia es una técnica que comprime o expande el espaciamiento de las armónicas de tono. En el dominio del tiempo, es un método que elimina o comprime, o por el contrario inserta o expande los periodos de tono en un tono de manera sincrónicamente interpolados.

El área de codificación híbrida está activamente estu-

diada, para que la codificación de forma de onda y la codificación de la fuente se reduzca dentro de un cuadro más unificado.

2.3.4.- SISTEMAS DE TEXTO ILIMITADO (TEXTO A VOZ) Y SISTEMAS DE TEXTO LIMITADO (RESPUESTA SONORA)

Como se dijo anteriormente, también los sintetizadores son clasificados como:

- Sintetizadores de Texto a Voz.
- Sintetizadores de Respuesta Sonora.

Los sistemas de texto a voz tienen la necesidad de un procesador lingüístico para convertir el texto dentro de una forma apropiada para el acceso de unidades de voz almacenada. Por lo que deben trasladar las oraciones de entrada dentro de una secuencia de códigos para extraer las apropiadas unidades almacenadas, y determinar los parámetros de entonación del texto. Para poder realizar el traslado, se maneja por medio del conjunto de reglas del lenguaje para convertir una serie de letras dentro de una secuencia de fonemas.

Las reglas examinan el contexto de cada letra para determinar como ser pronunciadas. Un lenguaje como el Español necesita pocas reglas pues se puede escribir con buena ortografía siguiendo cercanamente la fonética, ya que cada letra tiene una pronunciación solamente.

Los sistemas de respuesta sonora son ampliamente utilizados en aplicaciones como teléfonos con asistencia automática de directorio, juguetes que hablan, sistemas de protección de máquinas y automóviles.

Estos sistemas codifican esencialmente la voz del usuario con eficientes métodos de almacenamiento, por lo tanto, están limitados por las combinaciones de las unidades originales de voz, usualmente, con las entonaciones originales.

2.4.- CALIDAD DE VOZ

La calidad de la voz es un factor importante de considerar en los sintetizadores; se puede mostrar el *estado del arte* de síntesis en términos de calidad de voz (el rango 0 - 1, es una medida hipotéticamente simplificada, donde el cero implica una pobre inteligibilidad, y el uno implica una calidad que es indistinguible esencialmente de la entrada original para una señal en el ancho de banda telefónica) como una función de la razón de bit, lo cual se muestra en la figura 2.4.

En esta figura, se muestran los tres métodos de codificación, de forma de onda, de modelación de la fuente y la híbrida. La calidad se muestra como una razón de bit, debido a que la disminución de la razón de bit es importante para que las técnicas de codificación ocupen un almacenamiento menor en la memoria.

Los métodos de codificación de forma de onda, sin hacer uso de las propiedades de la tabla 2.1, se puede obtener buena calidad de voz a una razón de bit sobre los 64 Kbps. Si se toma ventaja de la amplitud no estacionaria, se puede lograr una buena calidad con esquemas simples de codificación de forma de onda con la misma razón de bit.

Tomando ventaja de las características espectrales estáticas y dinámicas, se puede lograr una buena calidad a una razón de bit en los rangos de 24 - 32 Kbps. Una calidad acep

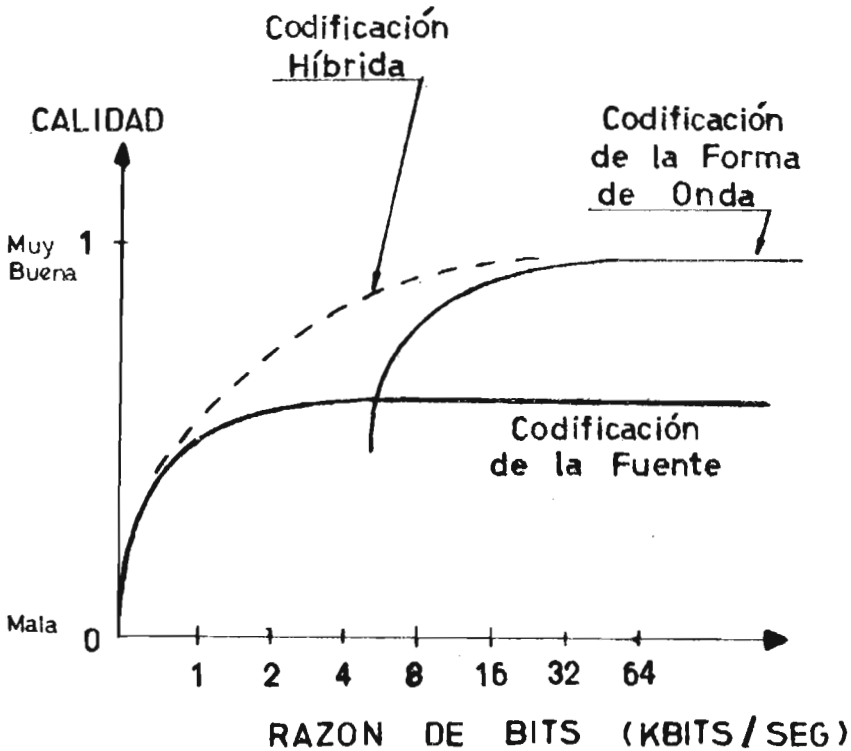


Fig. 2.4.- Representación gráfica de Calidad VS Razón de bit para las diferentes técnicas de codificación

table se puede lograrse con una razón de bit bajo de los 10 Kbps si se incluyen las propiedades de tono y ruido. La calidad de los métodos de forma de onda debajo del rango de los 10 Kbps, caen rápidamente como una función de la razón de bit y de las propiedades de la voz.

La razón de bit para una modelación de la fuente, pue-

de ser reducida al orden de los 2 Kbps, obteniéndose una voz como máquina. La calidad de esta técnica está limitada por el modelo de la señal "vocoder" y por las propiedades del lenguaje y la estructura fonética.

Las técnicas de codificación que han estado evolucionando son los métodos híbridos o combinados, los cuales pueden trabajar de 4 a 10 Kbps, donde al utilizar las ventajas de los métodos de forma de onda y de la modelación de la fuente, obtienen buena calidad de voz con este rango de razón de bit.

En todos los métodos de codificación, se pueden conseguir reducciones adicionales en la razón de bit en grandes periodos de tiempo, si se corta el codificador en las regiones de silencio, es decir, cuando el orador no está hablando. Realizando ésto, conduce a técnicas de rango variable de codificación, la razón promedio de bit dependerá de la actividad del orador y el porcentaje de tiempo que esté hablando.

2.5.- RESUMEN

La calidad de la voz que se desea lograr en el sistema de análisis y síntesis, la razón de transmisión, el costo de implementación, la tolerancia que existirá a los errores producidos en la transmisión, son factores que intervienen en la elección del tipo de codificador a utilizarse, e igualmente la técnica de compresión de datos, dependerá esencialmente de la aplicación que se le dará al sistema.

C
A
P
I
T
U
L
O

3.- TÉCNICAS DE COMPRESION DE SEÑALES DE FORMA DE ONDA

La representación de las señales en forma digital es en la actualidad de interés fundamental. Existen muchas maneras para representar las señales de voz en forma discreta. Como se vio en el capítulo anterior, estas formas discretas de representar la voz se clasifican en tres y son:

- REPRESENTACION DE FORMA DE ONDA
- _ REPRESENTACION DE LA FUENTE O VOCODER
- _ REPRESENTACION HIBRIDA

La representación de forma de onda como su nombre lo indica, implica preservar simplemente la *forma de onda* de la señal de voz analógica a través de un proceso de muestreo y cuantización.

El diagrama a bloques de la figura 3.1, describe en forma general la representación de la forma de onda. En ella, se ilustra la forma de onda de la voz, que puede ser idealizada como una función continua variable con el tiempo, la cual se muestrea generalmente a intervalos periódicos en el tiempo, para producir una secuencia de muestras $x(n)$. Para obtener una representación digital, es necesario cuantizar las muestras en un conjunto finito de valores apropiados.

3.1.- MUESTREO Y CUANTIZACION DE SEÑALES DE VOZ

En el capítulo 1, se explicó el teorema de muestreo. Ahí se mostró que una señal analógica puede ser reconstruida en forma exacta a partir de sus valores muestra, siempre y cuando la señal esté limitada en banda, y haya sido muestrada al rango de muestreo de Nyquist o mayor ($f \geq 2B$).

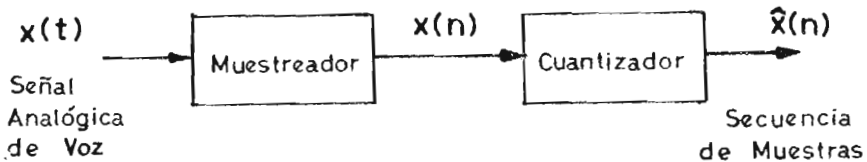


Fig. 3.1.- Diagrama a bloques que describe la representación de la forma de onda.

Dado que nos interesa la representación digital de las señales de voz, es necesario considerar sus propiedades espectrales y además, tener en cuenta que las señales de voz no son inherentemente de banda limitada, aunque el espectro tiende a caer rápidamente a altas frecuencias. En la práctica para asegurar la condición anterior, se introduce un filtro pasobajas de conte rápido antes del proceso de muestreo.

Por otro lado, para representar exactamente todos los sonidos de voz, se requiere un rango de muestreo mayor de 20 KHz, debido a que el espectro de sonidos insonoros no caen apreciablemente sobre los 8 KHz. Sin embargo, en la mayoría de las aplicaciones este rango no se requiere (por ejemplo, en telefonía), ya que dichas señales pueden ser especificadas con las tres primeras armónicas, por lo que solamente estaremos interesados sobre el rango del espectro sobre el límite de 3.5 KHz. Entonces, el rango de muestreo práctico de las señales de voz de acuerdo al teorema de Nyquist, es de 8 KHz.

Para un fácil entendimiento es conveniente considerar

en forma separada el proceso de muestreo y el de cuantización (en la práctica esto no es posible) en la representación digital de forma de onda. De esta manera se asume que la forma de onda de voz ha sido limitada en banda por un filtro pasobajas y después muestreada al rango apropiado dando una secuencia $x(n)$.

Esta señal $x(n)$ servirá entonces como una señal de entrada a cualquier proceso de PDS, o también para almacenarla directamente en una memoria digital, o bien para la transmisión sobre un canal de comunicaciones en forma digital.

Los valores de las muestras son cuantizadas en un conjunto finito de amplitudes, ya que éstas pueden ser representadas por un conjunto finito de símbolos. Dando como resultado el proceso de cuantización y codificación.

De la misma forma en que separamos el proceso de muestreo y el de cuantización, también es útil representar en forma separada el proceso de cuantización y codificación, esta acción se describe en la figura 3.2. Aquí, la secuencia de muestras $x(n)$, son representadas por un conjunto finito de símbolos $c(n)$.

También en la figura podemos apreciar las dos etapas para realizar la representación: Una etapa de cuantización, la cual produce una secuencia de amplitudes cuantizadas $\hat{x}(n) = Q \{ x(n) \}$, y una etapa de codificación, en la cual cada muestra cuantizada es representada por una palabra código $c(n)$. Todo lo anterior forma parte del codificador (figura 3.2a), donde Δ representa el tamaño de la etapa de cuantización en el cuantizador.

En el decodificador, la secuencia de palabras código

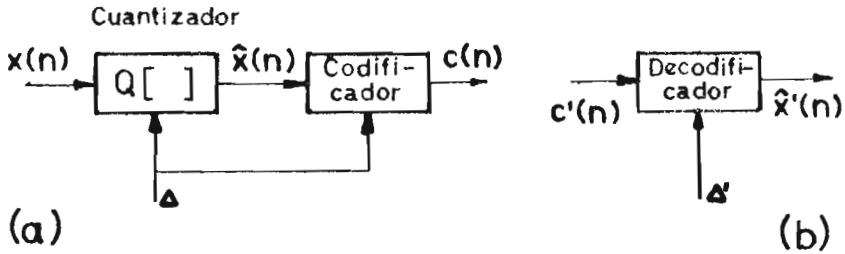


Fig. 3.2.- Proceso de cuantización y codificación:
a) Codificador b) Decodificador

$c'(n)$, se transforman dentro de una secuencia de muestras cuantizadas $\hat{x}'(n)$, que es descrito en la figura 3.2b. Si no hubo errores al recuperar $c'(n)$, entonces:

$$c'(n) = c(n)$$

por tanto $x'(n) = x(n)$

En la mayoría de los casos, se usan números binarios para representar las muestras cuantizadas. Con m bits palabras código binario, es posible representar 2^m niveles diferentes de cuantización.

La capacidad de información requerida para transmitir o almacenar la representación digital es:

$$I = (m)(f) = \text{RAZON DE BIT en bits/seg} \quad (3.1)$$

donde: f es el rango de muestreo (muestras/seg), y
 m es el número de bits/muestra

En general, es deseable mantener la razón de bit tan ba-

ja como sea posible, a la vez que se mantiene un nivel deseado de calidad. Por lo que, para un ancho de banda de voz, el rango de muestreo está fijado por el teorema de muestreo, por tanto, el camino para reducir la razón de bit solamente es reducir el número de bits/muestra. Lo anterior se logra utilizando una gran variedad de técnicas, las que a continuación se explican.

3.2.- MODULACION POR CODIFICACION DE PULSOS (PCM)

3.2.1. PCM LINEAL

Un método común y quizá la forma más simple de codificación de forma de onda es el PCM lineal, en el cual una señal analógica es uniformemente cuantizada, en este tipo los rangos y niveles de cuantización son distribuidos uniformemente.

Existen dos clases de cuantizadores uniformes:

- CUANTIZADOR UNIFORME DE MEDIA SUBIDA "MID-RISER"
- CUANTIZADOR UNIFORME DE MEDIO PASO "MID-TREAD"

En el cuantizador uniforme de media subida, el origen aparece en el centro de una parte de la subida de la función como escalera (figura 3.3a). Mientras que en cuantizador de medio paso, el origen aparece en el centro de una parte del escalón de la función como escalera (figura 3.3b).

Para un esquema de código binario, el cuantizador de media subida tiene el mismo número de niveles positivos y negativos, y está simétricamente posicionado alrededor del origen.

El cuantizador de medio paso, tiene un nivel más negativo, en este caso uno de los niveles de cuantización es cero,

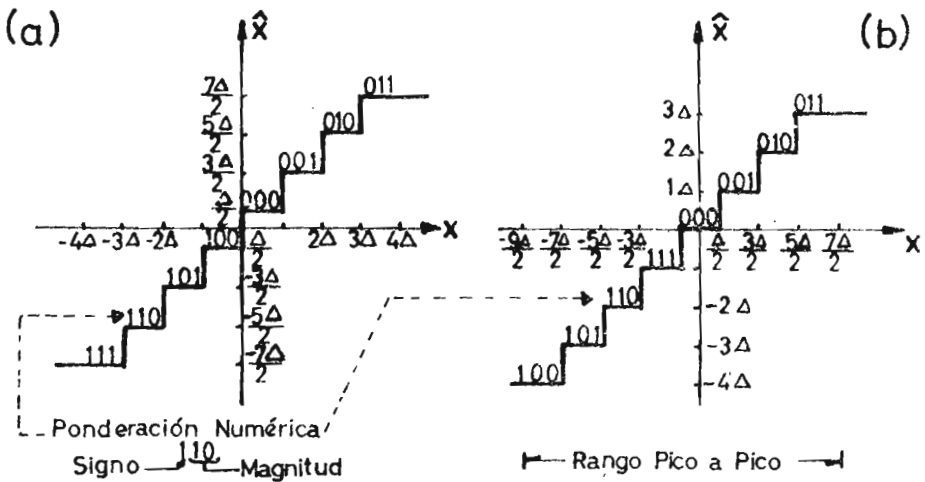


Fig. 3.3.- Representación de las características de los cuantizadores uniformes: a) Cuantizador de media subida; b) Cuantizador de medio paso.

mientras que el de media subida hay un nivel no cero.

En la figura también, se muestran las palabras código asignadas que sirven para tener una ponderación numérica, interpretando las palabras código binario como una representación signo - magnitud.

La ponderación es una palabra codificada en binario, en la cual nos da el signo (por ejemplo, "0" es un dígito que nos indica signo positivo, el dígito "1" nos indica signo negativo), y los dos bits restantes nos representan la magnitud entonces las muestras cuantizadas se relacionan con las palabras código por medio de lo siguiente:

$$\hat{x}(n) = \frac{\Delta}{2} \text{sign} \left\{ c(n) \right\} + \Delta c(n) \quad (3.2)$$

donde el sign $\{c(n)\}$ es igual a +1 si el primer bit de $c(n)$ es cero y -1 si el primer bit de $c(n)$ es uno.

En la figura 3.3b, las palabras codificadas en binario se interpretan como representaciones de dos complementos de 3 bits, donde las muestras cuantizadas son relacionadas con las palabras codificadas por la relación:

$$x(n) = \Delta \{ c(n) \} \quad (3.3)$$

Para los cuantizadores uniformes, únicamente existen dos parámetros, el número de niveles (generalmente elegidos de la forma 2^m), y el tamaño de la etapa de cuantización Δ . Por lo que se elegirán adecuadamente Δ y m , para cubrir el rango de muestras de entrada.

Por otro lado, la consideración que se hace para cualquier tipo de cuantización es la relación señal a ruido, de acuerdo a lo estipulado en el capítulo 1, Ec. (1.42), el SNR para cuantización uniforme es:

$$\text{SNR}_{\text{dB}} = 10.8 + 6 m \quad (3.4)$$

donde m es el número de dígitos binarios de las palabras de ponderación.

Si en esta ecuación $m = 7$, tendremos un SNR de alrededor de 52.8 dB, en el cual es apropiado para proporcionar una calidad adecuada en un sistema de comunicaciones. Aunque para la representación de la señal de voz con alta calidad, en un cuantizador uniforme se utilizará m igual a 11 o 12 bits, dado que no toma ventaja de las propiedades de la voz enunciadas en la tabla 2.1; y en consecuencia no tiene capacidad de compresión de datos.

En la cuantización uniforme, la relación señal a ruido es dependiente del nivel de la señal, por lo que será deseable que el error sea de varianza constante, es decir, independiente de la amplitud de la señal, obteniendo así un porcentaje de error constante que puede ser logrado con cuantización no uniforme o logarítmica.

3.2.2.- PCM LOGARITMICO

Cuando se hace uso de una cuantización no uniforme de la amplitud de la señal, de acuerdo a una escala logarítmica en lugar de una escala lineal se habla de un PCM logarítmico, donde los niveles de cuantización están logarítmicamente espaciados.

Como se puede apreciar en la figura 3.4, el logaritmo de la entrada, es decir, $\log x(n)$ es cuantizada en lugar de la entrada $x(n)$ como en el caso lineal, es decir, realizandose el "companding", debido a que en el codificador las amplitudes de entrada son comprimidas por la función logarítmica antes de la cuantización, y después son expandidas en el decodificador por la función exponencial. Esto conduce a una insensibilidad a la amplitud de la señal, por tanto, el porcentaje de error es constante, y la relación señal a ruido es independiente de la varianza de la señal y dependiendo solamente del tamaño de la etapa del cuantizador.

De la figura se ve que:

$$y(n) = \ln \{x(n)\} \quad (3.5)$$

por tanto:

$$x(n) = \exp \{y(n)\} \text{sign} \{x(n)\} \quad (3.6)$$

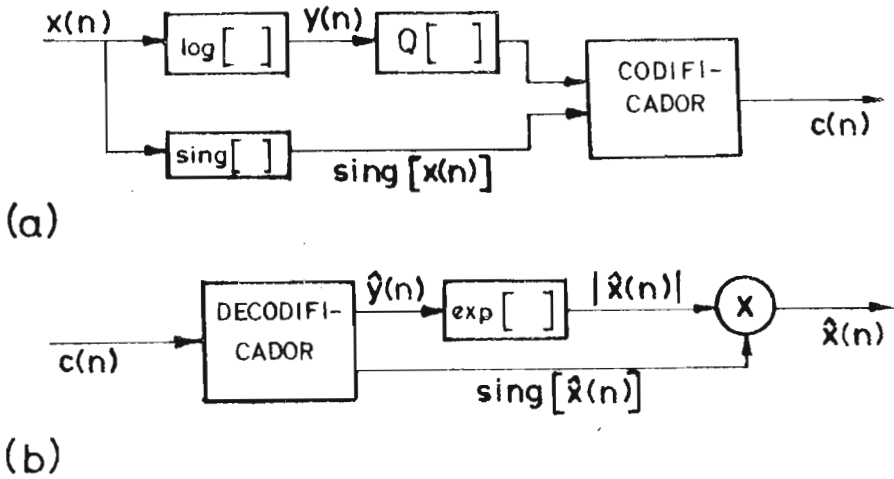


Fig. 3.4.- Diagrama a bloques de la cuantización logarítmica
 a) Codificador b) Decodificador

donde el sign $\{x(n)\}$ es +1 si $x(n)$ es positivo y -1 si $x(n)$ es negativo. La magnitud logarítmica cuantizada es:

$$\begin{aligned}
 y(n) &= Q \left\{ \log |x(n)| \right\} \\
 &= \log |x(n)| + \epsilon(n)
 \end{aligned}
 \tag{3.7}$$

pero $\epsilon(n)$ es independiente del $\log x(n)$, tenemos que:

$$\begin{aligned}
 \hat{x}(n) &= \exp \left\{ \hat{y}(n) \right\} \text{ sign } \{x(n)\} \\
 &= |x(n)| \text{ sign } \{x(n)\} \exp \left\{ \epsilon(n) \right\} \\
 &= x(n) \exp \left\{ \epsilon(n) \right\}
 \end{aligned}
 \tag{3.8}$$

donde $\text{sign } \{x(n)\} = \text{sign } \{\hat{x}(n)\}$.

En la actualidad existen varias técnicas logarítmicas de cuantización, siendo las más utilizadas la ley μ y la ley A PCM. En dichas codificaciones utilizan las características estáticas de la amplitud no estacionaria de la voz para conseguir buena calidad a razón de bit de 56 y 64 Kbps.

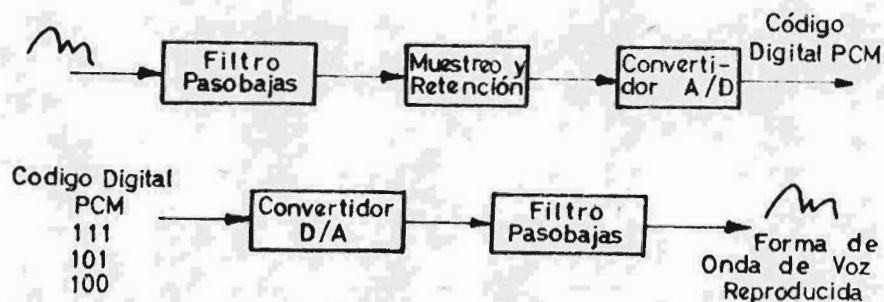


Fig. 3.5.- Diagrama a bloques de un sistema de reproducción de voz que emplea PCM.

3.3.- MODULACION DELTA (DM)

Métodos opcionales para convertir las señales analógicas a un formato digital se han propuesto para algunas aplicaciones, con la esperanza de reducir el ancho de banda requerido, mejorar la eficiencia o reducir el costo.

La modulación Delta (DM) es uno de estos métodos y ha sido adoptado para algunas aplicaciones que requieren la transmisión de voz o de imágenes. En tales aplicaciones, existe un alto grado de redundancias en la información que se va a transmitir.

La información pasada contiene también parte de la información actual, como para poder predecirla aceptablemente

bien, de manera que las señales nuevas sólo necesitan ser transmitidas si se producen cambios significativos en ella. Este proceso asume que las formas de onda de las señales de entrada serán bastante uniformes y con una pendiente predecible.

En la DM el rango de muestreo es muchas veces la razón de Nyquist para la señal de entrada, y ésto se lleva a cabo para reducir el ruido que introduce el sistema y aparece a la salida del receptor. Como resultado del alto rango de muestreo, las muestras adyacentes están altamente correlacionadas.

Este alto grado de correlación implica que seremos capaces de predecir mejor la muestra actual de las muestras pasadas, por tanto, la varianza de error de predicción será baja. De este modo, la razón de bit de un sistema DM es simplemente igual al rango de muestreo.

En la figura 3.6, se muestra el sistema más simple de DM, que es la Modulación Delta Lineal (LDM). Donde el cuantizador tiene solamente dos niveles y el tamaño de la etapa está fija.

Del diagrama a bloques, los valores de las muestras $x(n)$ se comparan con un valor muestra predicho denominado $\tilde{x}(n)$, y la diferencia $d(n)$ se hace pasar por el cuantizador. Obviamente, si $d(n)$ es pequeña durante la mayor parte del tiempo, de manera que la predicción es buena, se necesitarán pocos bits para representar esta señal diferencia.

El modulador delta, utiliza un cuantizador de dos niveles, de manera que sólo un bit se emplea para representar la señal. Los dos niveles del cuantizador son:

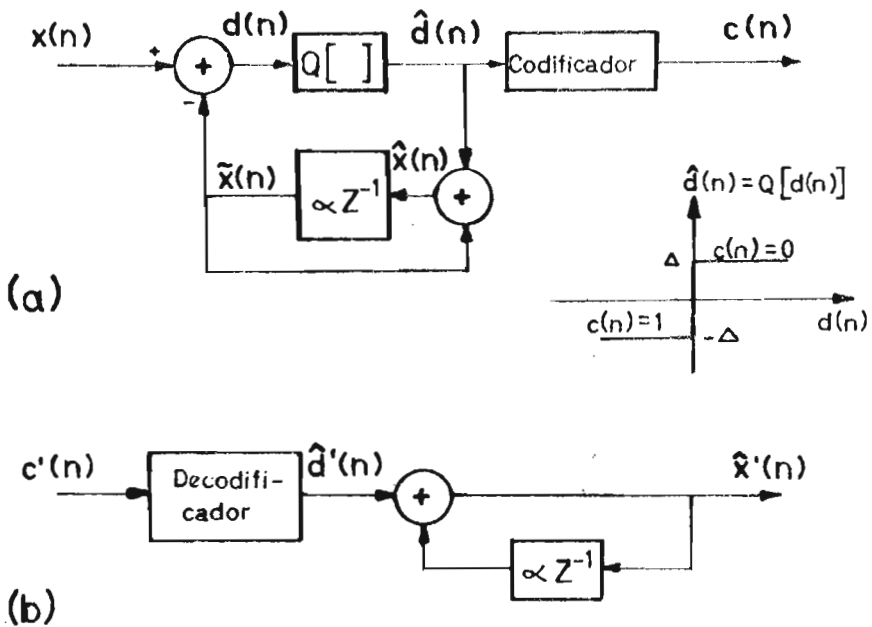


Fig. 3.6.- Diagrama a bloques de un sistema de Modulación Delta; a) Codificador b) Decodificador

$c(n) = 0$ si el nivel de cuantización es positivo, y
 $c(n) = 1$ si es negativo, entonces:

$$\begin{aligned} D \ d(n) &= \Delta && \text{si } c(n) = 0 \\ &= -\Delta && \text{si } c(n) = 1 \end{aligned}$$

En el receptor, la señal diferencia cuantizada se agrega a la salida del predictor para obtener una estimación discreta $\hat{x}'(n)$ de la señal muestreada deseada $x(n)$.

En la DM el predictor empleado es generalmente una suma ponderada de un cierto número de muestras pasadas estimadas.

Específicamente, siendo $x(n - 1)$ la muestra previa, $\hat{x}(n - 1)$ la estimación de ésta, $\hat{x}(n - 2)$ la estimación de dos muestras anteriores etc., la forma general de un predictor lineal puede escribirse:

$$\tilde{x}(n) = \sum_{\ell=1}^k h_{\ell} x(n - \ell) \quad (3.9)$$

Los coeficientes h_{ℓ} son los factores de ponderación

El predictor más simple es el que se usa la estimación de la muestra previa sólo como una estimación de la muestra actual $x(n)$. Para éste caso se tiene:

$$x(n) = h_1 x(n - 1) \quad (3.10)$$

Los coeficientes h_{ℓ} ($\ell = 1, 2, \dots, k$), pueden ser seleccionados para reducir en alguna medida los errores de modo que resulten tan pequeños como sea posible.

Como en el caso de los sistemas PCM, el ruido también se introduce cuando se utiliza DM. A causa de que la señal diferencia es cuantizada en dos niveles, aparece un ruido granular o de cuantización similar al del PCM.

Cuando el rango de muestreo es demasiado bajo o el tamaño de la etapa elegida es también demasiado pequeño para perseguir una señal cambiando rápidamente, ocurre otro tipo de ruido llamado *ruido de sobrecarga*; éste ruido causa un error llamado *error de obediencia*. El cual se adiciona al ruido de cuantización.

Para apreciar la ocurrencia de estos dos tipos de ruido, en la figura 3.7a, observamos la operación de un modulador

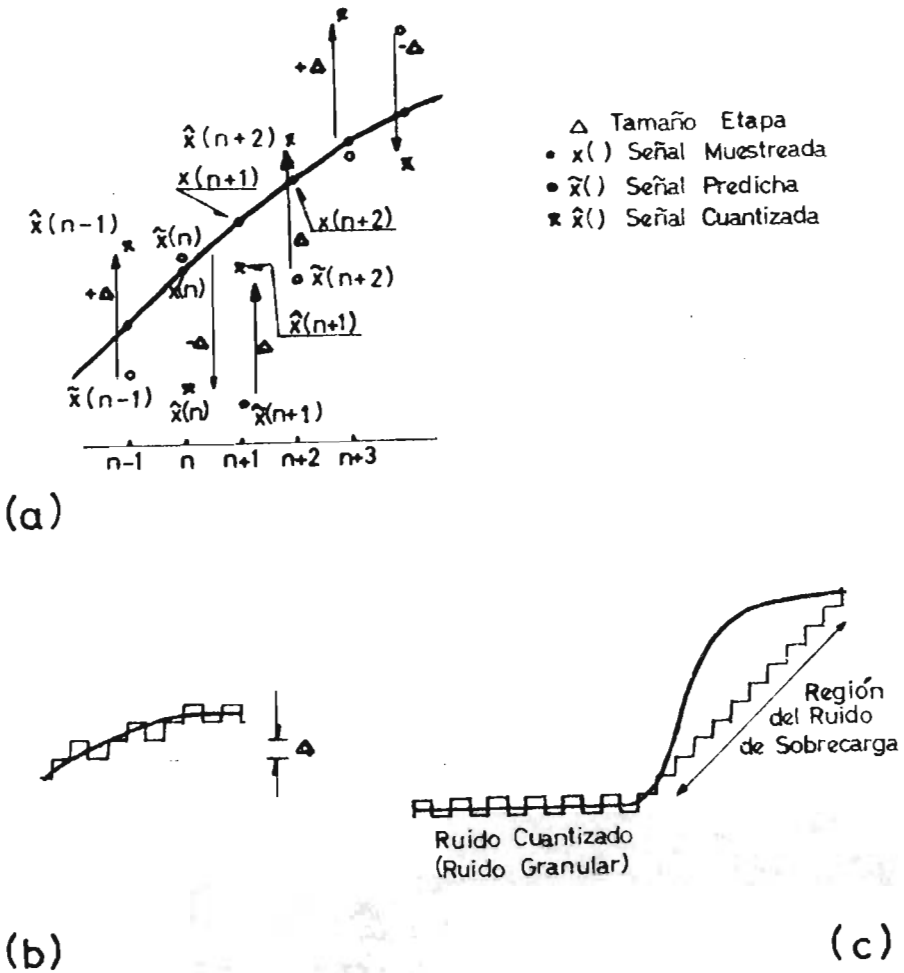


Fig. 3.7.- Funcionamiento de un modulador delta:

- a) Forma gráfica de la operación del modulador delta
- b) Forma de onda reproducida por un DM.
- c) Ruido de cuantización y ruido de sobrecarga.

delta. Debido a que la cuantización es diferencial se usa la estimación de las muestras previas. En el tiempo $(n - 1)$ de la figura el valor predicho $\tilde{x}(n - 1)$ es más pequeño que la señal $x(n - 1)$.

Debido a que :

$$d(n - 1) = x(n - 1) - \tilde{x}(n - 1)$$

es positivo, el tamaño de la etapa Δ es positiva, Esto es adicionado a $\tilde{x}(n - 1)$ para proporcionar la versión cuantizada de la entrada $\hat{x}(n - 1)$. La señal predicha al tiempo de muestreo (n) es entonces:

$$\tilde{x}(n) = \alpha_1 \hat{x}(n - 1)$$

Si asumimos que α_1 es algo menor que uno, se obtiene $\tilde{x}(n)$ como se ve en la figura. Si hacemos ésto para cada punto, encontramos que la versión cuantizada $\hat{x}(n)$ tiende a seguir la curva $x(t)$, algunas veces situada encima y otras por debajo. El resultado es una señal recibida que persigue la señal original, pero introduce ruido cuantizado.

La salida que aparecerá del modulador delta es como en la figura 3.7b, en el cual el ruido es proporcional al tamaño de la etapa Δ . Si se reduce Δ , el ruido también se reduce, pero si es demasiado pequeña, resulta el ruido de sobrecarga, y se debe a la inhabilidad del modulador para perseguir grandes cambios en la señal.

La región de sobrecarga en la figura 3.7c, ocurre porque el tamaño de la etapa fija un límite máximo sobre la pendiente de la señal de entrada que el modulador puede seguir. También Δ determina el error de pico, cuando la pendiente

es muy pequeña. Por tanto, si incrementamos Δ , el ruido de sobrecarga disminuye, pero el ruido cuantizado aumenta, por lo que será deseable una elección óptima de Δ .

Abate, dá la siguiente fórmula empírica para el tamaño de la etapa óptima:

$$\Delta_{\text{opt}} = \left\{ E \left[(x(n) - x(n-1))^2 \right] \right\}^{1/2} \ln(2F_o) \quad (3.11)$$

donde $F_o = F_s/2B$, tal que F_s es el rango de muestreo del modulador y B la frecuencia de Nyquist. También dá una relación señal a ruido en función del tamaño de la etapa variable normalizada:

$$\text{SNR} = \Delta / \left\{ E (x(n) - x(n-1))^2 \right\}^{1/2} \quad (3.12)$$

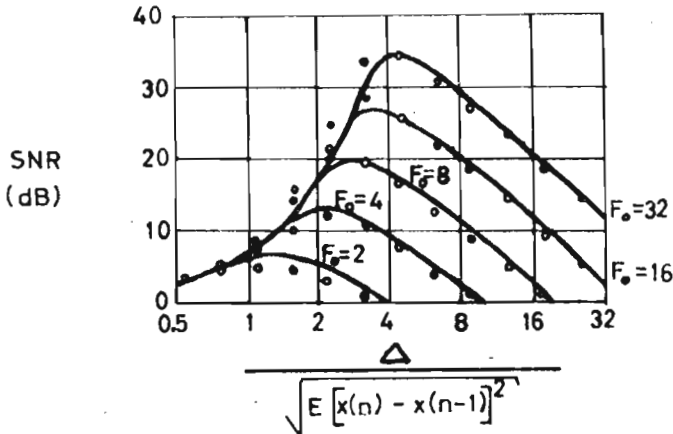


Fig. 3.8.- Gráfica del SNR para DM en función del tamaño de la etapa variable normalizada.

Como se observa en la figura 3.8, hay una localización del pico de la curva SNR para un valor de F_0 . Si F_0 se duplica, el SNR se incrementa a una razón de 9 dB. Al duplicar F_0 duplicaremos la razón de bit debido a que:

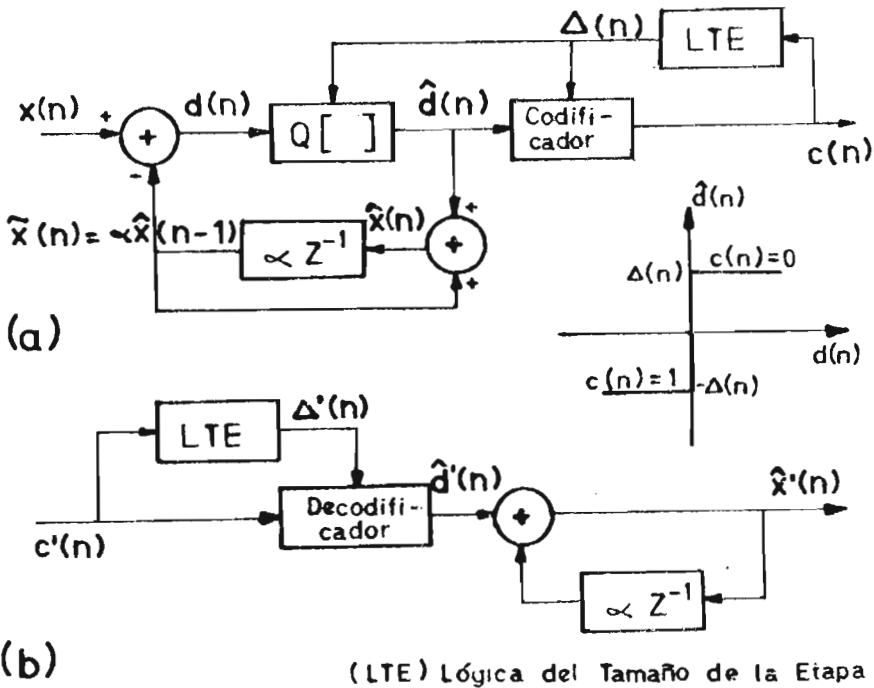
$$\text{RAZON DE BIT} = F_s (1 \text{ bit}) = F_s = 2BF_0 \quad (3.13)$$

Comparandolo con la técnica PCM, tenemos que si duplicamos la razón de bit, duplicando el número de bits muestra, realizaremos un incremento de 6 dB por cada bit adicionado, de esta forma, el SNR incrementado con la razón de bit es mucho más dramático para el PCM que para el DM.

Dado que solamente se utiliza un código de un bit en este sistema, se puede implementar con circuitos integrados simples analógicos y digitales, por lo que no requiere de sincronización de patrones de bit entre el transmisor y el receptor por tanto, la simplicidad de este sistema es su principal ventaja. Las limitaciones sobre el funcionamiento del sistema LDM, proviene principalmente de su cuantización tosca de la diferencia de la señal. Por lo que la cuantización adaptiva mejora grandemente el funcionamiento del modulador delta.

3.4.- MODULACION DELTA ADAPTIVA (ADM)

Una forma general de un sistema de Modulación Delta Adaptiva (ADM), es como el mostrado en la figura 3.9. Como en la figura, la mayoría de estos esquemas son del tipo de cuantización adaptiva regenerativa o de realimentación, en el cual el tamaño de la etapa para un cuantizador de dos niveles es adaptado de la salida de la palabra código. En este sistema, se mantiene la ventaja de que no se requiere de la sincronización de patrones de bit.



(LTE) Lógica del Tamaño de la Etapa

Fig. 3.9.- Diagrama a bloques de la Modulación Delta Adaptiva: a) Codificador b) Decodificador

En el algoritmo ADM, el tamaño de la etapa obedece a la regla:

$$\Delta(n) = M \Delta(n - 1) \quad (3.14a)$$

$$\Delta_{\min} \leq \Delta(n) \leq \Delta_{\max} \quad (3.14b)$$

En el cual el multiplicador es una función de las palabras código presentes y previas, $c(n)$ y $c(n - 1)$. Esto es posible debido a que $c(n)$ depende solamente del signo de $d(n)$, el cual está dado por:

$$d(n) = x(n) - \infty \hat{x}(n - 1) \quad (3.15)$$

De este modo el signo de $d(n)$ puede ser determinado antes del proceso de la determinación del valor cuantizado actual $\hat{d}(n)$, el cual debe esperar la determinación de $\Delta(n)$ de la Ec. (3.14). El criterio para elegir el multiplicador del tamaño de la etapa de la Ec. (3.14) es:

$$\begin{aligned} M = P > 1 & \quad \text{si} \quad c(n) = c(n - 1) \\ M = Q < 1 & \quad \text{si} \quad c(n) \neq c(n - 1) \end{aligned} \quad (3.16)$$

En la figura 3.10a, se muestra una señal analógica procesada por medio de LDM, y en la figura 3.10b, la misma señal pero procesada ahora pro ADM, donde:

- $x(n)$ es la señal de entrada
- o $\tilde{x}(n)$ es la señal predicha
- x $\hat{x}(n)$ es la señal cuantizada

Con esta figura podemos apreciar con gran facilidad el resultado de ambos procedimientos, y la diferencia que provocan en la señal original. La región de la izquierda es la de sobrecarga (que es más pronunciada en LDM que en ADM), y a la derecha la región de granularidad.

Delas Ecs. (3.14) y (3.16) vemos que los parámetros a controlar del sistema ADM son: P , Q , $\Delta_{\text{mín}}$, $\Delta_{\text{máx}}$. El tamaño mínimo de la etapa puede mantenerse tanto como la pendiente de la forma de onda se mantenga, o bien hacerla más pequeña que la requerida para el funcionamiento óptimo de un LDM, con lo que el ruido granular puede ser grandemente reducido. Igualmente el tamaño de la etapa máxima puede ser

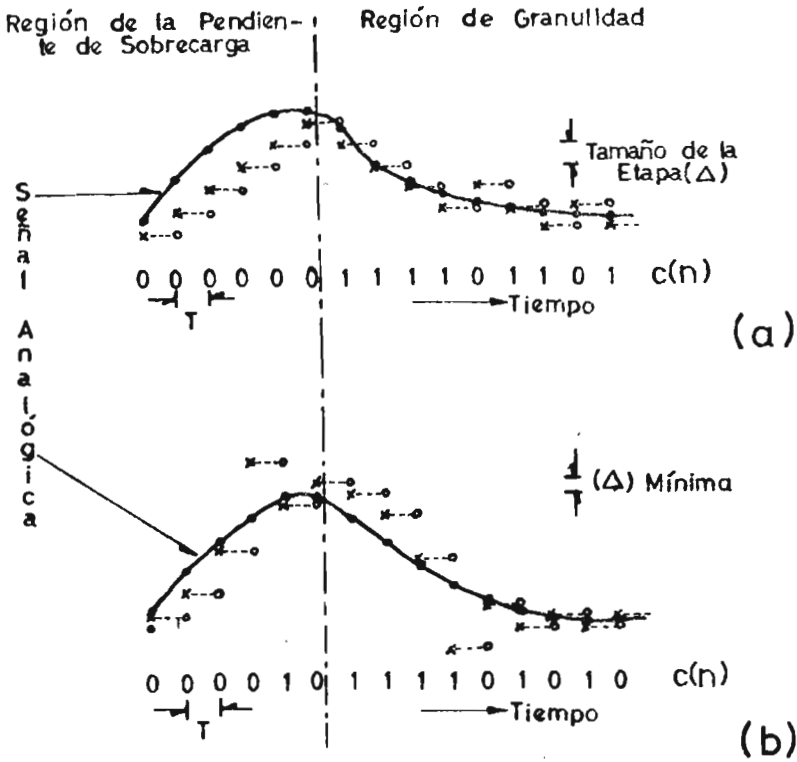


Fig. 3.10.- Señal analógica procesada mediante Modulación Delta:

a) Tamaño de la etapa fija (LDM)

b) Tamaño de la etapa adaptiva (ADM)

hecha más larga que la pendiente máxima de la señal de entrada para reducir la pendiente del ruido de sobrecarga.

Los límites del tamaño de la etapa, serán elegidos para proporcionar el rango dinámico deseado para la señal de entrada. La relación $\Delta_{\text{máx}} / \Delta_{\text{mín}}$ (que controla el rango dinámico del cuantizador), será lo bastante grande para mantener un SNR alto sobre un rango deseado de los niveles de la señal de entrada. El tamaño mínimo de la etapa será tan pequeño como sea práctico para minimizar el ruido que se produce en la condición de canal desconectado. (6)

3.5.- COMPARACION DE LOS SISTEMAS PCM LOGARITMICO, LDM Y ADM

En la figura 3.11, se comparan sistemas LDM, PCM log ($M = 100$) y ADM ($P = 1.5$), con un rango de muestreo de Nyquist ($2B = 6.6$ KHz). En la figura se muestra que el sistema ADM es superior al LDM por 8 dB a 20 Kbps y la ventaja de SNR aumenta a 16 dB a 60 Kbps. Para el LDM, se observa que duplicando el rango de muestreo y la razón de bit, incrementa 6 dB mientras el ADM el correspondiente incremento es de 10 dB.

Comparando el sistema ADM y el PCM logarítmico, notamos que el ADM excede el funcionamiento del PCM log para una razón de bit por debajo de los 40 Kbps. Para razones de bit más alto el PCM log tiene un SNR más alto, pues como se aprecia, un ADM requiere 60 Kbps para llevar a cabo la misma calidad que con un PCM log de 7 bits que la realiza con una razón de bit de alrededor de 46 Kbps.

Para muchas aplicaciones ADM puede ser preferida al PCM log, inclusive a expensas de una razón de información ligera

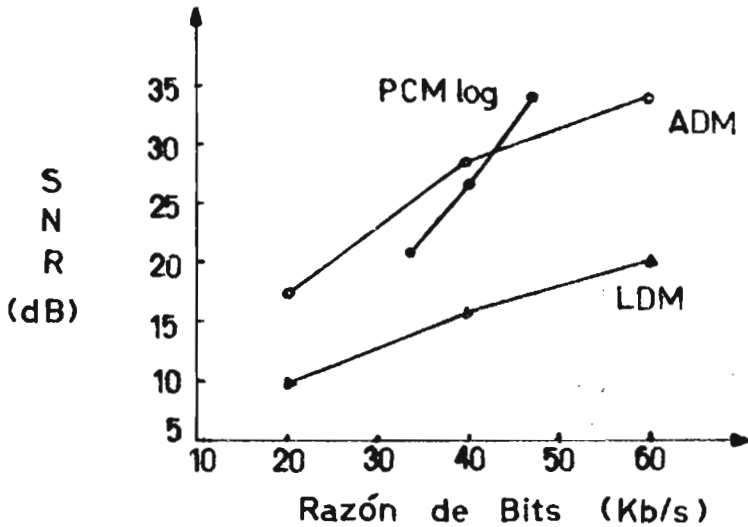
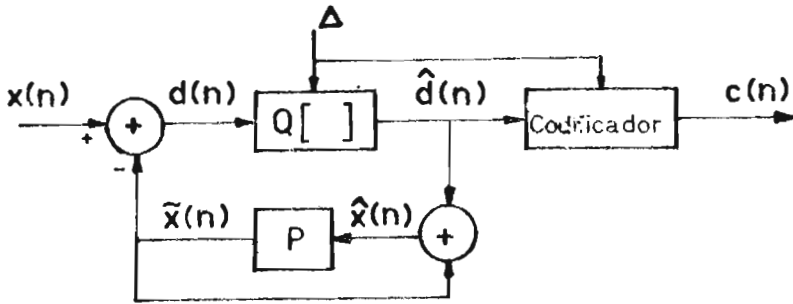


Fig. 3.11.- Comparación gráfica SNR (dB) contra razón de bit (Kbps) para los sistemas PCM logarítmico, LDM y ADM, con un rango de muestreo de 6.6 KHz.

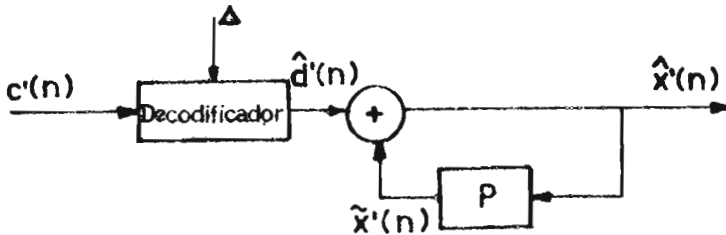
mente más alta. Debido a que el ADM mantiene la simplicidad del LDM, ya que la calidad mejorada del ADM se realiza solamente con un ligero incremento en complejidad, debido a que la adaptación del tamaño de la etapa se realiza usando la salida del bit corriente o actual.

3.6.- MODULACION POR CODIFICACION DE PULSOS DIFERENCIAL (DPCM)

Cualquier sistema que contenga un esquema de cuantización como el mostrado en la figura 3.12, será llamado un sistema PCM Diferencial (DPCM). El DPCM, en su forma más simple usa la muestra actual de voz como una predicción para la si-



(a)



(b)

Fig. 3.12.- Diagrama a bloques del esquema de cuantización diferencial:

a) Codificador

b) Decodificador

guiente y almacena el error de predicción es decir, la diferencia muestra a muestra.

La razón de datos puede ser reducida transmitiendo la diferencia entre muestras sucesivas en lugar de sus valores absolutos, por lo que, se requieren menos bits para la diferencia de la señal con cierta exactitud, debido a que no se consideran valores extremos. (16)

El sistema DPCM con predictor fijo, proporciona una mejora en el SNR de 4 a 11 dB sobre un PCM, debido a que se cuantiza la diferencia muestra a muestra de la señal de voz

$d(n)$. Es decir:

$$d(n) = x(n) - \tilde{x}(n) \quad (3.17)$$

donde $x(n)$ es la muestra de entrada no cuantizada y $\tilde{x}(n)$ es el valor predicho o estimado de la muestra de entrada, y es la salida de un sistema predictor P . La diferencia de la señal también puede ser llamada error de predicción de la señal, ya que esta diferencia es la cantidad por la cual el predictor fallará para predecir exactamente la entrada.

También se puede hacer uso de la predicción adaptiva cuando se da el caso de correlaciones dinámicas en la voz (cuando la estructura de "formant" varía con el tiempo). El predictor adaptivo tiende a trasladar redundancias de la señal de voz, por lo tanto, es inherentemente menos sensitivo que un predictor fijo.

Si la predicción está basada sobre las 4 a 20 muestras pasadas de voz, nos referimos a una predicción de tiempo corto. Por otro lado, si adicionalmente tenemos una correlación de periodo de tono (pitch) a periodo de tono, entonces tenemos una predicción de tiempo largo, y ésto puede ser utilizado para reducir la diferencia en una forma dinámica.

De aquí, obtenemos la técnica de Codificación por Predicción Adaptiva (APC), que es mostrada en la figura 3.13. Para que el algoritmo APC tome ventaja del rango dinámico espectral local y de las propiedades de enmascaramiento auditivo en la percepción de la voz, en la figura hay una tercera malla de realimentación alrededor del cuantizador, el cual proporciona realimentación de ruido cuantizado.

El APC, es un algoritmo de una complejidad relativamen-

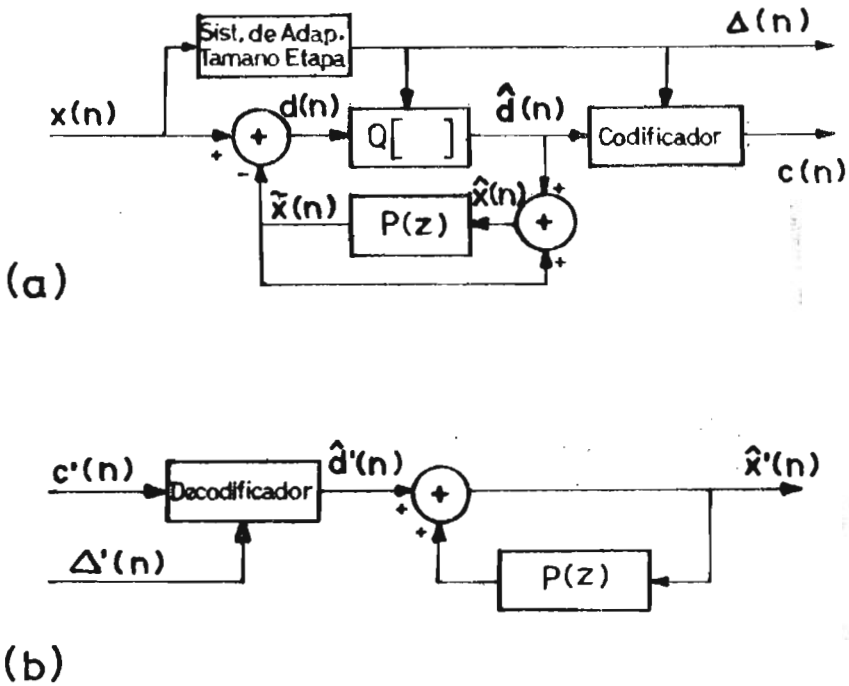


Fig. 3.14.- Diagrama a bloques del sistema ADPCM;
a) Codificador b) Decodificador

ADPCM, y en donde se aprecia que hace uso de la cuantización adaptiva básicamente, sin embargo, también es posible hacer uso de la predicción adaptiva, con lo que el funcionamiento mejora grandemente, debido a que el predictor adaptivo examina constantemente el flujo de la muestra actual y predice el siguiente valor de la muestra.

Esta predicción es altamente precisa, ya que el predictor en cierta forma, simula la voz humana y continuamente se ajusta a si mismo internamente para minimizar el error, haciendo consistente el funcionamiento, y las predicciones son

altamente seguras proque la diferencia entre las muestras actuales y las predichas $d(n)$, es siempre muy pequeña. Debido a ésto, el sistema ADPCM, mantiene una baja razón de bit, ya que varia dinámicamente la cuantización entre muestras, dependiendo sobre su razón de cambio.

Las variaciones en el valor del tamaño de la etapa de cuantización se regulan estimando las características complejas de las ondas senoidales que ocurren en la voz (Es por ésto, que la técnica no es aplicable a señales como la onda cuadrada).

Dado que esta técnica es la base principal de este trabajo, en el capítulo siguiente, se explicará con más detalle los conceptos antes mencionados.

3.8.- MODULACION DELTA CON PENDIENTE QUE VARIA CONTINUAMENTE (CVSD)

La Modulación Delta con Pendiente que Varia Continuatamente (CVSD), es otro tipo de cuantización adaptiva en DM y éste también se basa en un esquema como el mostrado en la figura 3.12, pero el tamaño de la etapa está definido por las ecuaciones:

$$\Delta(n) = \beta \Delta(n-1) + D_2 \quad (3.18a)$$

si $c(n) = c(n-1) = c(n-2)$

$$\Delta(n) = \beta \Delta(n-1) + D_1 \quad (3.18b)$$

de cualquier otra forma

donde $0 < \beta < 1$ y $D_2 \gg D_1 > 0$

El tamaño de la etapa máximo se obtiene si la entrada al filtro del tamaño de la etapa es constante a D_2 como ocurre en un periodo prolongado de la pendiente de sobrecarga. El tamaño de la etapa mínimo, se obtendrá si el arreglo $c(n) = c(n - 1) = c(n - 2)$, no ocurre para periodo prolongado como en la condición de canal desconectado.

El principio básico de funcionamiento del CVSD es como el del ADM, donde se incrementa el tamaño de la etapa en respuesta al patrón del bit actual, que indica la pendiente de sobrecarga.

En el sistema CVSD una carrera de tres unos consecutivos o de tres ceros consecutivos causan un incremento de D_2 para ser adicionado al tamaño de la etapa. En ausencia de estas condiciones, el tamaño de la etapa decae (debido a que $\beta < 1$), hasta alcanzar $\Delta_{\text{mín}}$. Al igual que ADM, $\Delta_{\text{mín}}$ y $\Delta_{\text{máx}}$ pueden ser elegidos para proporcionar el rango dinámico deseado y un ruido granular bajo durante la condición de canal desconectado.

El parámetro β controla la velocidad de adaptación. Si β está cercano a la unidad, será la razón de refuerzo y $\Delta(n)$ decae en forma lenta, mientras que si β es mucho menor que la unidad, la adaptación es más rápida. Este esquema de adaptación, puede ser ajustado para cualquiera de las dos formas siguientes:

- ADAPTACION SILABICA
- ADAPTACION INSTANTANEA

(Ver capítulo siguiente inciso 4.1 para mayor detalle).

Generalmente, este sistema ha sido utilizado en situa-

ciones donde se requiere baja sensibilidad para errores de canal, con necesidades de calidad de voz por debajo de los requisitos para canales de comunicación comercial. En esta situación, el coeficiente del predictor está fijado a un valor considerablemente menor a la unidad, y los parámetros del sistema son ajustados para proporcionar adaptación silábica, debido a que el efecto de los errores del canal desvirtuan la calidad.

El precio que paga un sistema CVSD por su insensitividad a errores, naturalmente es que la calidad decrece cuando no ocurren errores. Aun embargo, su mayor ventaja es que tiene la suficiente flexibilidad para proporcionar un compromiso (trade-off) efectivo entre la calidad y la robustez.

Todas las técnicas de codificación de forma de onda en el dominio del tiempo, son las más ampliamente utilizadas, y están englobadas en lo anteriormente explicado, sin embargo, también hay codificadores de forma de onda en el dominio de la frecuencia.

3.9.- CODIFICACION EN SUB-BANDAS (SBC)

Una de las formas más simples de codificación de la forma de onda en el dominio de la frecuencia es la Codificación en Sub-Bandas (SBC), la cual se muestra en el diagrama a bloques en la figura 3.15.

En la aproximación SBC, la señal de voz se subdivide dentro de un conjunto de componentes separados en el dominio de la frecuencia y estos componentes se codifican independientemente. La banda de voz se divide de 4 a 8 sub-bandas, por medio de un banco de filtros, las señales resultantes de estos filtros se codifican usando métodos de cuantización

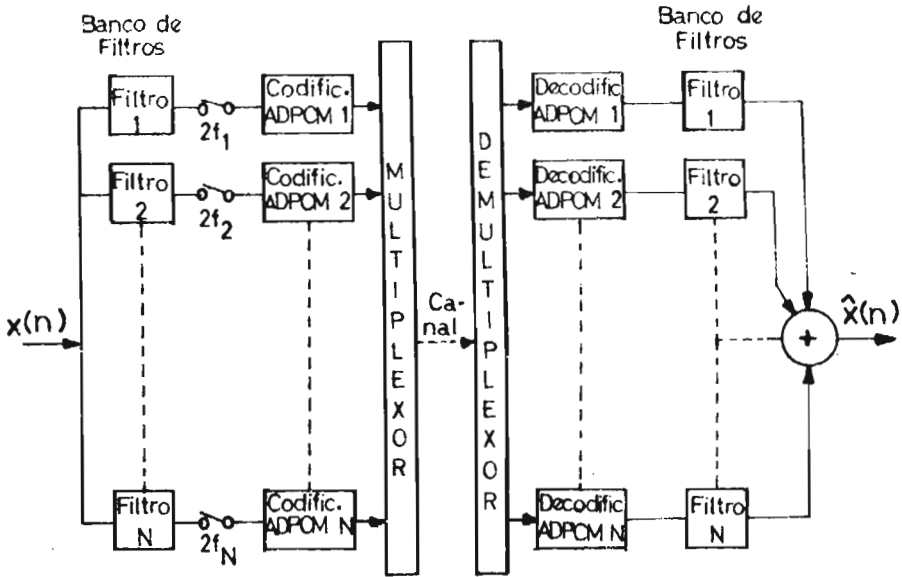


Fig. 3.15.- Diagrama a bloques del sistema de Codificación en Sub-Bandas (SBC).

adaptiva (codificadores ADPCM).

Debido al tipo de codificador usado, el tamaño de la etapa varia independientemente, las bandas con baja energía usan etapas de tamaño pequeño, produciendo ruido cuantizado menor que con sistemas menos flexibles. Por tanto, el ruido de una banda no afecta otras bandas de frecuencia.

Se debe elegir cuidadosamente el número de bits/muestra a utilizarse para codificar cada banda. Después de bajar el rango de muestreo en cada banda, se puede obtener una razón de bit total más pequeña, mientras se mantiene la calidad de la voz, reduciendo el número de bits/muestra en bandas de

alta frecuencia que perceptualmente son menos importantes.

Esta técnica puede adoptar dinámicamente las variaciones espectrales ordinarias (estructura de "formant") en la voz, y proporciona condición de ruido "noise shaping", tomando ventaja de las propiedades de la percepción de la voz (tabla 2.1), y de la insensibilidad de fase del oído.

Si se incrementa el número de las bandas de frecuencia y además se permite que el número de bits/muestra varíe dinámicamente de una banda a otra, obtenemos el método de Codificación Adaptiva por Transformada (ATC).

La aproximación ATC, se muestra por medio de un diagrama a bloques en la figura 3.16. Como se puede apreciar en la figura, a diferencia del sistema SBC, en el sistema ATC las muestras de voz son transformadas a la frecuencia por medio de un banco de filtros y lo realizan efectivamente por medio de un algoritmo de transformada rápida discreta, y entonces los coeficientes de transformación son codificados como un bloque.

En la síntesis se utiliza la información lateral, que afecta las características del cuantizador para transformar los coeficientes retrasados en una señal de voz.

Además, esta aproximación es capaz de adoptar dinámicamente y en forma completa las propiedades espectrales de la voz, incluyendo la estructura de "formant", la estructura de tono "pitch", y controla la condición de ruido "noise shaping" por medio de la colocación del bit de asignación o de signación en el algoritmo.

Tanto la aproximación SBC como la ATC proporcionan una buena calidad de voz a una razón de bit de 16 Kbps.

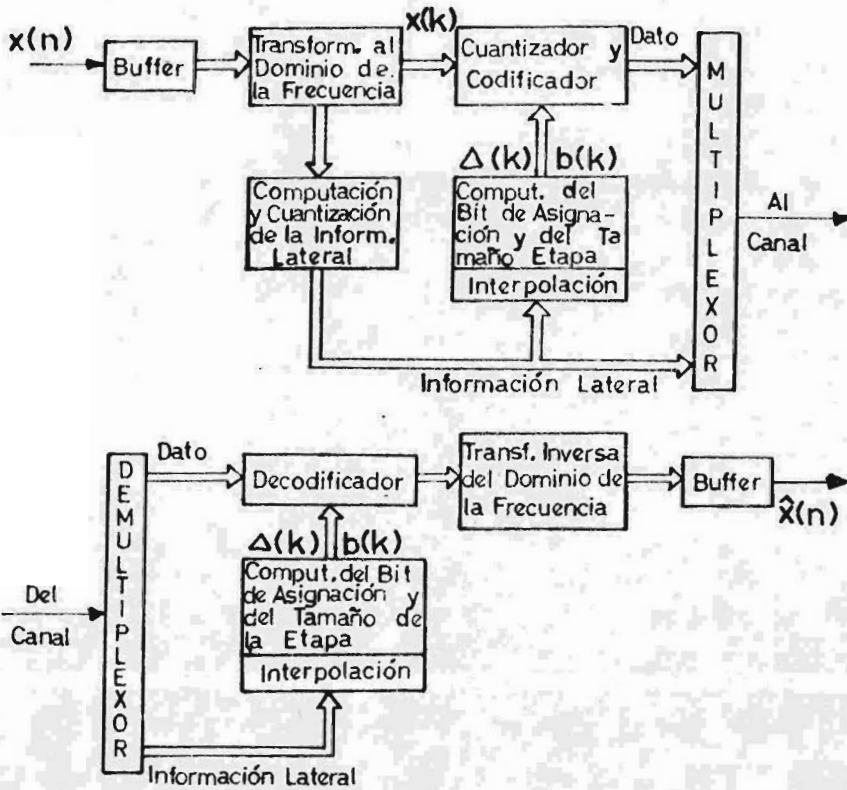


Fig. 5.16.- Diagrama a bloques del sistema de Codificación Adaptiva por Transformada (ATC)

5.10.- RESUMEN

Se ha desarrollado ampliamente la investigación de codificación de la forma de onda, enfocada en diversas áreas. Técnicas en el dominio del tiempo para razones de bit de 16 a 32 Kbps de mediana y baja complejidad han sido extensamente estudiadas y refinadas para máximo funcionamiento y robustez. Aproximaciones para razones de bit de 16 Kbps y menores con alta complejidad no se encuentran del todo bien refinadas, sin embargo, las variaciones básicas de estas téc

nicas están todavía siendo exploradas.

Igualmente, en el dominio de la frecuencia se ha desarrollado la investigación para mejoras principalmente en técnicas de codificación de transformación adaptiva y las técnicas relacionadas.

La elección de la técnica a utilizarse dependerá de la calidad deseada, la razón de datos a manejar, la complejidad o costo requerido, pero sobre todo de la aplicación del sistema.

C
A
P
I
T
U
L
O

4.- TECNICA DE COMPRESION DE SEÑALES MEDIANTE ADPCM

Una de las técnicas desarrolladas ampliamente en la última década es la Modulación por Codificación de Pulsos Diferencial Adaptivo (ADPCM).

Debido a que en los sistemas de comunicaciones la técnica más ampliamente utilizada es la PCM, el ADPCM adquiere gran importancia ya que, esta técnica es una forma especializada del PCM, sobreponiendo los defectos de la técnica DM y mejorando la DPCM, ofreciendo una calidad y una inteligibilidad significativamente alta a una razón de bits más bajo.

Por ejemplo, un PCM de 12 bits puede ser condensado en uno de 3 o 4 bits ADPCM, por lo que el canal de comunicaciones puede ser por lo menos doblemente usado.

Para comprender la aproximación ADPCM más fácilmente, se explicará con algo de detalle la cuantización adaptiva y la predictiva.

4.1.- CUANTIZACION ADAPTIVA

La amplitud de la señal de voz puede variar sobre una gran cantidad de valores, dependiendo del orador, de una pronunciación dada, de los segmentos de sonidos sonoros a sonidos insonoros y del medio ambiente de la comunicación.

Para satisfacer todas las fluctuaciones, será conveniente elegir el tamaño de la etapa lo más grande posible para acomodar el rango máximo pico a pico de la señal y por otro lado, hacer la etapa de cuantización pequeña para minimizar el ruido cuantizado. Por lo anterior, se puede utilizar un

cuantizador no uniforme, pero una alternativa es adoptar las propiedades del cuantizador al nivel de la señal de entrada.

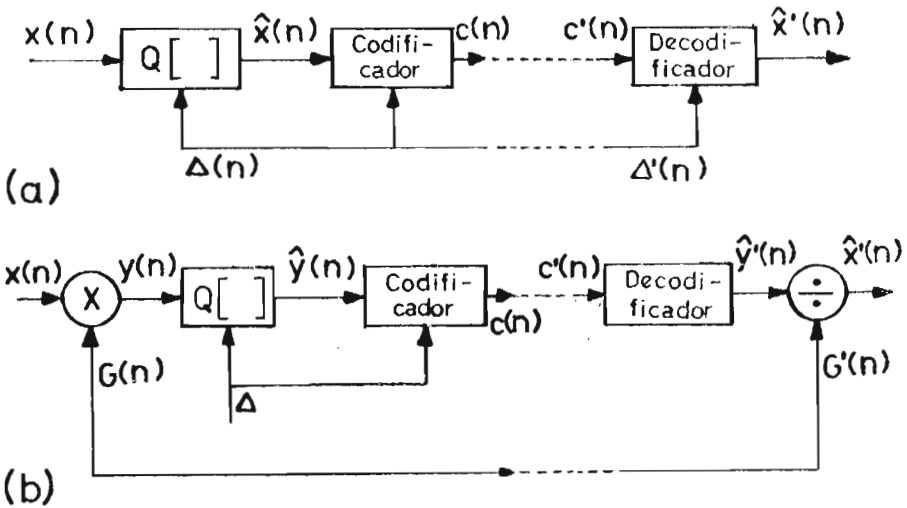


Fig. 4.1.- Representación en diagrama a bloques de la cuantización adaptativa:

- a) Representación con el tamaño de la etapa variable;
- b) Representación con la ganancia variable.

Para lograr la cuantización adaptativa, se debe permitir que el tamaño de la etapa $\Delta(n)$ o en forma más general que los niveles y rangos del cuantizador varíe para igualar la variancia de la señal de entrada, lo que se describe en la figura 4.1a.

Una forma alternativa es considerar fija las características del cuantizador, haciendo variar la ganancia $G(n)$ con el tiempo, con lo que se conservará la variancia constante,

ésto se describe esquemáticamente en la figura 4.1b.

En el primer caso, el tamaño de la etapa se incrementará y decrementará cuando la varianza de la señal de entrada se incremente o decremente; si en este caso tenemos un cuantizador no uniforme, los niveles y rangos de cuantización serán escalados linealmente para igualar la varianza de la señal. En el segundo caso, la modificación se aplica afuera, tanto para cuantizadores uniformes como no uniformes, donde para conservar la varianza de la señal de entrada al cuantizador relativamente constante, la ganancia cambia inversamente con los cambios en la varianza. En cada caso, es necesario obtener un estimado de las propiedades de amplitud variando con el tiempo de la señal de entrada.

Hasta la fecha han evolucionado dos clase de esquemas de cuantización adaptiva:

- CUANTIZADORES ADAPTIVOS DE ALIMENTACION ADELANTADA.
- CUANTIZADORES ADAPTIVOS REGENERATIVOS.

En la primera clase la amplitud o varianza de entrada es estimada de la entrada en sí misma. En la segunda clase, el tamaño de la etapa se adapta sobre las bases de la salida del cuantizador $\hat{x}(n)$, o en forma equivalente sobre la salida de las palabras código $c(n)$.

El tiempo de adaptación o la escala del tiempo sobre el cual el cambio tiene lugar, en cualquiera de las dos clases de cuantizadores adaptivos, puede ser:

- INSTANTANEA, si la adaptación es rápida.
- SILABICA, si la adaptación es lenta.

Cuando nos referimos a cambios instantáneos, nos estamos enfocando a cambios de amplitud muestra a muestra o a cambios rápidos dentro de pocas muestras.

Si en un intervalo sonoro o en un intervalo insonoro la amplitud pico se mantiene esencialmente inalterado para intervalos relativamente largos de tiempo, por su variación lenta es llamada silábica, esto implica que la variación ocurre con el rango comparable del rango silábico cuando se habla.

De aquí, los esquemas de cuantización adaptiva se clasifican de acuerdo que tan rápido se adopten a las variaciones de la amplitud de la voz.

4.1.1.- CUANTIZACION ADAPTIVA DE ALIMENTACION ADELANTADA

Una representación general del cuantizador adaptivo de alimentación adelantada se muestra en la figura 4.2. Donde se observa que el parámetro a variar es el tamaño de la etapa $\Delta(n)$, mientras que por conveniencia en el análisis el cuantizador es uniforme, sin embargo, se puede generalizar para el cuantizador no uniforme.

En la figura 4.3, se muestra un cuantizador adaptivo de alimentación adelantada representado en términos de la ganancia variando con el tiempo, donde las muestras cuantizadas son representadas por las palabras código $c(n)$ y la ganancia $G(n)$ juntas.

Por otro lado, cuando varía el tamaño de la etapa como en la figura 4.2, la muestra $x(n)$ se representa por la palabra código $c(n)$ y el tamaño de la etapa $\Delta(n)$.

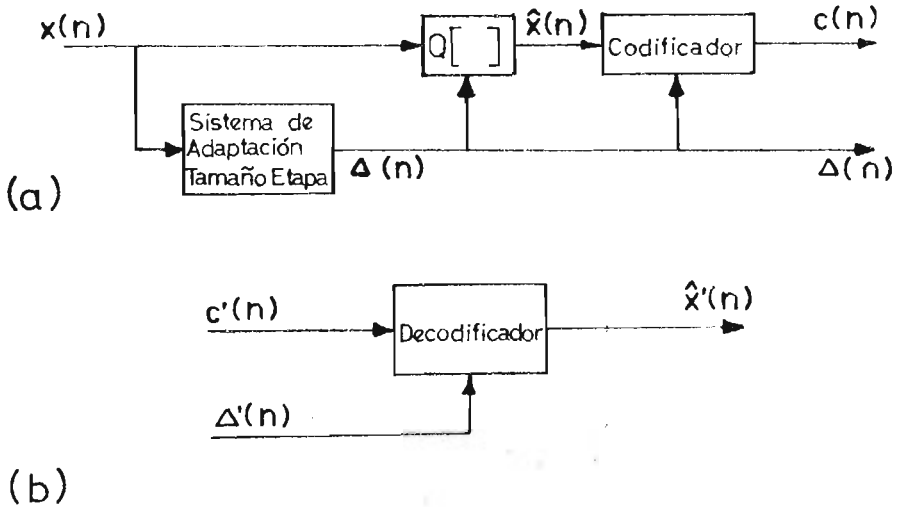


Fig. 4.2.- Representación general en diagrama a bloques del cuantizador de alimentación adelantada:

a) Codificador

b) Decodificador

Ahora si; $c'(n) = c(n)$ y $\Delta'(n) = \Delta(n)$
entonces $\hat{x}'(n) = \hat{x}(n)$

lo que indica que no existieron errores en la transmisión pero en el caso que los hubiera;

$c'(n) \neq c(n)$ y $\Delta'(n) \neq \Delta(n)$
entonces $\hat{x}'(n) \neq \hat{x}(n)$

Los esquemas de alimentación adelantada obtienen un estimado de la varianza en función del tiempo, por tanto, el tamaño de la etapa o niveles de cuantización se hacen proporcionales a la desviación estandar, mientras que con la ganancia aplicada a la entrada se hace inversamente propor

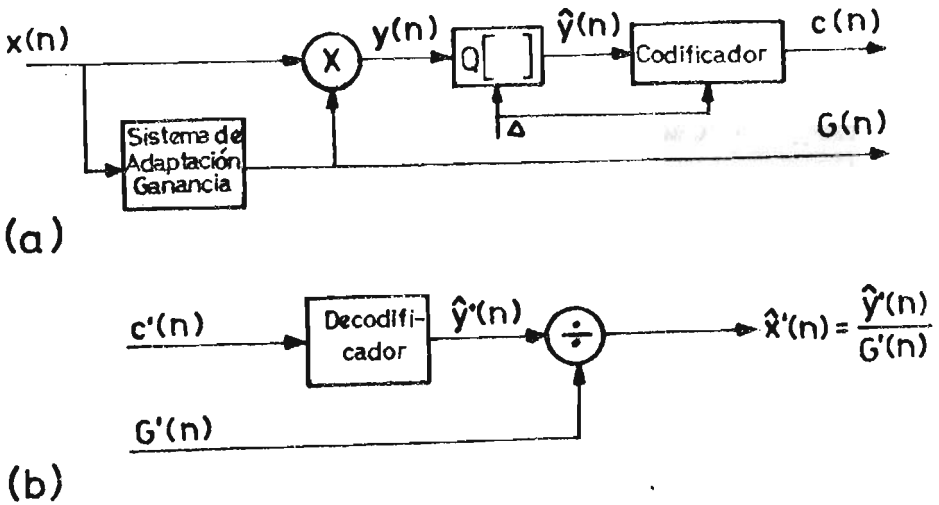


Fig. 4.3.- Representación general en diagrama a bloques del cuantizador adaptivo de alimentación adelantada con la ganancia en función del tiempo.

a) Codificador

b) Decodificador

cional a la desviación estandar.

Por tanto, el tamaño de la etapa en la figura 4.2 será de la forma:

$$\Delta(n) = \Delta_0 \sigma(n) \quad (4.1)$$

y la ganancia en función del tiempo en la figura 4.3, será de la forma:

$$G(n) = \frac{G_0}{\sigma(n)} \quad (4.2)$$

donde Δ_o y G_o describen la ganancia del filtro, ésto es debido a que la varianza es proporcional a la salida de un filtro pasobajas con entrada $x^2(n)$

El tamaño de la etapa $\Delta(n)$, o en su caso la ganancia $G(n)$ deben ser muestreadas y cuantizadas antes de la transmisión en las figuras 4.2 y 4.3. Debido a que el rango de información de la representación digital de la señal, es la suma del rango de información de la salida del cuantizador y el rango de información de la ganancia, es importante considerar un rango de muestreo lo más bajo posible para la ganancia o para el tamaño de la etapa.

A causa de las restricciones físicas de implementación y para permitir la cuantización de la ganancia o bien del tamaño de la etapa se limitará su variación, por lo que:

$$G_{\text{mín}} \leq G(n) \leq G_{\text{máx}} \quad (4.3)$$

$$\Delta_{\text{mín}} \leq \Delta(n) \leq \Delta_{\text{máx}} \quad (4.4)$$

Así, el rango dinámico del sistema lo determinan la relación de estos límites, de este modo, para obtener un SNR relativamente constante sobre los 40 dB, se requiere que:

$$G_{\text{máx}}/G_{\text{mín}} = 100 \quad \text{y} \quad \Delta_{\text{máx}}/\Delta_{\text{mín}} = 100$$

En un estudio comparativo realizado por Noll, hace evidente que la mejoría en el factor SNR que puede obtenerse con la cuantización adaptiva, se muestra en la tabla 4.1, donde para su estudio consideró un esquema de alimentación adelantada, en el cual la varianza estimada fué:

$$\sigma^2(n) = \frac{1}{M} \sum_{m=n}^{n+M-1} x^2(m) \quad (4.5)$$

Por lo que el tamaño de la etapa $\Delta(n)$ o la ganancia $G(n)$ fueron evaluados cada M muestras. Entonces el sistema necesitó de un buffer de M muestras que permite determinar el tamaño de la etapa del cuantizador o la ganancia en función de los términos de las muestras que van a ser cuantizadas.

TABLA 4.1.- CUANTIZACION ADAPTIVA DE 3 BITS CON ADAPTACION DE ALIMENTACION ADELANTADA

CUANTIZADORES NO UNIFORMES	NO ADAPTIVA SNR (dB)	ADAPTIVA M= 128 SNR (dB)	ADAPTIVA M = 1024 SNR (dB)
Gaussiana	7.3	15.0	12.1
Laplaciana	9.9	15.3	12.8
CUANTIZADORES UNIFORMES			
Gaussiana	6.7	14.7	11.5
Laplaciana	7.4	15.4	11.5

Como se puede ver en la tabla, el SNR de un cuantizador adaptivo es mejor por 8 dB que los cuantizadores no uniformes fijos. Es por ésto que los cuantizadores adaptivos tienen una ventaja evidente sobre el resto de los cuantizadores. Sin embargo, hay una ventaja adicional que no muestran los números de la tabla anterior. Esto es, si elegimos apropiadamente $\Delta_{\text{mín}}$ y $\Delta_{\text{máx}}$ (el SNR de la segunda columna, la adaptación del tamaño de la etapa está basada sobre la varianza estimada de la Ec. (4.5) es posible llevar a cabo una mejora en el factor SNR, mientras se mantiene un bajo ruido para la condición de canal desconectado y un amplio rango di

námico, todo ésto se da en sistemas bien diseñados. La com
binación de todos éstos factores hacen de la cuantización
adaptiva una alternativa atractiva.

4.1.2.- CUANTIZACION ADAPTIVA REGENERATIVA

Los cuantizadores adaptivos regenerativos o de realimentación
son mostrados en la figura 4.4.

La varianza de la entrada es estimada desde las pala-
bras código $c(n)$. Como en los sistemas de alimentación ade-
lantada el tamaño de la etapa y la ganancia son proporci-
onal e inversamente proporcional respectivamente a un esti-
mado de la desviación estandar de la entrada como en las
ecuaciones 4.1 y 4.2. La varianza para éste caso será:

$$\sigma^2(n) = \sum_{m=-\infty}^{\infty} \hat{x}^2(m) h(n-m) \quad (4.6)$$

donde $h(n)$ es la respuesta al impulso del filtro pasobajas.

Sin embargo, aquí no será posible utilizar un buffer
para implementar un filtro no causal, debido a que la varian-
za estimada estará basada solamente sobre valores pasados de
 $\hat{x}(n)$ puesto que el valor presente no estará disponible hasta
que la cuantización haya ocurrido y ésta estará después que
la varianza ha sido estimada.

Una aproximación basada en la figura 4.4 a y b, fue es-
tudiada por Jayant, en donde el tamaño de la etapa de un cuan-
tizador uniforme es adaptado a cada tiempo de muestra por
la relación:

$$\Delta(n) = P \Delta(n-1) \quad (4.7)$$

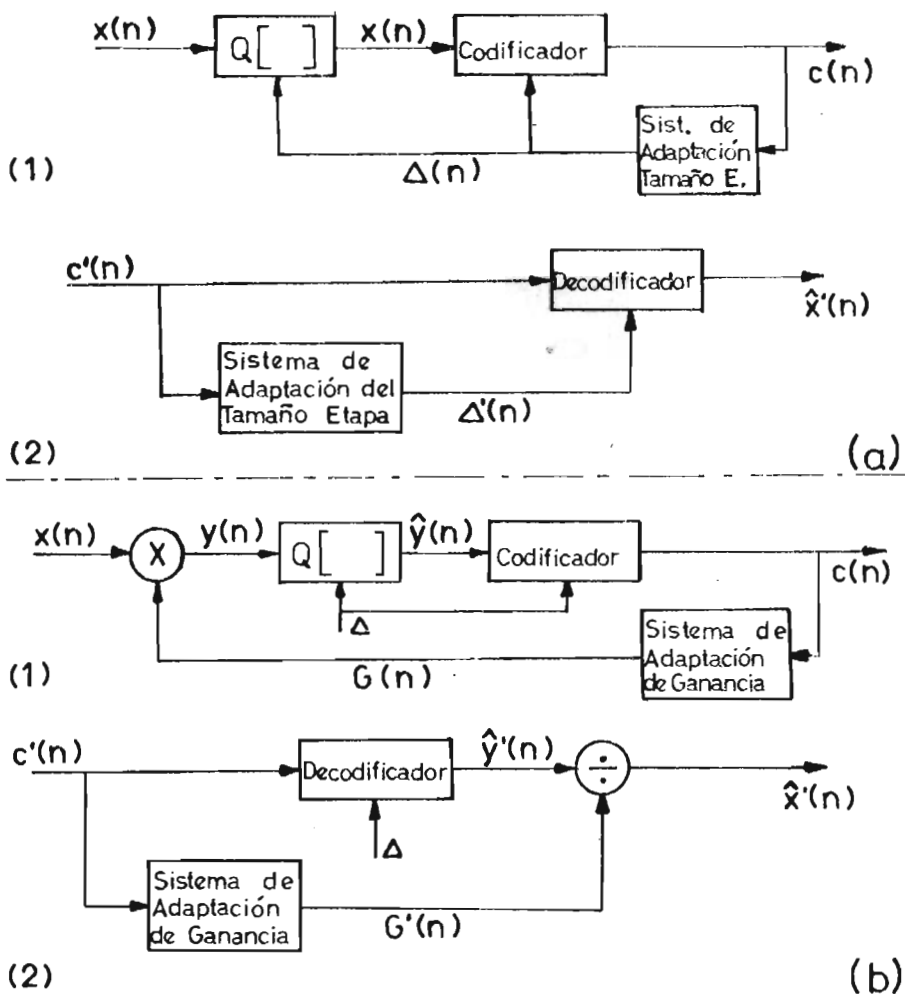


Fig. 4.4.- Representación general en diagrama a bloques del cuantizador adaptivo regenerativo:

a) Adaptando la etapa

1) Codificador

2) Decodificador

b) Adaptando la ganancia

1) Codificador

2) Decodificador

donde el multiplicador del tamaño de la etapa P, es una función solamente de la magnitud de las palabras código previas $c(n - 1)$. Eligiendo las palabras código como la de la figura 3.3a, donde el bit más significativo es el bit signo y el resto de la palabra es la magnitud, dando la ecuación 3.2, que por conveniencia la escribiremos de nuevo:

$$\hat{x}(n) = \frac{\Delta(n) \text{sing}\{c(n)\}}{2} + \Delta(n) c(n) \quad (4.8)$$

donde $\Delta(n)$ satisface la Ec. (4.7).

Para representar la señal, sólo se requiere de la secuencia de las palabras código ya que el tamaño de la etapa actual $\Delta(n)$ depende del tamaño de la etapa y de las palabras código previas.

La forma en la cual el multiplicador P en la Ec. (4.7), variará con $c(n - 1)$ será de la siguiente manera; Si la palabra código previa corresponde a cualquiera de los dos niveles del cuantizador más largo positivo o más largo negativo, entonces el cuantizador está sobrecargado, y así, el tamaño de la etapa es demasiado pequeña, por lo que el multiplicador será mayor que la unidad. Alternativamente, si la palabra código previa corresponde a cualquiera de los dos niveles, el más pequeño positivo o el más pequeño negativo, el tamaño de la etapa será grande y se deseará decrementarla, por lo que el multiplicador será menor que la unidad. El multiplicador del tamaño de la etapa depende de la cantidad Q la que está definida como:

$$Q = \frac{1 + 2|c(n - 1)|}{2^n - 1} \quad (4.9)$$

El diseño del cuantizador involucrará la elección de los multiplicadores que correspondan a cada una de la 2^n pa

labras código para un cuantizador de n-bits, así como se tendrá la consideración de imponer los límites como la Ec. (4.4) para que la relación $\Delta_{m\acute{a}x} / \Delta_{m\acute{i}n}$ controle el rango dinámico del cuantizador.

Este tipo de esquema de cuantización tiene la ventaja que el tamaño de la etapa o de la ganancia no necesitan ser retenidas o transmitidas explícitamente debido a que éstas son derivadas de las palabras código. La desventaja de éstos cuantizadores es que la sensibilidad se incrementa a errores en las palabras código, ya que tales errores implican no solamente un error en el nivel del cuantizador, sino también en el tamaño de la etapa.

La cuantización adaptiva tiene muchas características atractivas, sin embargo, no es razonable esperar ahorros dramáticos en la razón de bit, aún con esquemas sofisticados de cuantización adaptiva, ya que éstas técnicas simplemente aprovechan la distribución de amplitud de la señal de voz. Aunque también se puede aprovechar la correlación muestra a muestra por medio de la cuantización diferencial.

4.2.- CUANTIZACION DIFERENCIAL

El esquema general de cuantización diferencial que se describe en la figura 4,5, toma ventaja del modelo estático de la señal de voz, esto es, de la correlación entre muestras adyacentes de voz. El significado de la alta correlación es que, en una oración promedio la señal no cambia rápidamente muestra a muestra, ya que la diferencia entre muestras adyacentes tendrá una varianza más baja que la varianza de la señal en sí misma.

La señal de entrada al cuantizador del sistema es de

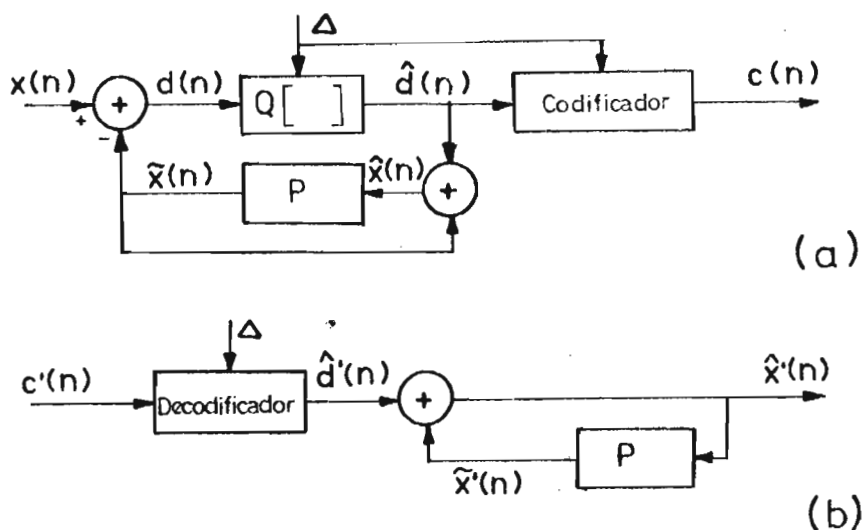


Fig. 4.5.- Representación en diagrama a bloques del esquema de cuantización diferencial.
 a) Codificador b) Decodificador

la forma:

$$d(n) = x(n) - \tilde{x}(n) \quad (4.10)$$

La Ec. (4.10), nos da la diferencia entre la muestra de entrada no cuantizada $x(n)$, y una estimada o predicción de la muestra de entrada $\tilde{x}(n)$, ésta señal predicha es la salida de un sistema predictor P , cuya entrada es una versión cuantizada de la señal de entrada $\hat{x}(n)$. La diferencia de la señal puede ser llamada el error de predicción de la señal, debido a que es la cantidad por la cual el predictor fallará para predecir exactamente la entrada.

El cuantizador puede ser fijo o adaptivo, uniforme o no uniforme, pero en cualquier caso los parámetros serán ajusta

dos para igualar la varianza de la diferencia de la señal $d(n)$. La diferencia de la señal cuantizada es representada como:

$$\hat{d}(n) = d(n) + e(n) \quad (4.11)$$

donde $e(n)$ es el error de cuantización.

Como se puede apreciar en la figura 4.5a, para producirse la versión cuantizada de la entrada, la diferencia de la señal cuantizada es adicionada al valor predicho:

$$\hat{x}(n) = \tilde{x}(n) + \hat{d}(n) \quad (4.12)$$

Sustituyendo las Ecs. (4.10) y (4.11) en la Ec. (4.12) tenemos:

$$\hat{x}(n) = x(n) + e(n) \quad (4.13)$$

que como se podrá ver, el error de cuantización es independiente de las propiedades del predictor y la muestra de voz cuantizada difiere de la entrada solamente por este error, que resulta de la diferencia de la señal. Por lo que, si la predicción es buena, la varianza de la diferencia de la señal $d(n)$, será más pequeña que la varianza de la señal de entrada $x(n)$. También se puede apreciar que para almacenar o transmitir, la diferencia de la señal cuantizada $\hat{d}(n)$ es codificada.

En la figura 4.5b, se describe el diagrama a bloques para decodificar $\hat{x}'(n)$, en donde se puede ver que para reconstruirla es necesario reconstruir la diferencia de la señal cuantizada $d'(n)$, para lo cual se utilizará el mismo tipo de predictor que fué usado en el codificador (figura 4.5a) Si $c'(n)$ es idéntico a $c(n)$, entonces $\hat{x}'(n)$ es igual a $\hat{x}(n)$ y solamente difiere de $x(n)$ por el error de cuantización

cuando la diferencia de la señal $d(n)$ es cuantizada.

Por definición, la relación señal a ruido cuantizado del sistema de la figura 4.5 es:

$$\text{SNR} = \frac{E \{x^2(n)\}}{E \{e^2(n)\}} = \frac{\sigma_x^2}{\sigma_e^2} \quad (4.14)$$

que también podemos escribir como:

$$\text{SNR} = \frac{\sigma_x^2}{\sigma_d^2} \cdot \frac{\sigma_d^2}{\sigma_e^2} = G_p \cdot \text{SNR}_Q \quad (4.15)$$

donde
$$\text{SNR}_Q = \frac{\sigma_d^2}{\sigma_e^2} \quad (4.16)$$

es la relación señal a ruido cuantizado del cuantizador

$$G_p = \frac{\sigma_x^2}{\sigma_d^2} \quad (4.17)$$

es la ganancia debida a la configuración diferencial.

La cantidad SNR_Q depende del cuantizador que en particular sea utilizado y puede ser maximizado con el cuantizador adaptivo. Si la cantidad de la ganancia (G_p) es mayor que la unidad, nos representa la ganancia obtenida en el SNR debida al esquema diferencial, por lo que un objetivo viable es maximizar la ganancia (G_p) por medio de la elección apropiada del sistema predictor P.

Para una señal dada, σ_x^2 es una cantidad fija debido a que G_p se puede maximizar si minimizamos la varianza del error de predicción σ_d^2 es decir, el denominador de la Ec. (4.17).

Debido a la importancia del predictor, es necesario es

pecificar la naturaleza de éste, donde el valor predicho $\tilde{x}(n)$ es una combinación lineal de valores cuantizados pasados:

$$\tilde{x}(n) = \sum_{k=1}^P \alpha_k \hat{x}(n-k) \quad (4.18)$$

Este valor predicho es la salida de un filtro FIR cuya función del sistema es:

$$P(Z) = \sum_{k=1}^P \alpha_k Z^{-k} \quad (4.19)$$

y cuya entrada es la señal cuantizada $\hat{x}(n)$.

De lo explicado anteriormente y dado que la cuantización diferencial toma ventaja de la correlación entre muestras, el mejoramiento de éste esquema sobre la cuantización directa depende sobre la cantidad de correlación existente y además, debido a la importancia del predictor, se deberá hacer que la ganancia (G_p) de la configuración diferencial sea óptima.

Para que la ganancia (G_p) sea óptima dependerá de la elección de los coeficientes de predicción α_k con lo que se minimiza la varianza del error de predicción (σ_d^2). El predictor fijo por su naturaleza no puede ser óptimo para todos los oradores, por lo que es más factible la utilización de la predicción adaptiva.

4.3.- SISTEMA ADPCM CON CUANTIZACION ADAPTIVA

Un sistema ADPCM con un cuantizador adaptivo del tipo de alimentación adelantada es mostrada en forma de diagrama a bloques en la figura 4.6.

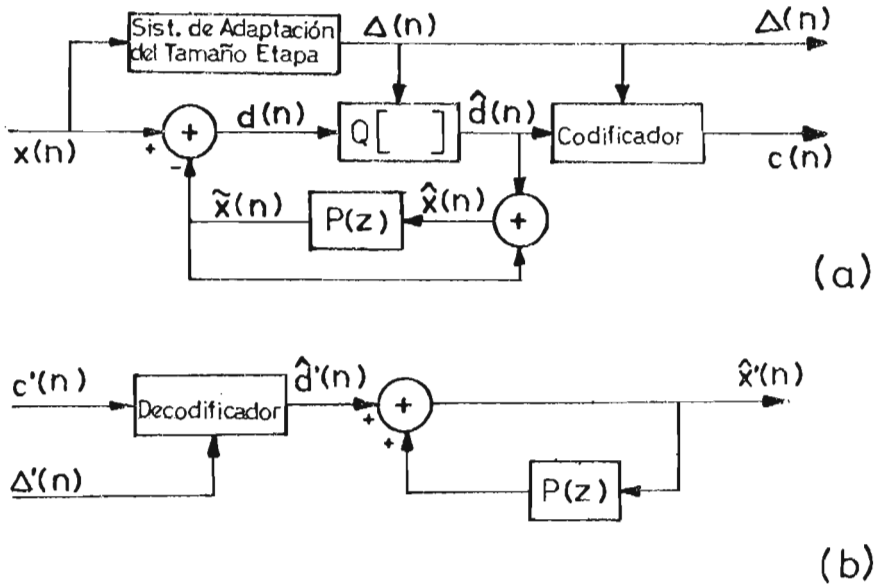


Fig. 4.6.- Representación en diagrama a bloques de un sistema ADPCM con cuantización adaptiva del tipo de alimentación adelantada:

a) Codificador

b) Decodificador

En este sistema el tamaño de la etapa es proporcional a la varianza de la entrada del cuantizador, sin embargo, como la diferencia de la señal $d(n)$ es proporcional a la entrada, el tamaño de la etapa $\Delta(n)$ se puede controlar desde cualquiera de las dos; desde la diferencia de la señal $d(n)$ o desde la señal en sí misma $x(n)$, como se aprecia en la figura.

El algoritmo de cuantización de alimentación adelantada se explicó en la sección 4.1.1, en donde vimos que con este algoritmo se obtiene una mejora de al menos 5 dB que adicionada a la mejora mínima de 6 dB obtenida por la configura-

ción diferencial con predicción fija, el sistema ADPCM realizará un SNR mayor en 10-14 dB que el obtenido con un cuantizador fijo, con el mismo número de niveles.

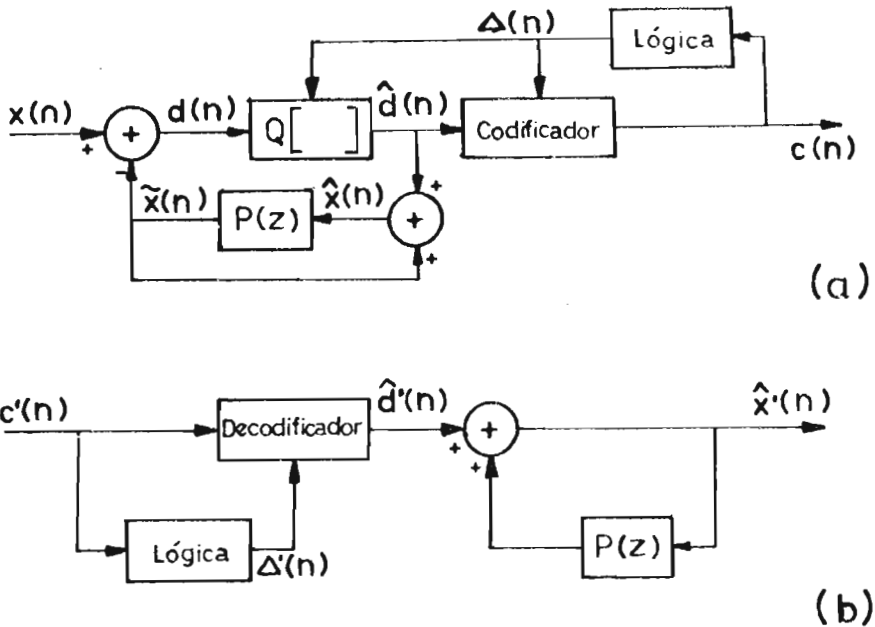


Fig. 4.7.- Representación en diagrama a bloques del sistema ADPCM con cuantización adaptativa del tipo de realimentación.

a) Codificador

b) Decodificador

En la figura 4.7, se muestra en forma de diagrama a bloques un sistema ADPCM con un cuantizador adaptativo del tipo de realimentación, cuyo algoritmo fué explicado en la sección 4.1.2, y donde la mejoría obtenida fué de 4-6 dB sobre un cuantizador fijo.

En cualquiera de los dos tipos de cuantización adaptativa

nos proporcionan una mejoría en el SNR de 10-14 dB además de proporcionar un mejoramiento del rango dinámico. Sin embargo, como ya vimos, con control de realimentación tiene la ventaja de que el tamaño de la etapa no necesita ser almacenada o transmitida, debido a que se deriva de la secuencia de las palabras código $c(n)$, pero la calidad de la señal reconstruida a la salida es más sensitiva a errores en transmisión.

Con un control de alimentación adelantada es posible transmitir el tamaño de la etapa $\Delta(n)$ con protección de error debido a que la señal se representa con las palabras código $c(n)$ junto con el tamaño de la etapa $\Delta(n)$.

4.4.- SISTEMA ADPCM CON CUANTIZACION Y PREDICCIÓN ADAPTIVA

En la figura 4.8, se muestra en diagrama a bloques un sistema general ADPCM con cuantización y predicción adaptiva, donde las líneas punteadas indican que ambos algoritmos de adaptación, en el predictor y en el cuantizador puede ser cualquiera de los dos tipos, de alimentación adelantada o de realimentación.

Si utilizamos el control de alimentación adelantada, para representar la señal de voz por completo, a las palabras código $c(n)$ se le adicionará el tamaño de la etapa $\Delta(n)$ o los coeficientes del predictor $\alpha(n) = \{\alpha_k(n)\}$, o ambos si utilizamos cuantización y predicción adaptiva.

El valor predicho es:

$$\tilde{x}(n) = \sum_{k=1}^p \alpha_k(n) \hat{x}(n-k) \quad (4.20)$$

donde los coeficientes del predictor son asumidos para ser

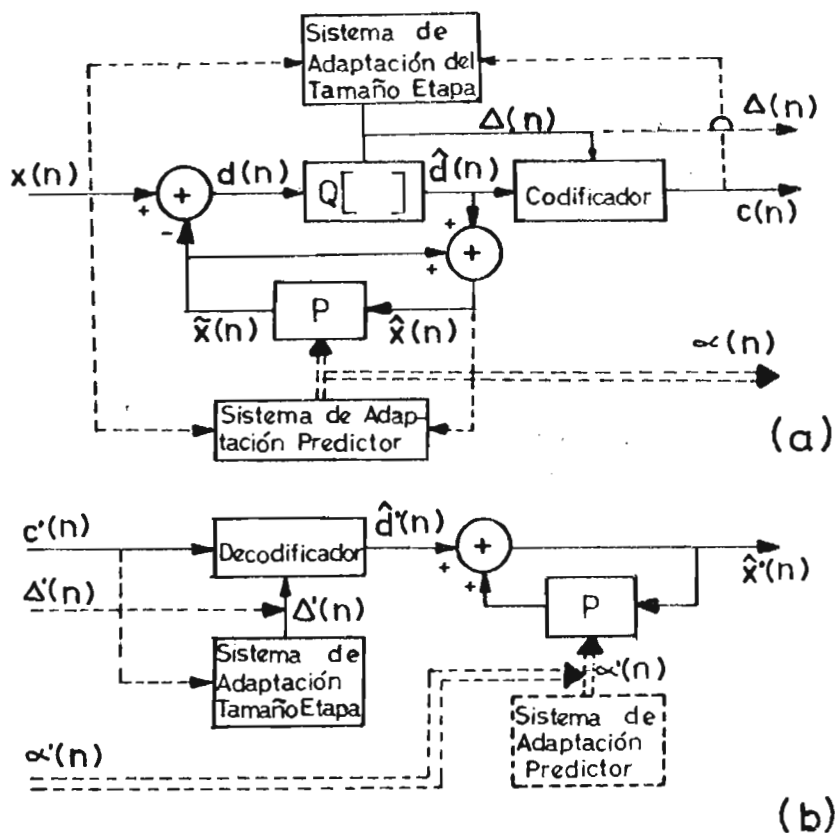


Fig. 4.8.- Representación general en diagrama a bloques de un sistema ADPCM con cuantización y predicción adaptiva:

a) Codificador

b) Decodificador

dependientes del tiempo. También es común asumir que las propiedades de la señal de voz se mantienen fijas en intervalos cortos de tiempo cuando se adoptan los coeficientes del predictor $\alpha(n)$, por tanto estos coeficientes se eligen para minimizar el error medio cuadrático de predicción.

Usando control de alimentación adelantada, la adaptación del predictor está basado sobre mediciones de la señal de entrada y dado que en este sistema se toma ventaja de la alta correlación de la voz, los coeficientes del predictor óptimo satisfacen:

$$R_n(j) = \sum_{k=1}^p \alpha_k(n) R_n(j-k) \quad j = 1, 2, \dots, p \quad (4.21)$$

donde $R_n(j)$ es la función de autocorrelación en tiempo corto;

$$R_n(j) = \sum_{m=-\infty}^{\infty} x(m) w(n-m) x(j+m) w(n-m-j) \quad 0 \leq j \leq p \quad (4.22)$$

donde $w(n - m)$ es una función ventana, que puede ser una ventana rectangular o una Hamming de longitud N , y está posicionada a n muestras de la secuencia de entrada. La duración de la ventana puede ser igual a los intervalos entre valores estimados o puede ser algo más larga.

Debido a que los parámetros de voz varían bastante lento, los parámetros del predictor $\alpha(n)$ no se ajustan con mucha frecuencia, por ejemplo, mantenerlos fijos entre estimaciones.

Si utilizamos el control de realimentación en el predictor adaptivo, la función de correlación estará basada sobre la señal cuantizada en lugar de la de entrada, por lo que la Ec. (4.22) será reemplazada por:

$$R_n(j) = \sum_{m=-\infty}^{\infty} \hat{x}(m) w(n-m) \hat{x}(j+m) w(n-m-j) \quad 0 \leq j \leq p \quad (4.23)$$

donde la ventana debe ser de la forma:

$$\begin{aligned} w(m) &= 1 & 0 \leq m \leq N - 1 \\ &= 0 & \text{de otra forma} \end{aligned} \quad (4.24)$$

donde el estimado de los coeficientes del predictor debe estar basado sobre valores cuantizados pasados, pues los valores actuales no pueden ser obtenidos hasta que los coeficientes del predictor estén disponibles.

Como en el caso del control de cuantizador adaptivo, el control del predictor adaptivo del tipo de realimentación, tiene la ventaja que solamente se necesitan ser transmitidas las palabras código $c(n)$. Sin embargo, no ha sido ampliamente utilizado debido a su inherente sensibilidad a errores y al inferior funcionamiento que resulta basando el control sobre una entrada ruidosa.

Expresando cuantitativamente los beneficios de la predicción adaptiva, Noll examinó la dependencia de la ganancia de predicción (G_p) sobre el orden del predictor (P), para predicción fija y adaptiva, mediante la siguiente relación:

$$10 \log_{10} (G_p) = 10 \log_{10} \left[\frac{E\{x^2(n)\}}{E\{d^2(n)\}} \right] \quad (4.25)$$

Esta dependencia se representa en forma gráfica en la figura 4.9. donde la curva inferior se obtiene computando un término largo estimado de la autocorrelación para la pronunciación de un material de voz dado y resolviendo el conjunto de coeficientes del predictor que satisfacen:

$$\rho = C \alpha \quad (4.26a)$$

donde:

$$\rho = \begin{bmatrix} \rho(1) \\ \rho(2) \\ \vdots \\ \rho(P) \end{bmatrix} \quad (4.26b)$$

$$C = \begin{bmatrix} (1 + \frac{1}{SNR}) & \rho(1) & \dots & \rho(P-1) \\ \rho(1) & (1 + \frac{1}{SNR}) & \dots & \rho(P-2) \\ \vdots & \vdots & \dots & \vdots \\ \rho(P-1) & \rho(P-2) & \dots & (1 + \frac{1}{SNR}) \end{bmatrix} \quad (4.26c)$$

$$\alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_P \end{bmatrix} \quad (4.26d)$$

donde: $SNR = \frac{\sigma_x^2}{\sigma_e^2}$

Así, el vector de los coeficientes del predictor óptimo se obtienen como la solución de la matriz de la Ec. (4.26a)

$$\alpha = C^{-1} \rho \quad (4.27)$$

por lo que, obtenemos una ganancia máxima de aproximadamente 10.5 dB como se aprecia en la figura.

La curva superior fué obtenida encontrando el valor de la longitud de la ventana, y los coeficientes de predictor $\alpha(n)$ que maximizan G_p a través de la pronunciación com-

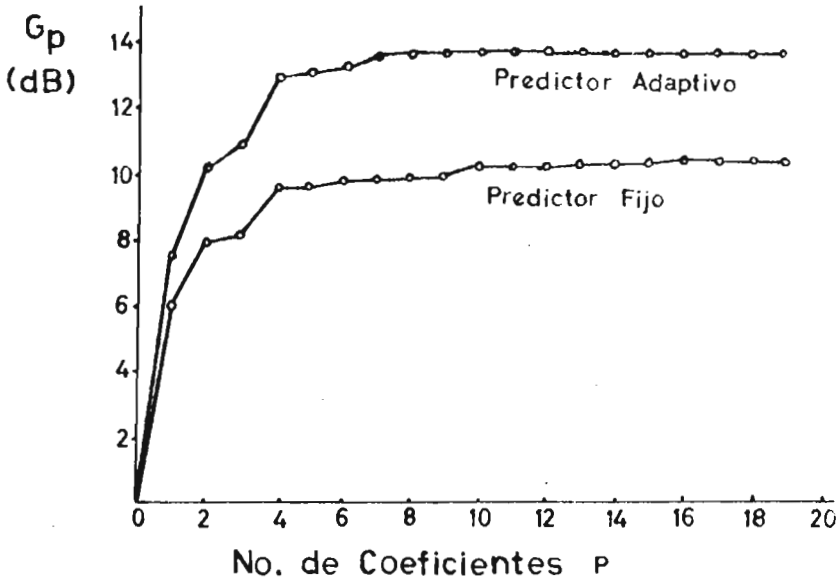


Fig. 4.9.- Representación gráfica para un orador femenino de la dependencia de la ganancia de predicción (G_p) sobre el orden del predictor (P), para predicción fija y para predicción adaptiva.

plata para un valor fijo para el orden del predictor (P), la ganancia para predicción adaptiva como se aprecia, es de aproximadamente 14 dB.

De éste modo, los límites superiores sobre el funcionamiento del sistema ADPCM y DPCM con predicción fija y adaptiva es de 10.5 y 14 dB respectivamente. Sin embargo, no se hace evidente (en las curvas de la figura 4.9) que el predictor fijo óptimo es muy sensitivo al orador y al material de voz, mientras el esquema de predicción adaptiva es inherentemente menos sensitivo. El predictor adaptivo tiende a trasladar redundancias de la señal de voz, pero si fuera posible una predicción perfecta, la señal de error de pre

dicción $d(n)$, será completamente no correlacionada (ruido blanco).

4.5.- CONVERSION DIRECTA DE CODIGO DIGITAL

Para la utilización de la técnica digital ADPCM en sistemas donde hay una representación digital diferente, será necesario convertir una representación digital a otra evitando la degradación de la calidad de la voz sin intervenir procesamiento analógico, es decir, usar técnicas enteramente digitales. Además, en situaciones donde se procesa señales de voz, es necesario representar la señal en forma de PCM uniforme, debido a que la filtración digital es desarrollada sobre las muestras de la señal de voz,

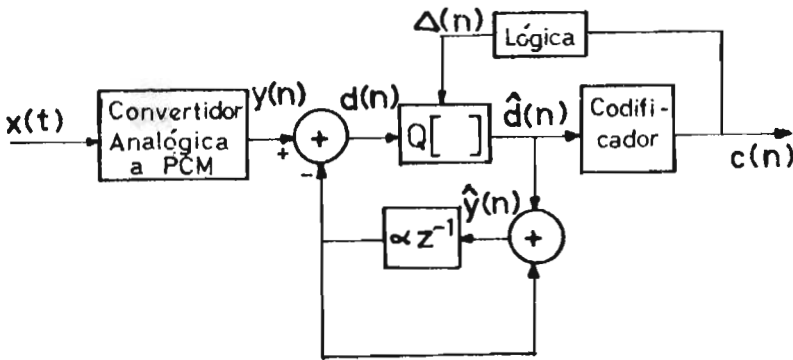


Fig. 4.10.- Representación en diagrama a bloques de un convertidor de código PCM ADPCM.

En la figura 4.10, se muestra la aproximación básica para implementar el algoritmo ADPCM directamente de las mues

tras PCM, donde las operaciones aritméticas y lógicas son realizadas con hardware digital ordinario. Sin embargo, hay una consideración importante en el proceso de conversión de código, ésto es la degradación introducida por el procedimiento adicional.

De la figura resulta claro, ver que la entrada PCM tendrá algún error de cuantización, el cual puede ser caracterizado por una relación señal a ruido SNR_1 , similarmente el algoritmo ADPCM introducirá un error adicional caracterizado por SNR_2 . Si asumimos que dicho error es independiente de la señal original y del error de cuantización, entonces el SNR total del sistema se puede aproximar mediante la ecuación siguiente:

$$SNR = \frac{SNR_1}{1 + \frac{1 + SNR_1}{SNR_2}} \quad (4.28)$$

El convertidor directo de un código a otro, cobra gran importancia, pues sería demasiado costoso que con el avance tecnológico se realizaría continuamente un cambio total de todos los sistemas en el mundo. Debido al incremento de necesidades y de usuarios de los canales de comunicación en banda de voz, es conveniente la utilización de un convertidor para el uso más eficiente de los canales ya existentes.

4.6.- RESUMEN

Una forma sintética de entender el funcionamiento y el comportamiento del sistema ADPCM explicado en las secciones anteriores de este capítulo, se resumen en la tres etapas siguientes:

- 1.- Se obtiene un valor diferencial $d(n)$ que resulta de la sustracción de la entrada no cuantizada y el valor de predicción de la entrada como se aprecia en la Ec.(4.11)
- 2.- La obtención del valor de cuantización o tamaño de la etapa $\Delta(n)$, que resulta multiplicando el valor de la etapa previa por un multiplicador que depende solamente de las palabras código previas como se ve en la Ec.(4.7)
- 3.- Finalmente el valor diferencial es expresado en términos de los valores de cuantización, por lo que la diferencia entre el valor actual y el valor predicho es siempre pequeño y por tanto puede ser codificada usando solamente cuatro bits.

Sin embargo, si de entrada tenemos un código PCM, las etapas anteriores quedarán de la siguiente forma:

- 1.- El valor diferencial $d(n)$ se obtiene de la diferencia del valor código PCM previo del valor actual.
- 2.- El valor de cuantización $\Delta(n)$ se obtiene multiplicando el valor de la etapa previa por un coeficiente que depende del valor previo del código ADPCM.
- 3.- Esta etapa es exactamente igual a la ya explicada.

Las etapas anteriormente mencionadas se muestran en forma gráfica en la figura 4.11.

La técnica de codificación de voz ADPCM ofrece esquemas alternativos de implementación basados sobre la configuración básica de la figura 4.6. Por su buena calidad y su baja razón de bit, ofrece ventajas para transmisión y sobre todo en

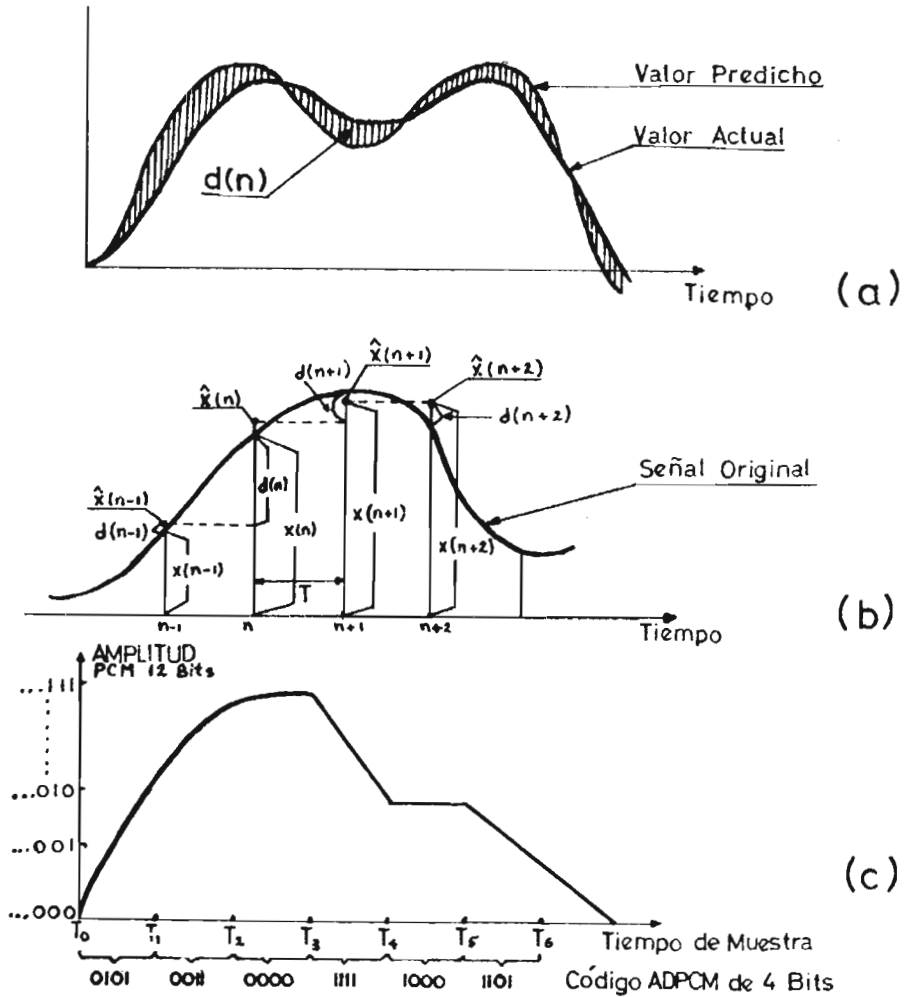


Fig. 4.11.- Forma gráfica de las etapas de una señal procesada mediante ADPCM:
a) Muestra gráfica de la etapa 1;
b) Forma de onda en donde se puede apreciar el proceso de las dos primera etapas;
c) Señal procesada mediante ADPCM (etapa 3).

almacenamiento de datos.

Un ejemplo palpable de lo anteriormente dicho, es el algoritmo ADPCM de la CCITT, en donde el codificador acepta en la entrada un PCM de 64 Kbps y lo convierte en un ADPCM de 32 Kbps, figura 4.12.(12)

En el codificador, la señal de entrada PCM es convertida primeramente de la ley μ (o ley A) a una forma lineal. El valor predicho se obtiene del predictor adaptivo, el cual se sustrae del valor actual para obtener la señal diferencia $d(n)$.

La señal diferencia tiene un rango dinámico más pequeño que la entrada $x(n)$, con lo cual puede ser codificada usando solamente un cuantizador de 4 bits, con un cuantizador adaptivo. El cuantizador inverso, generará la versión cuantizada de la señal diferencia, la cual se utiliza para manejar el predictor adaptivo y para obtener la señal reconstruida.

Los 4 bits del cuantizador adaptivo que se utiliza en el algoritmo de la CCITT es del tipo no uniforme; lo que es, los niveles de decisión y de reconstrucción no están uniformemente espaciados. Por tanto, el rango de adaptación está de acuerdo a la señal diferencia estática. Con esto, se tienen dos velocidades de adaptación; rápida para señales de voz y lenta para señales de datos en banda de voz. El predictor adaptivo tiene 2 polos y 6 ceros en su función de transferencia.

La estructura particular de polos-ceros del predictor y del esquema de adaptación empleado en este algoritmo, es para optimizar y asegurar la estabilidad y el buen funcionamiento en presencia de los errores de transmisión y con diferen-

tes clases de señales de entrada.

Como se puede apreciar en la figura, el decodificador es la contraparte del codificador. Sin embargo, el algoritmo de CCITT agrega en el decodificador un *ajuste síncrono de codificación*, éste reduce la acumulación de errores de cuantización durante la conversión de código PCM y ADPCM.

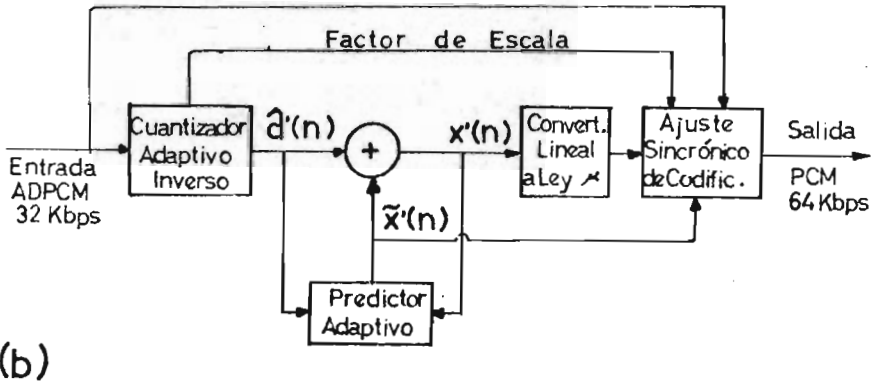
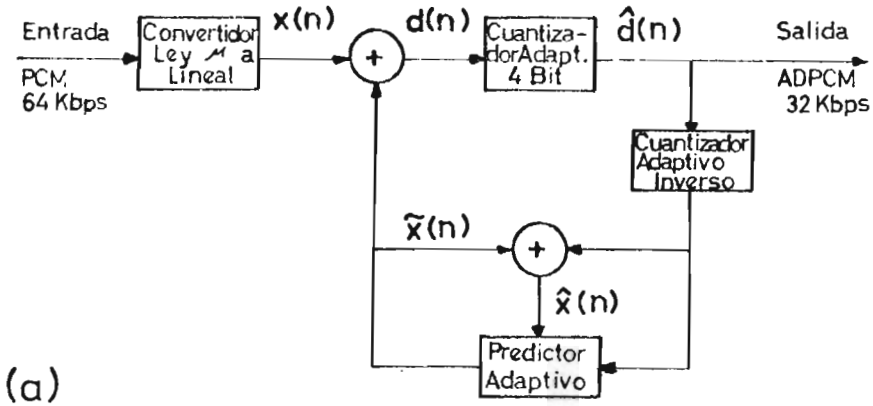


Fig. 4.12.- Algoritmo ADPCM de la CCITT.

a) Codificador

b) Decodificador

C
A
P
I
T
U
L
O

5.- ESTADO ACTUAL DE LA TECNOLOGIA PARA IMPLEMENTAR SINTESIS DE VOZ

El procesamiento digital de señales de voz se conoce desde hace poco más de 25 años con una gran actividad y con un amplio rango de aplicaciones en las comunicaciones y en el desarrollo de las máquinas lectoras y de respuesta hablada. Sin embargo, actualmente el gran auge tomado es porque la tecnología tiene la factibilidad de realizar algoritmos complejos en forma económica.

Los avances logrados en microelectrónica permite implementar algoritmos en unos pocos circuitos integrados, por lo que, todos los tipos de representaciones de las señales se han desarrollado, pudiendose implementar en tiempo real y a un razonable costo. En la figura 5.1, se muestra como se han incrementado y desarrollado las técnicas de codificación de voz, que se han podido implementar en forma económica en el transcurso de los últimos años.

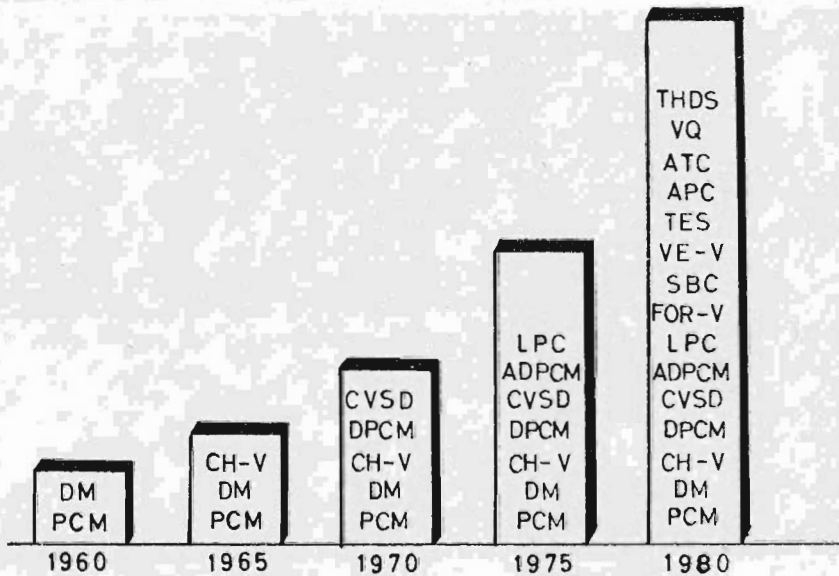
Por las perspectivas que se presentan de un gran mercado para esta área, los fabricantes de semiconductores a nivel mundial se han avocado al desarrollo de procesadores de señales cada vez más rápidos y poderosos.

El avance de la tecnología VLSI proporciona un gran camino a generaciones de procesadores y ha sido tan rápida que la primera generación inicia en 1979 con el 2920 de INTEL, en 1982 se da la segunda generación con el TMS 32010 de TEXAS INSTRUMENT, y se menciona que en 1986 se tendrá ya la tercera generación. Esto puede observarse gráficamente en la figura 5.2.

La tecnología VLSI ha hecho posible el uso de chips en

Fig. 5.1.- Gráfica del crecimiento de las Técnicas de Codificación de Voz.

PCM	Modulación por Codificación de Pulsos
DM	Modulación Delta
CH-V	Vocoder de Canal
DPCM	PCM Diferencial
CVSD	Delta de Pendiente Variable Continua
ADPCM	PCM Diferencial Adaptivo
LPC	Código de Predicción Lineal
FOR-V	Vocoder de Formantes
SBC	Codificación en Sub-Bandas
VE-V	Vocoder Excitado por Voz
TES	Voz Codificada en Tiempo
APC	Código de Predicción Adaptiva
ATC	Codificación Adaptiva por Transformada
VQ	Vector de Cuantización
THDS	Escalamiento de Armónicas en el Dominio del Tiempo



los años recientes para aplicaciones digitales de la voz. En el área de codificación de forma de onda, se han introducido chips para conversión completa A/D y D/A en PCM μ ley que incluye filtrado anti-aliasing, chips para ADM, ambos con interés de aplicación en telefonía y también con creciente interés en esta área a 32 Kbps chips para ADPCM. En el área de "vocoders" también están disponibles los sintetizadores en chips, como el utilizado en juguetes educativos con respuesta sonora.

El área de la tecnología VLSI con fuerte impacto en voz digital, es el desarrollo de microprocesadores de alta velocidad y circuitos integrados programables, con los que es posible implementar algoritmos como el ADPCM y el SBC (Codificación en Sub-Bandas). Para algoritmos de alta complejidad como el ATC (Codificación Adaptiva por Transformada), APC (Codificación Predictiva Adaptiva) y Analizadores "Vocoder", son realizados en tiempo real usando procesadores o "hardware" de propósito especial.

Ejemplos de este tipo de circuitos integrados de propósito especial de procesamiento de voz son:

- MSM 5218 (ADPCM Analizador y Sintetizador) de OKI
- TMS 5220 (Síntesis por LPC) de TEXAS INSTRUMENT
- D 2911 (Codec PCM ley A) de INTEL
- MC 3518 (CVSD) de MOTOROLA

y su número tiende a aumentar en el mercado de semiconductores.

5.1.- COMPROMISO ENTRE COMPLEJIDAD Y FUNCIONAMIENTO

En el diseño de un codificador de voz, existen tres clases de compromisos principales:

- 1.- CALIDAD TOTAL
- 2.- COSTO O COMPLEJIDAD
- 3.- RAZON DE BIT REQUERIDO

Estos compromisos se ilustran en la figura 5.3, para algunas realizaciones en "hardware". La escala vertical corresponde a la calidad de la voz, medida de 1 a 5 basada sobre un tanteo de opinión media para el juicio de los oyentes. La escala horizontal ofrece una medida brusca de complejidad "hardware".

Las diferentes curvas representan la razón de bit y los puntos sobre las curvas corresponden a la realización "hardware" que ha sido implementada en el laboratorio. Por tanto, de la figura es posible obtener una valoración relativamente buena de los compromisos de calidad, complejidad y razón de bit.

5.2.- RESUMEN

El desarrollo constante de la tecnología digital de la voz ha hecho posible que una vez que la voz se encuentra en forma digital, puede ser fácilmente manipulada, revisada, almacenada y transmitida, con lo cual puede ser posible corregir y compilar letras y mensajes sonoros, almacenar y transmitir información y datos con una razón de bit según necesidades y facilidades con tendencias de ser lo más baja posible.

Todas estas facilidades e innovaciones al ser adicionadas a la computadora hacen posible una interacción más estrecha HOMBRE-MAQUINA, y con nuevas posibilidades para que las redes digitales puedan ser manejadas y conmutada con voz y con información de datos.

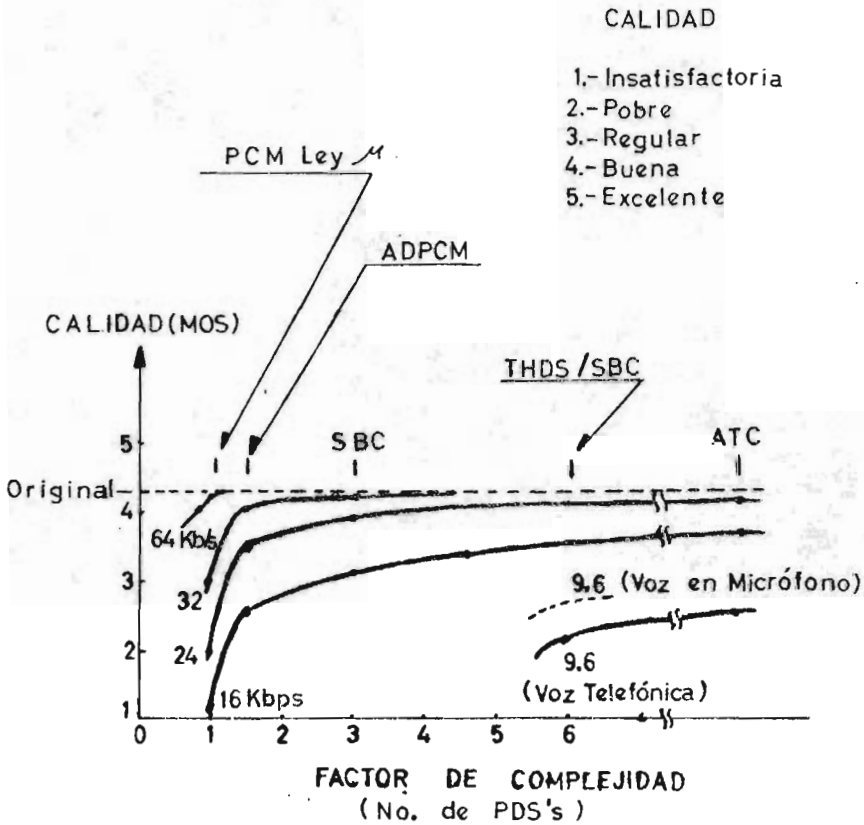


Fig. 5.3.- Compromiso (trade-off) en forma gráfica para algunas realizaciones "hardware", Calidad contra la razón de bit y la complejidad.

6.- DISEÑO DE UN SIMULADOR PARA EVALUACION DE VOZ MEDIANTE ADPCM

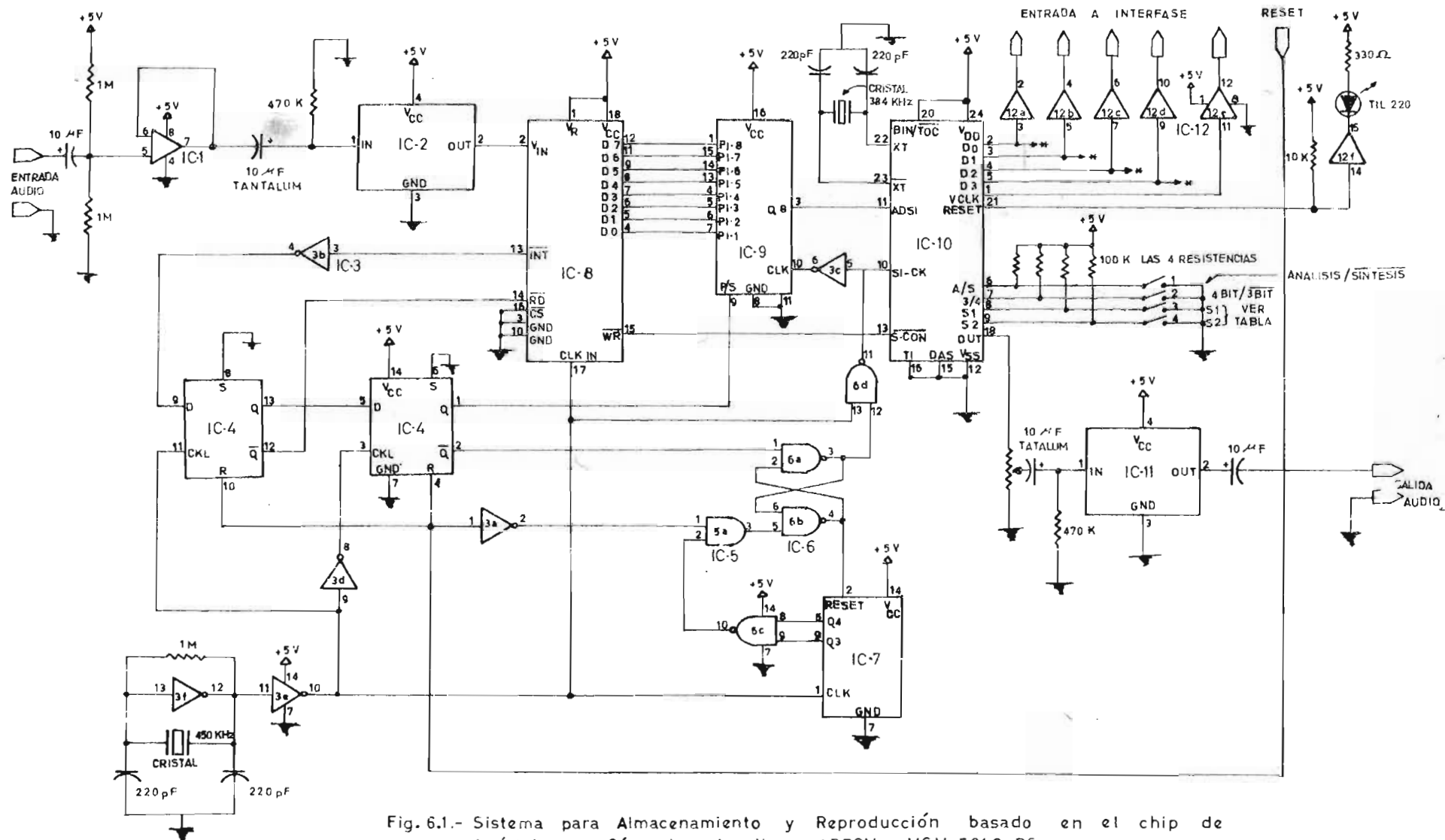
La generación de voz por medio de datos digitales -SIN-
TESIS DE VOZ - mediante la técnica de compresión ADPCM, que
fué expuesta en forma teórica ahora se mostrará la realiza-
ción práctica de dicha teoría. Pues debido a los avances
tecnológicos, ahora es posible tener el algoritmo ADPCM en
circuitos integrados con una facilidad real para tener acce-
so a ellos, a un costo efectivo que produce una salida alta-
mente inteligible y hace bastante eficiente la utilización
de la memoria. (26).

En la figura 6.1, se expone el diseño de un sistema de
Codificación y Decodificación ADPCM. Este sistema de sínte-
sis de voz utiliza componentes CMOS (Semiconductor Oxido-Me-
tal Complementario), con lo cual se obtiene un sistema de
costo efectivo con las características descritas.

En este capítulo se analiza el circuito de la figura
6.1, sin embargo, en el apéndice C se tratan los conceptos
básicos de funcionamiento de algunos dispositivos utilizados
en el sistema, como el amplificador operacional, los conver-
tidores A/D y D/A, algunas técnicas de conversión, etc.

El circuito de síntesis de voz está basada en el circui-
to integrado (IC) MSM 5218 RS de OKI Semiconductor, con el
cual se efectúa el análisis y síntesis. Debido a ésto tene-
mos la alternativa de almacenar datos de voz o bien reproducir
voz sонтética. La figura 6.1 podemos dividirla en las
siguientes etapas para su explicación:

- 1.- Sincronización
- 2.- Proceso de Limitación del Ancho de Banda



- IC-1.- MC 1458 AMPLIFICADOR OPE-RACIONAL
- IC-2 e IC-11.- ALP-2 FILTRO PASOBAJAS
- IC-3.- CD 4069 INVERSORES
- IC-4.- CD 4013 FLIP-FLOP D
- IC-5.- CD 4081 COMPUERTAS AND
- IC-6.- CD 4011 COMPUERTAS NAND
- IC-7.- CD 4024 CONTADOR BINARIO
- IC-8.- MSM 5204 RS CONVERTIDOR A/D 8 BITS
- IC-9.- CD 4014 REGISTRO DE CORRIMIENTO ESTADICO
- IC-10.- MSM 5218 RS ANALISIS / SINTESIS DE VOZ ADPCM
- IC-12.- CD 4050

S1	S2	FREC. MUESTREO
0	0	4 KHZ
0	1	5 KHZ
1	0	8 PHZ
1	1	EXTERNA

* BUS DE DATOS BIDIRECCIONAL

Fig. 6.1.- Sistema para Almacenamiento y Reproducción basado en el chip de Análisis y Síntesis de Voz ADPCM MSM 5218 RS.

- 3.- Digitalización de la Señal (Muestreo, Cuantización y Codificación)
- 4.- Análisis y Síntesis
- 5.- Almacenamiento y Reproducción

6.1.- SINCRONIZACION

La sincronización tiene la función como su nombre lo indica de proporcionar sincronía a todos los dispositivos componentes del sistema de síntesis de voz, como compuertas, el contador, el flip-flop y al convertidor A/D.

La frecuencia del reloj no es un valor crítico por lo que puede oscilar entre 450 y 500 KHz, y se obtiene por medio de un cristal, el cual es un dispositivo piezoeléctrico que transforma la energía de voltaje en oscilaciones de voltaje.

6.2.- LIMITACION DEL ANCHO DE BANDA

La etapa de limitación del ancho de banda de la señal de voz, es significativamente importante. Sin embargo, resulta fácil realizarla como se muestra en la figura 6.2.

Como se explicó en el capítulo 1, para poder hacer uso de los métodos PDS, y por lo tanto para que sea posible el muestreo de la señal analógica, ésta debe ser de banda limitada. Sin embargo, la señal analógica de voz no es inherentemente de banda limitada, por lo que la señal para ser muestreada es filtrada por un filtro pasobajas con una frecuencia menor o igual a la mitad de la frecuencia de muestreo, con lo que además garantizaremos la no ocurrencia de traslapamiento de la señal debido al muestreo.

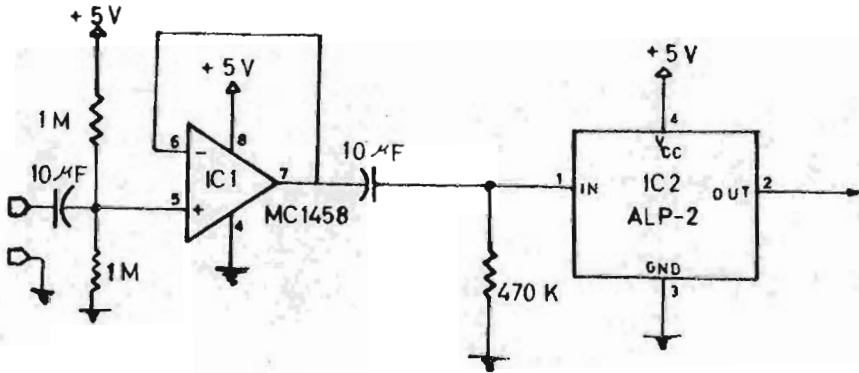


Fig. 6.2.- Limitación del Ancho de Banda.

El filtro activo pasobajas (IC 2) ALP-2 tiene una frecuencia de corte de 1.7 KHz. Aunque teóricamente con un rango de muestreo de 8 KHz, todas las armónicas necesarias para reproducción son capturados con 4 KHz, sin embargo, una frecuencia de corte más baja da una reproducción de sonidos mejor.

El circuito amplificador que funciona como un seguidor de voltaje sirve para separar las impedancias, es decir:

Impedancia de entrada tiende a infinito $Z_i \rightarrow \infty$

Impedancia de salida tiende a cero $Z_o \rightarrow 0$

Con lo que se proporciona un aislamiento entre la fuente de la señal y la carga, de esta forma se impiden interacciones perjudiciales o efectos de carga.

6.3.- DIGITALIZACION DE LA SEÑAL (MUESTREO, CUANTIZACION Y CODIFICACION)

La etapa de digitalización de la señal se muestra en la figura 6.3. En esta etapa obtendremos la representación digital de la señal de voz, es decir, se realizará el proceso de muestreo, cuantización y codificación. Sin embargo, como se explicó en el capítulo 3 en la práctica no es posible mostrar o realizar separadamente cada uno de estos procesos. Estos son realizados únicamente con el convertidor A/D.

El convertidor A/D utilizado aquí, es uno de 8 bits de bajo costo en lugar de uno de resolución más alta como lo es un convertidor A/D de 12 bits que es más costoso.

El diagrama a bloques funcional del convertidor A/D CMOS MSM 5204 RS se muestra en la figura 6.4, en donde se aprecia que consta de un sistema de muestreo y retención que es un circuito que mide las cantidades analógicas en intervalos pequeños y retiene dichos valores constantes, este circuito facilita la entrada directa de las señales analógicas que varían en forma rápida, las cuales vienen del filtro activo pasobajas. Aquí se realiza el muestreo.

El convertidor utiliza una técnica de conversión sucesiva con un arreglo o escalera de capacitores (apéndice C). La señal de salida de este arreglo es comparada en el comparador con una señal de referencia, que en este caso es de + 5V. La señal que sale de aquí está cuantizada y codificada.

La conversión de datos realizada por el convertidor A/D se inicia cuando la siguiente etapa (Análisis y Síntesis en el 5218) la requiere. Por lo que el registro de corrimiento recibe por medio de la pata \overline{WR} (15) la orden de comenzar la conversión, donde el margen de elevación dispara el proceso.

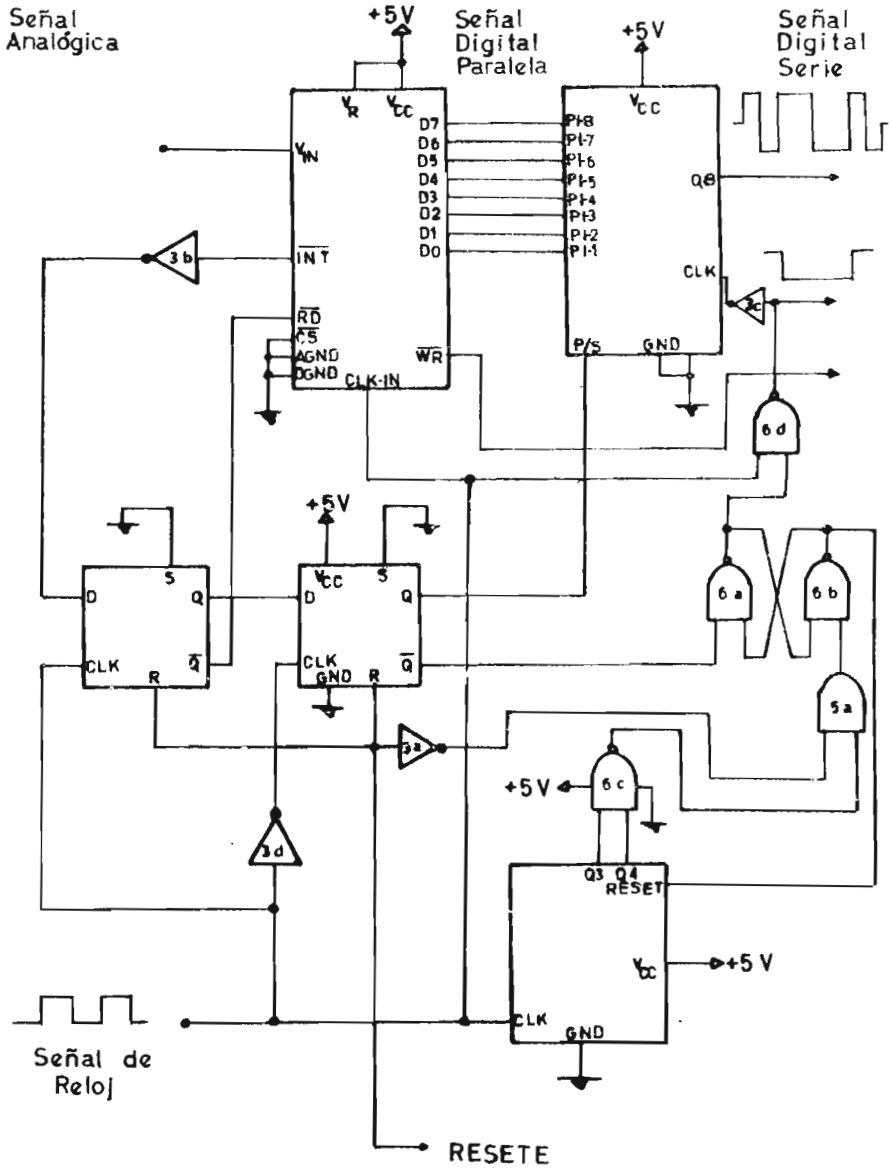


Fig. 6.3.- Digitalización de la Señal.

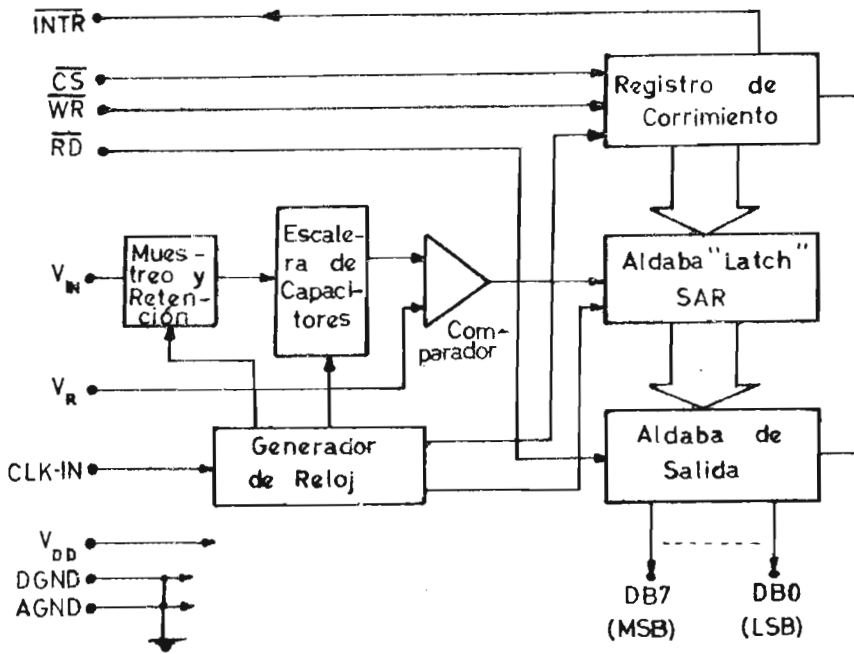


Fig. 6.4.- Diagrama a bloques funcional del Convertidor A/D CMOS MSM 5204 RS.

Cuando la conversión está completa, con la conexión existente entre el registro de corrimiento y la aldaba (latch) de salida se accesan los 8 bits del dato convertido en las líneas de salida $D_0 - D_7$ del 5204. Esto es posible cuando la línea de lectura A/D, \overline{RD} (14) se pone en un nivel bajo. Igualmente con esta misma conexión se pone fin a la conversión, debido a que cuando ésta termina envía una señal de nivel bajo la cual sale del 5204 por medio de la pata \overline{INTR} (13), y después de pasar por el inversor llega al flip-flop tipo D (retraso) IC CD 4013 con lo cual cambia el nivel de la señal de salida \overline{Q} que está conectada a la pata \overline{RD} .

La conversión completa del dato digital de 8 bits a una frecuencia de 450 KHz el 5204 la realiza en aproximadamente en 73 μ seg, pero si la frecuencia es de 500 KHz, la realizará en 66 μ seg.

Los 8 bits de la representación digital del dato que sale del 5204 se encuentra en forma paralela, sin embargo, el 5218 la requiere en serie. Por tanto, los 8 bits en paralelo entran al IC CD 4014 que es un registro de corrimiento estático, el cual puede recibir como entrada una señal serie o paralela síncrona y entregará una salida serie.

Aquí, el 4014 es un registrador de 8 etapas de entradas en paralelo y salidas en serie con reloj común. Cada etapa del registrador es un flip-flop amo-esclavo tipo D.

La entrada serie como la paralela se hacen síncronas dentro del registrador con la línea de transición positiva del reloj y con el control del control de entrada. Cuando el control de entrada serie/paralelo es un nivel bajo, el dato es trasladado dentro del registrador de 8 etapas sincrónicamente con la transición positiva de la línea del reloj. Cuando el control de entrada paralela/serie es un nivel alto, el dato es aglomerado dentro del registrador de 8 etapas vía las líneas de entradas paralelas y síncronas con la transición positiva del reloj.

De esta forma los 8 bits en paralelo son transformados a una forma serie, la cual sale por la pata de salida Q_8 (3) del 4014.

Sin embargo, el 5218 espera la entrada de 12 bits y no de 8, por tanto los cuatro bits faltantes son proporcionados como ceros por el contador CD 4024 y las cuatro compuertas NAND del IC CD 4011 (con lo cual formamos un enclavamiento).

El 4024 es un contador binario, en el cual todas las etapas son flip-flop tipo amo-esclavo. El estado del contador depende de la transición negativa de cada pulso de entrada que en este caso es el del reloj. Un nivel alto en la línea RESET reajusta el contador a un estado cero.

Las salidas del contador van a las compuestas NAND. Estos componentes proporcionan los cuatro pulsos extras como datos de nivel lógico cero sobre la entrada SI-CK (10), con lo cual el 5218 recibe los 12 bits que necesita.

6.4.- ANALISIS Y SINTESIS

El MSM 5218 RS es un IC completo, donde efectúa el análisis y la síntesis de voz usando el método de compresión de datos ADPCM.

En la figura 6.5, se muestra el diagrama a bloques funcional del 5218 que contiene: Un sincronizador y control al cual como se ve, por medio de las patas XT (22) y \overline{XT} (23) entra una señal de un resonador (que puede ser cerámico o de cristal) externo de 384 KHz con lo cual funciona el oscilador interno y pasa al sincronizador. También a éste llegan otras dos patas S_1 (8) y S_2 (9), con lo cual con estas entradas se selecciona la frecuencia de muestreo o el modo externo de acuerdo a la siguiente tabla:

TABLA 6.1.- FRECUENCIA DE MUESTREO DEL 5218

S_1	S_2	Frecuencia de Muestreo
0	0	4 KHz
0	1	6 "
1	0	8 "
1	1	Modo Externo

Con estas dos patas S_1 y S_2 se selecciona la frecuencia de muestreo y así mismo la frecuencia de trabajo tanto para el 5218 como para el convertidor A/D 5204, pues por medio de la pata $\overline{S-CON}$ (13) el 5218 solicita al convertidor el inicio de la conversión. Igualmente, con la pata $VCLK$ (1) que es una pata de dato de entrada/salida (E/S) en la cual sale una señal cuya frecuencia es igual a la seleccionada por S_1 y S_2 y que va a la interfase de la computadora para la grabación de la memoria.

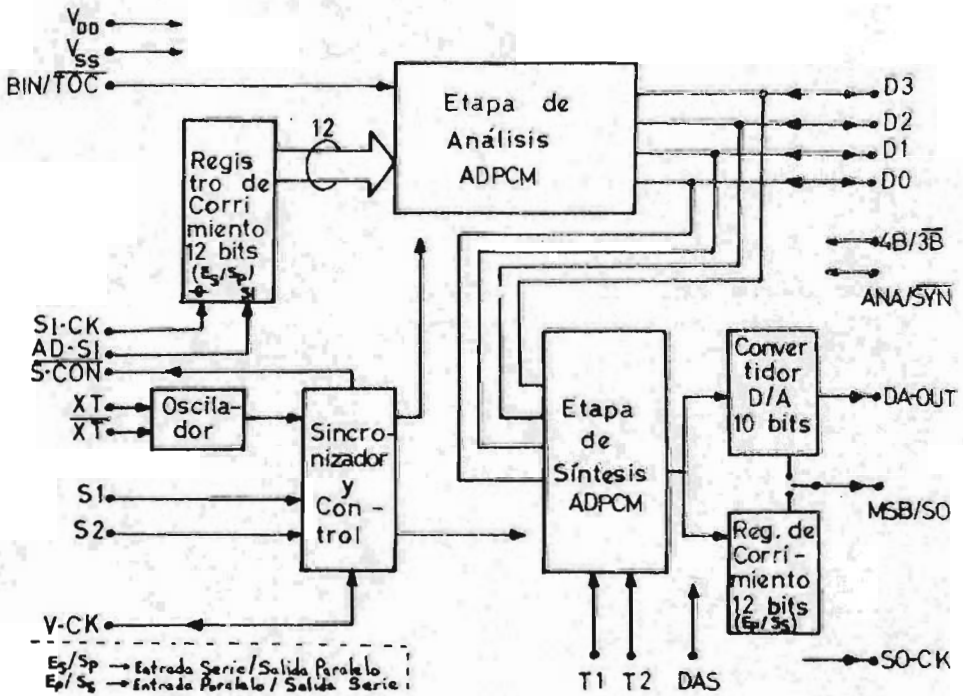


Fig. 6.5.- Diagrama a bloques funcional del MSM 5218 RS.

El 5218 se compone también de un registro de corrimiento de 12 bits, con entrada en serie y salida en paralelo. A

este registro le llegan los 8 bits PCM en serie del 4014 por medio de la pata AD-SI (11) y los 4 bits restantes desde el contador 4024 por medio de la pata SI-CK (10) y salen los 12 bits en paralelo a la etapa de análisis.

La etapa de análisis ADPCM comprime un dato de 12 bits PCM en un dato de 3 o 4 bits ADPCM paralelo los cuales pueden salir directamente por las patas $D_0 - D_3$ (2 - 5), y también estos 4 bits ADPCM pueden ser enviados a la etapa de síntesis. Por medio de la pata $\overline{\text{BIN/TOC}}$ se le especifica a la etapa de análisis si los datos PCM que recibe están en forma binaria o en forma de dos complementos.

La etapa de síntesis, sintetiza datos PCM de datos ADPCM. De esta etapa salen los datos en forma PCM cuando la pata DAS (15) está en nivel alto entonces, los datos pasan por el registro de corrimiento de 12 bits, el cual recibe los datos en paralelo y los saca en serie por medio de la pata MSB/SO (19). O bien si se desean los datos en forma analógica (generalmente) la pata DAS estará en nivel bajo y entonces los datos pasan al convertidor D/A de 10 bits interno y salen forma analógica por la pata DA-OUT (18).

En el 5218 existen parámetros seleccionables como se puede apreciar en la figura 6.1, y son:

- Análisis/ $\overline{\text{Síntesis}}$ (6)
- S_1 y S_2 (8) y (9)
- $4B/\overline{3B}$ (7)

Si en la pata $\overline{\text{ANA/SYN}}$ (6) elegimos un nivel alto, entonces el 5218 por medio de su puesto de direcciones de dato E/S, el dato estará saliendo y ocurre simultáneamente el

análisis y la síntesis, con lo cual tenemos salida de datos en $D_0 - D_3$ (forma digital) que van a interfase de la computadora y consecuentemente grabar en memoria, y además tenemos salida en la pata- DA-OUT (forma analógica). Sin embargo, si la pata (6) la ponemos en un nivel bajo, el dato E/S estará entrando es decir, no ocurre el análisis y los datos que vienen de la memoria son pasados a la etapa de síntesis y salen por la pata DA-OUT en forma analógica.

Otro rango seleccionable es el rango de frecuencias de muestreo que puede ser seleccionados con los niveles lógicos en las patas S_1 y S_2 , pudiendose elegir tres rangos de frecuencias internas preestablecidos como se aprecia en la tabla 6.1, o bien suministrarle externamente hasta 384 KHz, y con ésto tendremos que el 5218 controla el estado y transición de sincronización con la señal VCKL. Debido a que la computadora principal o bien cualquier "hardware" externo se sincroniza a si misma con el 5218 por medio de esta señal VCKL.

Es posible seleccionar si la codificación de datos ADPCM es de 3 o 4 bits, ésto se elige con el nivel lógico sobre la línea $4B/\overline{3B}$ (7), pues con un *1* lógico, elegimos una codificación de 4 bits, mientras que con un nivel *0* seleccionamos una codificación de 3 bits ADPCM. Cuando elegimos 3 bits de codificación, la pata D_0 (2) no se usa y D_3 (5) es el bit más significativo.

Con estas variables que se mencionaron tenemos:

- Una función dual del 5218, debido a que la línea de datos $D_0 - D_3$ son bidireccionales, pues usadas como salidas estas patas, los datos analizados ADPCM sirven para almacenamiento. Y usadas como entrada, los datos van a la circuiteria del sintetizador para su reproducción.

- Con la determinación del nivel lógico de S_1 y S_2 y con la elección de $4B/\overline{3B}$, podemos tener una razón de bit variable como:

Con 8KHz y 4 bits de codificación ($S_1 = *1*$, $S_2 = *0*$) la razón de bit será:

$$8\ 000 \times 4 \text{ bits} = 32\ 000 \text{ bps}$$

Pero con un pequeño sacrificio en fidelidad se puede reducir la razón de bit, con una frecuencia de muestreo de 4 KHz y codificarla a 3 bits ADPCM y tendríamos una razón de:

$$4\ 000 \times 3 \text{ bits} = 12\ 000 \text{ bps}$$

Con los parámetros seleccionables del 5218 tenemos algunas opciones de trabajo del sistema y dependerá de las necesidades de trabajo y fidelidad requerida.

6.5.- ALMACENAMIENTO Y REPRODUCCION

Con el circuito de la figura 6.1, tenemos un sistema dual de análisis y síntesis, y cuando este circuito lo integramos a una computadora basada en microprocesador Z-80 se puede analizar y almacenar 10 segundos de voz con 8 KHz de frecuencia de muestreo o 20 segundos a una razón de muestreo de 4 KHz, haciendo uso de la rutina LOAD (algoritmo de la figura 6.6) del programa 6.1 que se enlista aquí.

Con el programa ensamblador para controlar el circuito de síntesis de voz podemos lograr que el algoritmo LOAD se ponga en marcha cuando el reset se pone en un nivel bajo, y entonces en cada transición de nivel negativo de la señal VCLOCK que sale del 5218 se lee un nybble (4 bits ADPCM) y se almacenan en la memoria.

**Prog. 6.i.- Programa ensamblador basado en el microprocesa
dor Z-80 para análisis y síntesis de voz:**

```

00010 ;
00020 ; THESE ROUTINES ARE USED TO ALLOW A Z-80
00030 ; BASED MICRO TO LOAD AND DUMP THE 4 BIT
00040 ; NYBBLES CREATED BY THE OKI ADPCM CHIP.
00050 ;
00060 ; ONE I/O PORT IS USED FOR BOTH INPUT AND
00070 ; OUTPUT. THE BIT MAP FOR THE PORT IS AS
00080 ; FOLLOWS :
00090 ;
00100 ;
00110 ;   B7   B6   B5   B4   B3   B2   B1   B0
00120 ;           #   #   #   #
00130 ;           D3   D2   D1   D0
00140 ;
00150 ;           * CLOCK ON INPUT, RESET
00160 ;           ON OUTPUT
00170 ;
00180 ;   *   *   *   UNL _ED
00190 ;
00200 ;
00210 ; SINCE THE MASTER CLOCK OF THE OKI CHIP IS
00220 ; USED BY THIS PROGRAM, ANY OF THE CLOCK
00230 ; RATYES MAY BE SELECTED WITHOUT MODIFYNG
00240 ; THIS CODE.
00250 ;
00260 ; AT 8KHZ, YOU WILL FILL 4K BYTES PER SECOND
00270 ;
00280 ;
4000 00290 ;   ORG 4000H           ; START AT 4000H ON D/G CPU.
4000 00300 LOAD EQU 0
4000 F5 00310 PUSH AF           ; SAVE ANY REGS
4001 05 00320 PUSH BC           ;   THAT I WILL Clobber
4002 E5 00330 PUSH HL           ;   DURING THE ROUTINE
4003 2641 00340 LD H,41H       ; BUFFER STARTS AT
4005 3E00 00350 LD L,00H       ; 4:00H
4007 06DF 00360 LD B,06FH      ; AND ENDS AT DF00H
00370 ;
00380 ;
4009 3E10 00390 LD A,00H       ; PICK UP 0.
400E D311 00400 OUT (11H),A     ; AND TURN OFF RESET TO ADPCM
00410 ;
400D 0F2740 00420 LOOP1 CALL NIN   ; BIT UP NYBBLE FROM ADPCM
4010 07 00430 RLCA           ; PUT THE
4011 07 00440 RLCA           ; NYBBLE AT THE
4012 07 00450 RLCA           ; TOP OF THE
4013 07 00460 RLCA           ; REGISTER.
4014 4F 00470 LD C,A         ; AND SAVE IN THE C REG.
4015 0F2740 00480 CALL NIN   ; PICK UP ANOTHER NYBBLE
4018 B1 00490 OR C           ; COMBINE WITH SAVED NYBBLE
4019 77 00500 LD D,D1,A     ; AND SAVE IN THE BUFFER
401A 26 00510 INC HL        ; OVER TO NEXT LOCATION
401B 78 00520 LD A,H       ; PICK UP END POINTER
401C 2C 00530 OR D         ; AND OR AT THE END
401D 20EE 00540 JR NZ,LOOP1   ; AND DO ANOTHER

```

```

401F 3EFF      00550      ; YES, PICK UP ALL 1'S
4021 D311      00560      ; AND RESET THE ADDR0M
4023 E1        00570      ; RECOVER THE
4024 C1        00580      ; SAVED REGISTERS
4025 F1        00590      ; WE SCRAMBLED
4026 C9        00600      ; AND RETURN TO THE CALLER.
                00610 ;
                00620 ;
                00630 ;
4027          00640 NIN
4027 DB11      00650 INHIW   IN A,(11H) ; PICK UP INPUT FROM ADDR0M
4029 CB67      00660      BIT 4,A    ; CHECK FOR A 1 ON THE CLOCK
402B 28FA      00670      JR Z,INHIW ; WAIT FOR ONE TO COME
402D DB11      00680 INLOW   IN A,(11H) ; PICK UP INPUT FROM ADDR0M
402F CB67      00690      BIT 4,A    ; CHECK FOR A 0 ON THE CLOCK
4031 28FA      00700      JR Z,INLOW ; WAIT FOR ONE TO COME
                00710 ;
00720 ; WHEN A 1 TO 0 TRANSITION HAS OCCURRED, THE DATA
00730 ; IN THE A REG IS THE NYBBLE WE WANT TO STORE.
00740 ;
4033 E60F      00750      AND OFH   ; MASK OFF THE HI BITS
4035 09        00760      RET       ; AND RETURN TO CALL F9
                00770 ;
                00780 ;
                00790 ;
                00800 ;
                00810 ;
                00820 ;
                00830 ;
4036          00840 DUMP    EQU 8
4036 F5        00850      PUSH AF   SAVE THE REGS
4037 C5        00860      PUSH BC   JUST LIKE
4038 E5        00870      PUSH HL   BEFORE
4039 2641      00880      LD H,41H  SET UP START
4039 2E00      00890      LD L,00H  ADDRESS OF BUFFER
403B 06DF      00900      LD B,0DFH SET UP END TOO
403F 3E00      00910      LD A,00H  PICK UP ALL 0'S
4041 D311      00920      OUT (11H),A AND TURN OFF RESET TO ADDR0M
                00930 ;
                00940 ;
4043 78        00950 LOOP2. LD A,(HL) ; PICK UP STORED DATA BYTE
4044 E6F0      00960      AND OF0H ; ISOLATE THE UPPER NYBBLE
4046 07        00970      RLCA    ; AND SWING
4047 07        00980      RLCA    ; IT TO
4048 07        00990      RLCA    ; THE LOWER
4049 07        01000      RLCA    ; NYBBLE
404A CD6040    01010      CALL NOUT ; WRITE THE NYBBLE TO THE ADDR0M
404B 7E        01020      LD A,(HL) ; PICK UP THE BYTE AGAIN
404E E60F      01030      AND OFH  ; LOWER NYBBLE THIS TIME
4050 CD6040    01040      CALL NOUT ; WRITE IT TO THE ADDR0M
4053 28        01050      INC HL   ; STEP TO NEXT BUFFER POSITION
4054 78        01060      LD A,B   ; PICK UP END OF BUFFER POINTER
4055 BC        01070      CP H    ; ARE WE THERE YET?
4056 20E2      01080      JR NZ,LOOP2 ; NO, DO SOME MORE

```

```

4058 3EFF      01090      LD A,0FFH      ; PICK UP ALL 1'S
405A D311      01100      OUT (11H),A    ; TURN ON RESET TO ADPCM
405C E1        01110      POP HL         ; RECOVER THE
405D C1        01120      POP BC         ; REGS THAT WE
405E F1        01130      POP AF         ; HAVE Clobbered
405F C9        01140      RET           ; AND RETURN TO CALLER
              01150 ;
              01160 ;
              01170 ;
4060          01180      NOUT      EQU 0
4060 4F        01190      LD C,A        ; SAVE A FOR LATER
4061 DB11      01200      OLOW      IN A,(11H)    ; PICK UP ADPCM BYTE
4063 CB67      01210      BIT 4,A      ; TEST FOR 0 BIT
4065 D0FA      01220      JR NZ,LOW     ; WAIT FOR ONE TO COME
4067 DB11      01230      OHIW      IN A,(11H)    ; PICK UP ADPCM BYTE
4069 CB67      01240      BIT 4,A      ; TEST FOR 1 BIT
406B 28FA      01250      JR Z,OHIW     ; WAIT FOR ONE TO COME
              01260 ;
              01270 ; WHEN 0 TO 1 TRANSITION IS DETECTED ,WE PUT OUR DATA
              01280 ; OUT TO THE ADPCM
              01290 ;
406D 79        01300      LD A,C        ; RECOVER THE SAVED DATA
406E D311      01310      OUT (11H),A  ; WRITE TO THE ADPCM
4070 C9        01320      RET           ; AND GO BACK TO CALLER
              01330 ;
              01340 ;
              01350 ;
              01360 ;
              01370 ; END OF LOAD/DUMP ROUTINES
              01380 ;
              01390 ;
0000          01400 ;      END
00000 TOTAL ERRORS

DUMP  4036 00840
INHIW 4027 00650 00670
INLOW 402D 00680 00700
LOAD  4000 00300
LOOP1 400D 00420 00540
LOOP2 4043 00950 01080
NIN   4027 00640 00420 00480
NOUT  4060 01180 01010 01040
OHIW  4067 01230 01250
LOW   4061 01200 01220

```

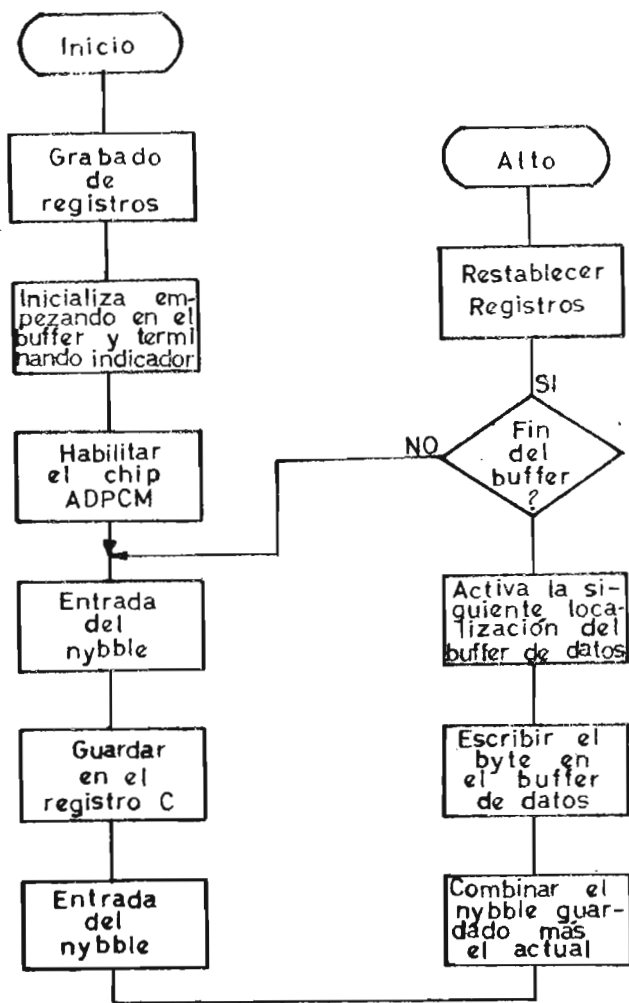
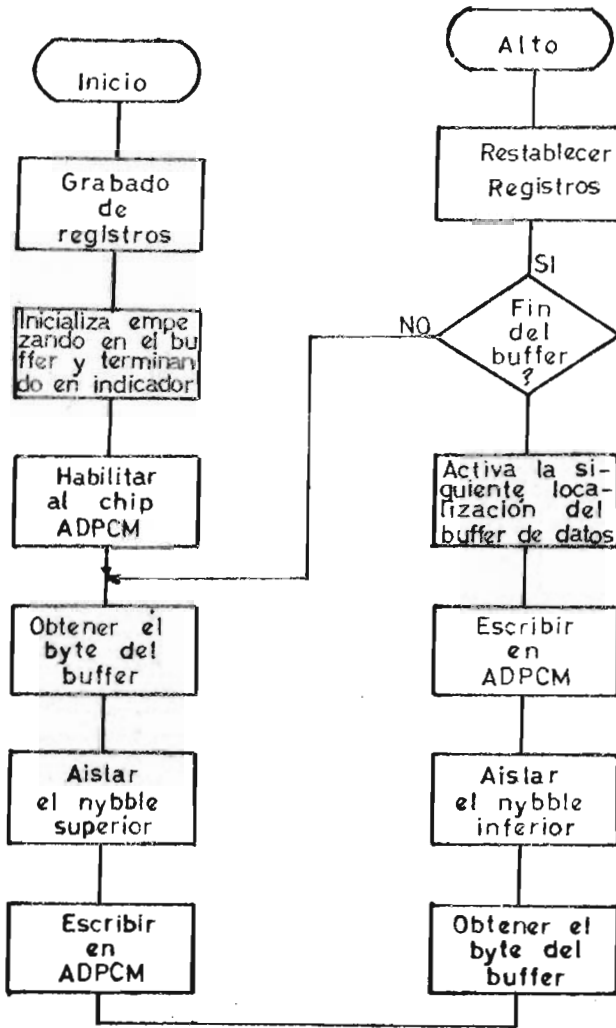


Fig. 6.6.- Algoritmo de la rutina LOAD del programa 6.1, y se usa para codificar la señal de voz y almacenar la en memoria en forma ADPCM.



- Fig. 6.7.- Algoritmo de la rutina DUMP del programa 6.1 y se usa para decodificar datos de voz ADPCM guardados en memoria para reconstruir la forma de onda de voz analógica.

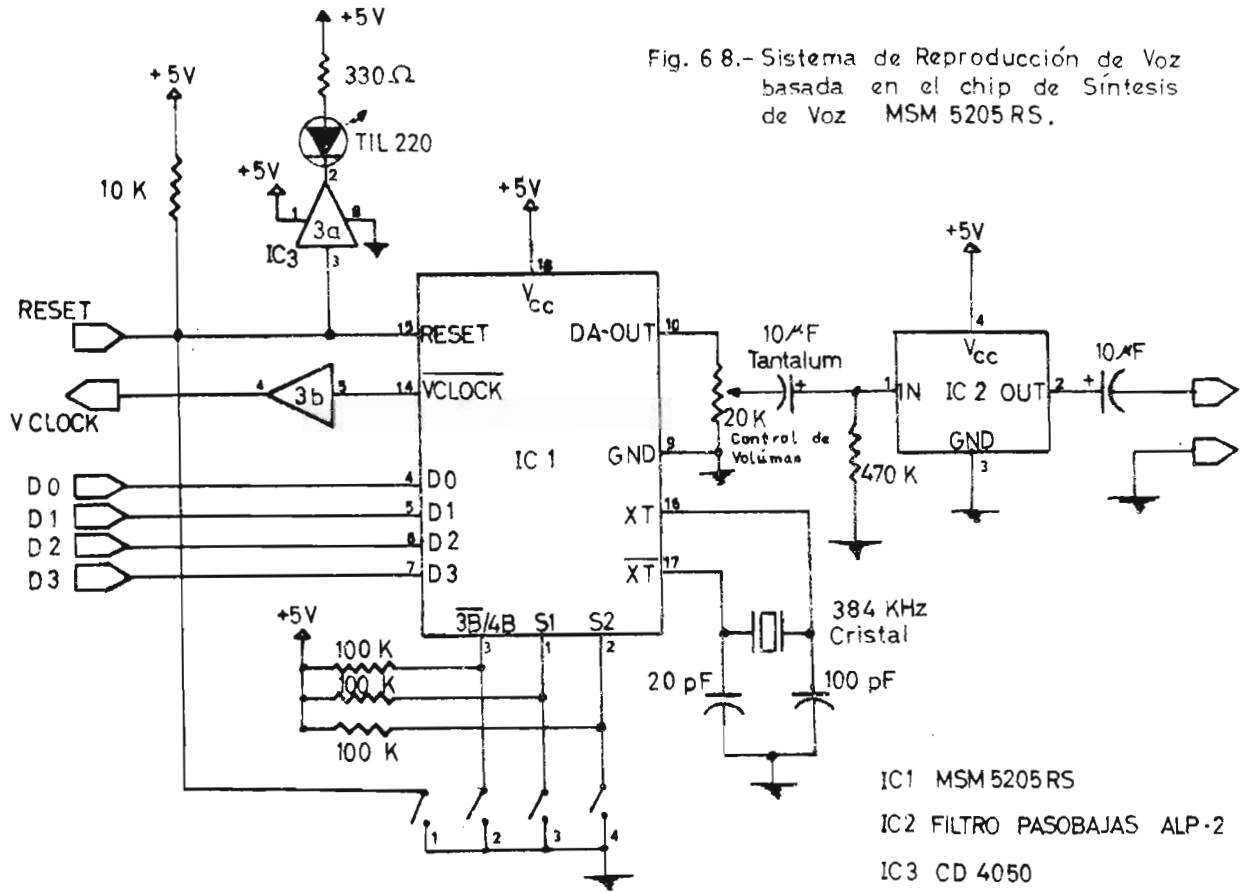


Fig. 6 8.- Sistema de Reproducción de Voz basada en el chip de Síntesis de Voz MSM 5205 RS.

- IC1 MSM 5205 RS
- IC2 FILTRO PASOBAJAS ALP-2
- IC3 CD 4050

Después de que ya ha sido almacenada la voz, si deseamos reproducir lo que hay en memoria entonces podemos sacarla de aquí por medio de la rutina DUMP (algoritmo de la figura 6.7) del mismo programa, entonces los códigos de voz ADPCM son secuencialmente sacados en cada márgen de elevación de la VCLK. Los códigos ADPCM pasan o entran al 5218 cuando se encuentra en el modo de síntesis.

Sin embargo, si se desea tener datos almacenados en forma permanente y reproducirlos de acuerdo a requerimientos como puede ser anuncio, mensajes telefónicos, etc. Ya no será necesaria la etapa de análisis del 5218 y para no cambiar el nivel de la pata ANA/SYN (6) después de haber almacenado los datos en memoria podemos hacer uso del IC MSM 5205 RS.

En la figura 6.8, se muestra un sistema de síntesis de voz basada únicamente en el IC 5205 que como se aprecia, a comparación del sistema de la figura 6.1 es más sencillo y barato pues el 5205 que realiza solamente la síntesis es más simple que el 5218 como se aprecia en el diagrama a bloques funcional de la figura 6.9.

Con la rutina DUMP del programa enlistado, también se puede sacar los datos en memoria e introducirlos al 5205 en lugar del 5218 y sacarlos en forma analógica.

6.6.- RESUMEN

En este capítulo que es la realización práctica de la síntesis de voz mediante ADPCM se vieron dos sistemas. Uno en donde podemos realizar todo el proceso, se puede almacenar en memoria datos de voz y después reproducirla, éste basado en el IC MSM 5218 RS de la figura 6.1.

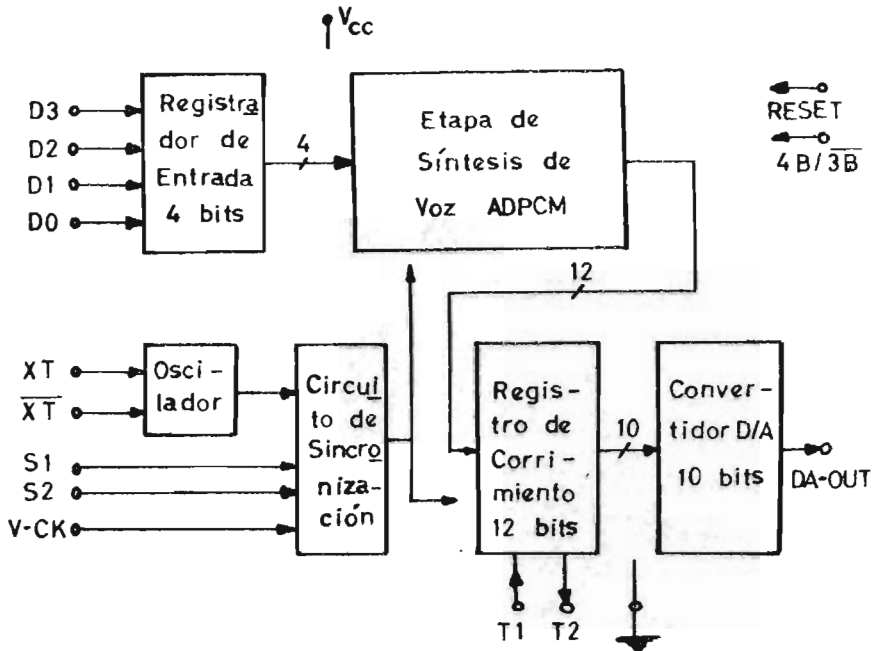


Fig. 6.9.- Diagrama a bloques funcional del MSM 5205 RS.

Otro sistema que es más simple y que realiza únicamente la síntesis, basado en el IC MSM 5205 RS de la figura 6.8, en el cual ya se tiene una memoria grabada y servirá únicamente para reproducción.

Sin embargo, teniendo acceso a los dos sistemas tendremos la posibilidad de poder manufacturar o producir un vocabulario o mensaje cuantas veces se desee, conservando la inteligibilidad y el acento regional del orador en memoria. Y después decodificar la voz que se encuentra en memoria dentro del simple circuito de síntesis de voz de la figura 6.8 para su salida en el lugar que se necesite.

C
O
N
C
L
U
S
I
O
N
E
S

CONCLUSIONES

La importancia de la comunicación hablada se ha incrementado en años recientes con el interés que ahora se muestra en la comunicación humana con las máquinas; ésto ha estimulado la investigación de la generación de señales de voz por medios artificiales que es conocida como SINTESIS DE VOZ

La síntesis de voz debe desarrollarse para que pueda haber avance y mayor utilización en una forma más general de algunas de las aplicaciones del procesamiento digital de la voz como lo es el reconocimiento de voz, la verificación e identificación de una persona, en la educación o bien en la ayuda a impedidos. Debido a que los sistemas de reconocimiento, de verificación e identificación hacen uso de la síntesis. Con lo que respecta al uso pedagógico, a la ayuda a impedidos o bien simplemente para una mayor facilidad e interacción hombre-máquina, la síntesis de voz llegará a ser en un futuro no muy lejano algo tan común y necesario como lo está siendo actualmente la computadora.

La inquietud por el desarrollo de los sistemas de comunicaciones y el gran auge que ha tenido en las últimas décadas la tecnología de los circuitos integrados han permitido grandes progresos en las técnicas de procesamiento digital de señales, con el consiguiente progreso de las computadoras y con ésto el deseo de llegar a tener comunicación con ellas de una forma tan fácil como se hace con cualquier persona.

Por tal razón las técnicas de codificación de voz desde la década de los sesentas que inicia con el PCM se han ido incrementando en forma extraordinaria, buscando la técnica por medio de la cual se tenga una buena calidad, una baja razón de bit y un bajo costo. Sin embargo, la mayoría de los

sistemas de comunicación mundial están basados en el PCM, y hacer grandes cambios en los sistemas sería muy costoso. Con lo cual la técnica ADPCM es la alternativa más viable para lograr una reducción en la razón de bit manteniendo una buena calidad de voz, al ser ésta una forma especializada del PCM.

La técnica ADPCM ofrece una razón de bit mucho más baja que la PCM con una calidad y una inteligibilidad significativamente alta, y el uso de los canales de comunicación puede ser más efectivo ahora que éstos son altamente requeridos por el aumento de los usuarios de los sistemas de comunicaciones, y por el aumento progresivo que hacen las industrias de todo tipo de los canales para la transmisión de datos en banda de voz vía los satélites o bien las líneas terrenas.

Una prueba del uso como alternativa del ADPCM es el algoritmo realizado por la CCITT mostrado en el resumen del capítulo 4.

Sin embargo, la técnica de procesamiento digital de señales a elegir dependerá a la calidad de voz deseada, al uso que se le dará al sistema, a la complejidad y al costo. En la figura 1 se dan algunas técnicas, sus razones de bit y la calidad obtenida de ellas.

Las técnicas de compresión de datos sirven para ahorrar espacio de almacenamiento y tiempo de transmisión, debido a que compactan un código como lo realiza el ADPCM al transmitir o almacenar la misma información con 3 o 4 bits que con un código PCM de 8 bits.

El desarrollo y diseño de los sistemas de procesamiento de voz al igual que cualquier tipo de sistema se debe a que en la tecnología de los circuitos integrados se ha logrado

un sorprendente avance, lograndose una alta velocidad de procesamiento, chips relativamente pequeños y compatibilidad de tecnologías. Con ésto es posible tener ahora en CI's memorias, microprocesadores, microcontroladores, dispositivos periféricos, dispositivos para telecomunicaciones, procesadores de voz, etc.

Como ejemplos del avance de la tecnología son los CI de procesamiento de voz mediante ADPCM MSM 5218 RS y MSM 5205 RS, con los cuales se mostraron dos sistemas diferentes en el capítulo 6.

Con el 5218 se forma un sistema de grabación y reproducción de voz mediante ADPCM que en diagrama a bloques el sistema puede mostrarse como se ve en la figura 2. Con el 5205 se vió un sistema que solamente es reproductor de voz y en diagrama a bloques se muestra en la figura 3.

El uso de estos dos sistemas puede ser diverso y quedará a la imaginación del usuario. Una de las aplicaciones puede ser como un recurso pedagógico en la instituciones educativas, pues con el primer sistema se puede almacenar información de diferentes áreas en discos para que después los estudiantes tengan acceso a ellos y puedan reproducirlos en cubículos, con el segundo sistema.

Sin embargo, cualquier industria puede hacer y acondicionar estos sistemas a sus necesidades, pues ellos dan la facilidad de manufacturar cuantas veces sea necesario un vocabulario, mensaje u ordenes, así como reproducirlos. Pudiendose realizar recepcionistas telefónicas automáticas, sistemas de despacho (en centrales camineras, aeropuertos, correo, tiendas, etc), publicidad, y otros.

La inquietud por el desarrollo del procesamiento de voz es muy amplio, y la tecnología de la voz digitalizada seguirá en aumento pues al encontrarnos de lleno ya dentro de la era de las computadoras se hace necesario que la comunicación con ellas, cualquiera que sea su tamaño sea fácil y accesible para todas las personas y la forma más simple es mediante la palabra hablada.

Por todo esto, y debido a los avances tecnológicos en el campo de las telecomunicaciones, en la informática, en la electrónica, se hace necesario que en las carreras de ingeniería se actualicen los planes de estudio para incluir materias que proporcionen conocimientos y actualicen información sobre temas de procesamiento de señales, sistemas de transmisión, sistemas y medios de comunicación, etc. Pues al no poseer éstos, el país y los profesionistas de estas áreas seguirán siendo dependientes tecnológicamente de otros países.

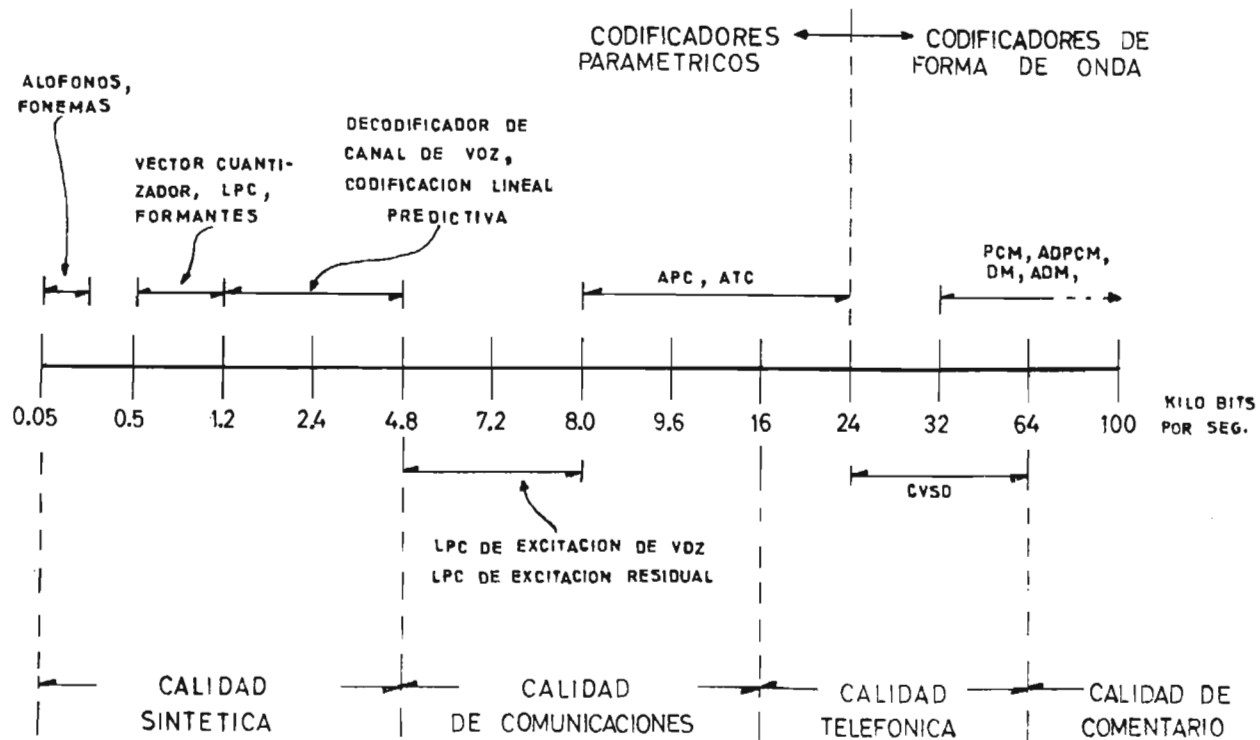


Fig.1.- Técnicas para Procesamiento Digital de Señales.

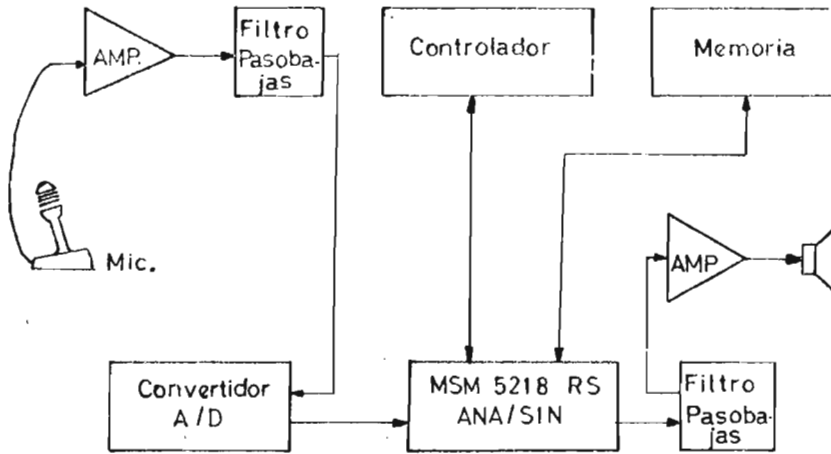


Fig. 2.- Grabador y Reproductor de Voz .

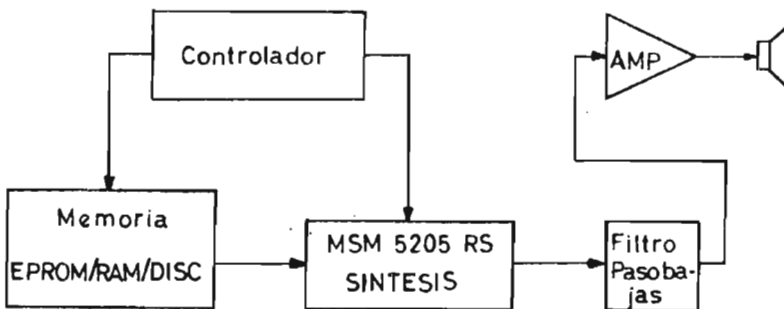


Fig.3.- Reproductor de Voz .

A
P
E
N
D
I
C
E

A

APENDICE A

El uso de algunos términos y expresiones técnicas hacen la necesidad de la elaboración de un glosario de términos y así evitar confusiones.

ACCESO - ACCESAR

Accesar es almacenar, modificar o recuperar información en algún dispositivo o archivo o bien hacer uso de él.

Acceso es la acción de acceder. Se dice que se tiene acceso a un dispositivo si es posible utilizarlo en alguna labor, o si se puede leer o escribir alguna información en él.

ALGORITMO

Método para resolver un problema, un algoritmo es un conjunto de acciones que deben ejecutarse en un orden específico.

AMPLIFICADOR OPERACIONAL

El OPAM es un amplificador cuya función de transferencia está determinada por la entrada externa y el circuito de realimentación.

ANALISIS ESPECTRAL

Hay muchas formas de analizar la voz, las cuales están cercanamente relacionadas al análisis espectral, las cuales logran y muestran sus resultados por medios diferentes.

/ *TRANSFORMADA DE FOURIER*-

La transformada de Fourier es una forma de realizar el análisis espectral de la voz. El cálculo de la transformada es esencialmente un proceso discreto, dando el resultado correspondiente a un segmento de tiempo particular. ASin embargo, es posible, obtener un análisis sucesivo de la señal, tan cercana en tiempo como se desee.

-* AUTOCORRELACION*-

La función de autocorrelación de un segmento de señal es la transformada de Fourier de su potencia espectral, y básicamente contiene la misma información del espectro del mismo segmento.

-*CEPSTRUM*-

El cepstrum es un método importante de análisis de voz la cual es definida como la transformada de Fourier del logaritmo de la potencia espectral. En la producción de voz, la fuente de sonido con su propia estructura espectral está modificada por el efecto formado del espectro del tracto vocal; cada uno de los componentes espectrales de la fuente está multiplicada por la ganancia del tracto vocal.

En el análisis cepstrum, por la ley de los logaritmos convierte esta multiplicación a una suma lineal, así se permite una filtración lineal para separar las propiedades espectrales de la fuente y la respuesta del tracto vocal.

-*CEPSTRAL*-

Se le llama cepstral al procedimiento de suavisar o alisar la función espectral que se obtiene del cepstrum.

Todas las formas de análisis aquí mostradas pueden ser desarrolladas sobre diferentes medidas de segmentos de tiempo. Todas ellas dependen sobre la teoría de la Transformada de Fourier.

ANCHO DE BANDA

Capacidad de transmisión de un canal, el ancho de banda indica la cantidad de información por unidad de tiempo que puede llevar una línea de transmisión.

BINARIO

Dos; binario es referirse a un sistema de numeración de base dos en el que sólo existen dos dígitos, "1" y "0" y es ampliamente utilizado en los computadores por la facilidad para formar circuitos capaces de asumir dos estados.

BIT

Bit es la contracción de binary unite (unidad binaria). Es el componente más pequeño de la clave binaria, un bit es un solo dígito ("0" o "1") en un número binario. Físicamente un bit se caracteriza por una celda de almacenamiento un pulso, un punto magnético, o una burbuja magnética.

BITS POR SEGUNDO

Los bits por segundo se emplea como unidad para medir la velocidad de transferencia de información.

BUFFER

Area de almacenamiento que conserva información temporalmente, los buffers en un programa, son áreas reservadas

que reciben información y la conservan para su procesamiento.

BUS

Ducto, ruta o canal común entre dispositivos del hardware en una computadora.

BYTE

Unidad de almacenamiento equivalente a 8 bits o a un carácter de información. El byte es una unidad común de almacenamiento en un sistema de computo y es sinónimo de carácter de datos o de texto. Los bytes se emplean para hacer referencia al tamaño del hardware, el software o la información.

CCITT

Comité Consultor de Telefonía y Telegrafía Internacional, la CCITT es uno de los cuatro órganos permanentes de la ITU (Unión Internacional de Telecomunicaciones). La CCITT fija las normas internacionales para las comunicaciones.

CHIP

Circuitos electrónicos miniaturizados que contienen desde unos cientos hasta varios cientos de miles de componentes electrónicos (transistores, resistencias, etc.), sinónimo de circuito integrado y pastilla microelectrónica.

CODIFICACION DE PREDICCIÓN LINEAL (LPC)

Un esquema que codifica parámetros de un modelo matemá

tico, el cual define la forma del tracto vocal a intervalos de análisis fijos.

CODIFICADOR

El codificador es un dispositivo para cambiar una cantidad analógica dentro de una representación digital equivalente.

CODIGO

Se define un código como la correspondencia de todas las secuencias posibles de símbolos de un alfabeto dado a una secuencia de símbolos de algún otro alfabeto.

COMPARADOR

Un circuito para comparar magnitudes de dos voltajes y generar un voltaje de salida de dos niveles dependientes del voltaje más largo o más grande.

COMPRESION DE DATOS

La compresión de datos es compactar un código, generalmente binario; la compresión se utiliza para ahorrar espacio de almacenamiento y tiempo de transmisión. Se usan varias técnicas, pero el principio siempre es el mismo; Se analizan los datos contenidos en un grupo de bits (por ejemplo PCM) y se codifican de diferente manera para representarlos mediante un número de bits menor (ADPCM). Cuando los datos se procesan, deben ser decodificados para llevarlos a su formato original.

COMPUTADORA ANALOGICA

El término analógico implica señales continuas. La computadora analógica es aquella que acepta y procesa señales análogas a las del mundo real; Las entradas y salidas de estas computadoras son señales continuas, tales como las fluctuaciones de voltajes.

COMPUTADORA DIGITAL

El término digital implica señales discretas. La computadora digital, sólo puede aceptar información en código binario. Sin embargo, al volverse éstas más veloces y accesibles, están reemplazando las funciones que realizaban las computadoras analógicas haciendo uso del procesamiento digital de señales.

COMPUTADORA DE PROPOSITO GENERAL

La versatilidad y flexibilidad de una computadora digital estandar se basa en el hecho de que es una computadora de propósito general, que resuelve un problema específico al seguir un conjunto de instrucciones (programa). La mayoría de las computadoras digitales son de propósito general, aún cuando existan algunas especializadas en aplicaciones específicas.

CONCATENACION

Significa encadenamiento. La concatenación encadena estructuras de datos.

CONVERTIDOR ANALOGICO A DIGITAL

Algunas veces usado en intercambio con el codificador.

El significado preferido es un sistema completo de conversión el cual incluye un codificador, almacenaje de palabras digitales, fuentes de voltaje de referencia, etc.

CONVERTIDOR DIGITAL A ANALOGICO

El convertidor D/A es algunas veces usado en intercambio con el decodificador. El significado preferido es un sistema de conversión que incluye el decodificador, almacenaje de la palabra digital, fuente de voltaje de referencia, amplificador de salida, etc.

CUANTUM

El cuantum está definido como el valor analógico equivalente a la cantidad o incremento más pequeño que puede ser resuelto o representado por una palabra digital. Es igual al valor del bit menos significativo de la palabra digital.

DATO

Los datos son unidades de información que pueden definirse con precisión, desde el punto de vista técnico los datos son la materia prima que al ser procesada dan lugar a la información.

DECODIFICADOR

El decodificador es un dispositivo para cambiar una palabra digital, representando la magnitud de alguna cantidad, dentro de su valor equivalente analógico.

DIGITIZAR O DIGITALIZAR

Transformar a clave binaria (digital), casi siempre el término digitizar se refiere a la conversión de una señal analógica como la voz o la imagen a una clave binaria.

DUMP (VACIADO)

Copiar a otro medio de almacenamiento o imprimir tal y como está. De hecho el término vaciado se refiere a una copia sin modificación de datos o instrucciones.

ERROR DE CUANTIZACION

Es el error en la representación digital de una cantidad. El error de cuantización mínimo es cero, es decir, el valor analógico es exactamente el valor equivalente de la palabra digital. El error de cuantización máximo es \pm un medio del nivel cuántico, de donde el valor analógico es la mitad entre una palabra digital y otra.

FONEMA

Una del conjunto de las unidades más pequeñas de voz que sirve para distinguir una articulación o expresión de otra en un lenguaje o dialecto.

FORMANT

Las resonancias del tracto vical que causan concentraciones de energía a ciertas frecuencias son conocidas como formants, y se presentan como picos cuando se realiza el análisis del espectro en frecuencias en la voz. Para la mayoría de la voz masculina, el primer formant está situado en alguna

parte entre los 200 y 800 Hz, el segundo se encuentra entre los 800 y 2 300 Hz, y el tercero estará entre los 2 300 y 3 000 Hz. La frecuencia resonante produce una relación directa a la forma del tracto vocal. Los formants están combinados a través de la suma de la filtración y la señal, para crear fonemas.

HARDWARE

Todos los equipos y componentes físicos de un sistema de cómputo, como unidades de entrada, unidad central de procesamiento CPU y unidades de salida, así como dispositivos externos de almacenamiento.

LOAD (CARGA)

Significa hacer una transferencia a un dispositivo de cómputo. Cargar el programa, significa copiar el programa en la memoria, cargar el disco significa escribir datos en el disco.

MODULACION

Mezcla de una señal con una portadora, la modulación es el proceso de entremezclar una señal de voz o una serie de datos con una portadora para su transmisión a través de una red.

MODULACION DELTA

Un esquema de codificación de la forma de onda que codifica la diferencia en magnitud de las muestras adyacentes de la señal.

MODULACION DELTA ADAPTIVO

Un algoritmo que usa un integrador en la malla de realimentación para compensar las no linealidades del modulador delta.

MODULACION POR CODIFICACION DE PULSOS PCM

Un esquema de codificación de la forma de onda, que digitaliza muestras de la forma de onda analógica. Entonces usa éstas para *modular* un flujo de unos y ceros para que proporcionalmente espaciados representen la amplitud de la muestra.

MODULACION POR CODIFICACION DE PULSOS DIFERENCIAL ADAPTIVO ADPCM

Un algoritmo que usa la muestra de voz actual para predecir el valor de la siguiente muestra; la diferencia muestra a muestra es usada con cuantización adaptiva, el cual explota la información acerca de la amplitud de la señal.

MUESTREO Y RETENCION

Circuitería para muestrear el valor de un voltaje analógico para un periodo de tiempo dado, y después que el muestreo está completo, se almacena el voltaje analógico para el tiempo de retención requerido. Puede ser usado para reducir el tiempo de apretura de los convertidores A/D.

NYBBLE

Se le llama nybble a la representación de la palabra digital de cuatro bits. En un byte hay dos nybbles.

PALABRA DIGITAL PARALELO

Una palabra digital con todos los bits existiendo simultáneamente sobre líneas separadas.

PALABRA DIGITAL SERIE

Una secuencia sincronizada de unos y ceros sobre una simple línea representando una palabra digital.

PROCESO - PROCESAR

Manipular información; el término procesamiento indica cualquier acción que se realiza con la información. El procesamiento puede modificar el contenido o la forma de la información o sólo puede transferirla de una fuente a otra sin cambiarla. El término procesamiento puede significar:

- *- Cualquier trabajo realizado por el sistema de cómputo.
- *- Sólo el trabajo realizado internamente con la información.

PROCESAMIENTO DE DATOS

Captura, almacenamiento, manejo y recuperación de datos e información; el procesamiento de datos se refiere al empleo de los datos.

PROCESAMIENTO DIGITAL DE SEÑALES PDS

Procesamiento en forma digital de señales del mundo real el PDS es una categoría de técnicas que se aplican al análisis de formas y señales del mundo real, como la voz e imágenes. Por principio el PDS convierte los datos de los objetos que se van a analizar en un formato digital para procesarlos utilizando diversos métodos matemáticos. Cuando una señal ha

sido reducida a una serie de números, sus componentes pueden aislarse y analizarse con mayor facilidad que en una analógica.

PROGRAMA

El programa consiste en una o más rutinas que dirigen la solución de un problema completo.

RAZON DE BITS

La razón de bits en la conversión es el número de bits en la palabra digital resultante, codificada por unidad de tiempo.

RESOLUCION

Para palabras digitales, el número de valores discretos que pueden ser representados por la palabra digital de cero a la escala completa. Una palabra binaria de 10 bits tiene una resolución de una parte en 1 024.

RUTINA

Es un conjunto de instrucciones que dirige la realización de una tarea específica.

SIMULACION

- * Emulación.- La simulación la realiza un programa que reduce el lenguaje de un programa extraño al lenguaje de máquina de la computadora que se está empleando.
- * Planteamiento de un modelo científico.- La simulación es también una técnica científica para simular matemáticamente

el comportamiento del mundo real.

SINTESIS DE VOZ

Combinación o arreglo de componentes elementales de la voz humana (fonemas), transformados en señales digitales especiales, para formar palabras y frases. La síntesis de voz sirve para dar respuesta sonora a las máquinas con lo que existe una mayor interacción hombre-máquina.

SOFTWARE

El software son los programas para realizar aplicaciones específicas que contienen las instrucciones para dirigir las operaciones del equipo de cómputo.

SONIDO DE VOZ

Un sonido producido cuando el aire pasa lentamente a través de la laringe. El tracto vocal, de la faringe a los labios actúan como una cavidad resonante, amplificando ciertas frecuencias y atenuando otras.

VENTANA

La ventana es una secuencia la cual determina la porción de la señal de entrada que recibe énfasis en un tiempo señalado en particular.

A

P

E

N

D

I

C

E

B

MSM5204RS 8 BIT ANALOG-TO-DIGITAL CONVERTER

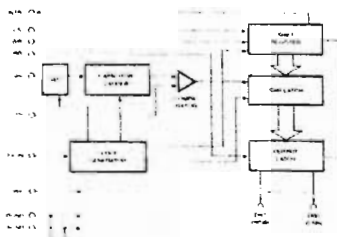
GENERAL DESCRIPTION

The MSM5204RS is a CMOS 8 bit analog-to-digital converter which requires a minimum of external ladder components system. A built-in sample and hold function enables direct input of analog signals. To illustrate analog-to-digital conversion in a designed-in or stand-alone system, the additional two power consumption modes for MSM5204R are available for portable battery operated systems.

FEATURES

- 8 bit resolution
- $\pm 1/2$ LSB accuracy
- Typical conversion time of 100 ns at 400 kHz clock rate
- Tri-state outputs
- Low power consumption (4.5 mW typical)
- Tri-state compatible
- On-chip sample & hold circuit

FUNCTIONAL BLOCK DIAGRAM



PIN CONFIGURATION



MSM5204RS 8 BIT ANALOG-TO-DIGITAL CONVERTER

PIN DESCRIPTION

V+	Reference supply input. Nominality - 5V	RD	Read output. Falling edge of RD enables the 8-bit converted data. In tri-state mode, RD enables the 8-bit converted data to be placed in tri-state.
V-	Analog signal input	WR	Write output. Falling edge of WR enables the 8-bit converted data to be placed in tri-state.
AGND	Analog ground. For use with V+ and V- inputs	CS	Chip select input. Falling edge of CS enables the 8-bit converted data to be placed in tri-state.
DB[0:7]	8-bit digital outputs. Each 8-bit converted data can be read	CLKIN	Clock input
D[0:7]	Digital feedback. Must be at the same potential as AGND	VDD	Power supply pin. Nominality - 5V
INTR	An active-low signal which signals end of conversion		

ABSOLUTE MAXIMUM RATINGS

Parameter	Symbol	Conditions	Rating	Unit	Notes
Supply voltage	V _{DD}	T _a = 25°C	0.3 to +7	V	Stresses above those listed under Absolute Maximum Ratings may cause permanent damage to the device. This is a stress rating only and functional operation of the device at these or any other conditions above those indicated in the operational specifications of this specification is not recommended. Exposure to maximum rating conditions for extended periods may affect device reliability.
Input voltage	V _{in}	T _a = 25°C CS, RD, WR pins	0.3 to +15	V	
Input voltage	V _{in}	T _a = 25°C	0.5 to V _{DD} + 0.3	V	
Storage temperature	T _{stg}		-55 to +150	°C	

OPERATING CONDITIONS

Parameter	Symbol	Condition	Rating	Unit
Supply voltage	V _{DD}		4.5 to 5.5	V
Operating temperature	T _{op}		40 to +85	°C

D.C. CHARACTERISTICS

(V_{CC} = 5V ± 10%, T_a = 40°C to +85°C, unless otherwise noted)

Parameter	Symbol	Condition	Limits			Unit
			Min.	Typ.	Max.	
Input High Voltage	V _{ih}	V _{DD} - 5.5V CS, RD, WR pins only	3.6		15	V
Input High Voltage	V _{ih}	V _{DD} - 5.5V CLKIN pin only	3.6		V _{DD} + 0.3	V

D.C. CHARACTERISTICS Continued

Parameter	Symbol	Condition	Limits			Unit
			Min.	Typ.	Max.	
Input Low Voltage	V _{IL}	V _{DD} - 4.5V			0.8	V
Input High Current	I _{ih}	V _{ih} = 5.5V			1	μA
		V _{ih} = 5.0V				
Input Low Current	I _{il}	V _{ih} = 5.5V			1	μA
		V _{ih} = 5.0V				
Output High Voltage	V _{oh}	V _{oh} = 4.5V	2.4			V
		I _{oh} = 400μA				
Output Low Voltage	V _{ol}	V _{ih} = 4.5V I _{oh} = 1.6mA			0.4	V
Output Leakage Current	I _{oh}	V _{ih} = 5.5V			3	μA
		V _{ih} = 5V DBU = DB7				
Output Leakage Current	I _{ol}	V _{ih} = 5.5V			3	μA
		V _{ih} = 5V DBU = DB7				
Output (Signal) Circuit Low Current	I _{oh}	V _{oh} = 3V			8	mA
		V _{oh} = 0V				
Output (Signal) Circuit High Current	I _{oh}	V _{oh} = 5V			18	mA
		V _{oh} = 5V				
Supply Current	I _{DD}	V _{ih} = 5V f _{CLK} = 500KHz		0.9	3	mA

A.C. CHARACTERISTICS

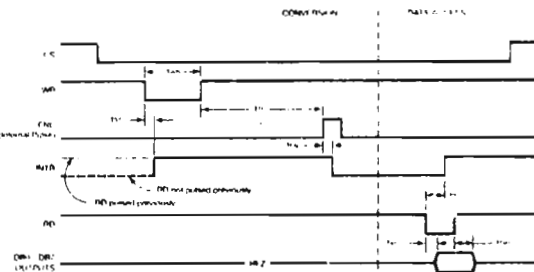
(V_{ih} = 5V ± 10%, f_{CLK} = 40°C to +85°C)

Parameter	Symbol	Condition	Limits			Unit
			Min.	Typ.	Max.	
Max. clock frequency	f _{CLK}		500	700		KHz
Aperture time	T _A	f _{CLK} = 500 KHz		9		ns
Total conversion time	T _{TC}	CS = 0V		98		μs
		f _{CLK} = 500 KHz				
WFI pulse width	t _{WFI}	CS = 0V	200			ns
INTR Delay time	t _{INT}			360		ns
		f _{CLK}		100		ns
INTR Delay time	t _{INT}			280		ns
		f _{CLK}		200		ns
Data access time	t _{DA}	CS = 0V		200		ns
		CS = 0V		200		ns

CONVERTER CHARACTERISTICS

(V_{ih} = 10%, f_{CLK} = 40°C to +85°C)

Parameter	Symbol	Condition	Rating			Unit
			Min.	Typ.	Max.	
Conversion accuracy						LSB
Max input resistance						Ω
Max input voltage						V
Max input voltage						V
Max input voltage						V
Power supply sensitivity		V _{DD} = 5.0V ± 10% V _{SS} = 0V		± 1b	± 1b	LSB



650 North Mary Ave., Sunnyvale, CA 94088
Tel (408) 720-1500

OKI SEMICONDUCTOR, INC.

TELEPHONE (408) 984-4842 TELEX (25) 910-3380508

CLARA, CALIF 95051

OKI Semiconductor reserves the right to make changes in specific areas at any time and without notice. The information furnished by OKI Semiconductor in this publication is believed to be accurate and reliable. However, no responsibility is assumed by OKI Semiconductor for its use for any infringement of patents or other rights of third parties resulting from its use. No license is granted under any patents or patent rights of OKI.

© 1982 OKI SEMICONDUCTOR, INC.

PRINTED IN U.S.A.

MSM5218RS ADPCM SPEECH ANALYSIS/SYNTHESIS IC

GENERAL DESCRIPTION

The MSM5218RS is a complete speech analysis/synthesis LSI featuring the Adaptive Filterbank Pulse Code Modulation (ADPCM) method of data compression. The MSM5218RS contains an analysis stage where serial PCM data is compressed to 2, or 4 bit serial ADPCM data. In addition, a synthesis stage synthesizes PCM data from ADPCM data. This PCM data can be output directly or routed to an external 80-bit DAC for analog signal output.

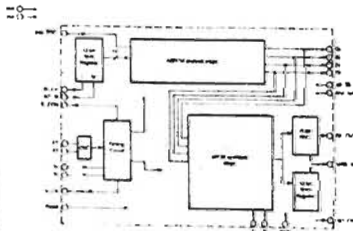
An advanced sampling speech analysis and synthesis system control enables users to develop their own speech analysis and synthesis systems.

The MSM5218RS is fabricated using OKI's advanced CMOS process for low power consumption and is housed in a 24 pin standard DIP for versatility.

FEATURES

- On-chip speech analyzer/synthesizer
- 2, or 4 bit ADPCM system
- ADPCM data compatible with OKI's synthesis LSI MSM5204RS
- Single power supply
- Variable sampling frequency (4 kHz, 8 kHz, & user's external)
- Low power consumption CMOS process (15 mW typical)
- Built-in 80-bit D/A converter for analog output
- Handshaking signals provided for synchronous operation with an external A/D converter

FUNCTIONAL BLOCK DIAGRAM



PIN CONFIGURATION



PIN DESCRIPTION

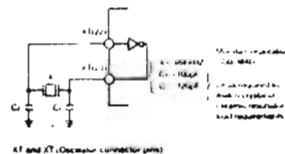
V _{CC}	Pin 1 (Dip) - This pin outputs a signal whose frequency is lower to the sampling frequency specified by S1, S2 inputs. Note: The selected sampling frequency is not to be applied through this pin when the external mode is selected via the S1, S2 inputs.	DAS	Serial PCM data output (pins 21-24) serial PCM data output (DAS - 4H)
D0-D3	Data 0-3 port for the ADPCM data. For 2-bit ADPCM mode, D0 input is not used. (S1 - MSB)	D11-D12	Serial PCM data output (pins 21-24) serial PCM data output (DAS - 4H)
Anal. Syn	Analyze/synthesis function selector. Conditions data 0-3 port direction: When high - data 0-3 are outputs and serial PCM analysis and synthesis indicator. When low - data 0-3 are inputs and no analysis indicator.	DAOUT1	Analog digital output pin
4B, 5B	Specifies whether 3 bit or 4 bit ADPCM data is to be used. (High - 4 bit)	MSB, MS	MSB serial data output pin. MSB is the data to the external 80-bit DAC with 4-bit and 8-bit serial PCM data output (DAS - 4H) serial PCM data output (DAS - 4H) serial PCM data output (DAS - 4H) serial PCM data output (DAS - 4H)
S1, S2	These inputs select the sampling frequency of the external mode according to figure 1.	80-BIT DAC	80-bit DAC. The serial PCM data is output to the external 80-bit DAC.
D1-D4	Data output for receiving or serial PCM data from an external A/D converter to the internal 12 bit shift register.	RES1	Reset pin. High input pulse indicates to the internal 80-bit DAC internal circuitry. To be effective must be held low for at least 1.0 μ s.
AD, SI	Serial PCM data input	XT, XT	Crystal input pins for a 344 kHz crystal or ceramic resonator (4 pins 2).
VSS	Ground (0V)	VDD	Power supply pin (approximately -1V)
S/CLOCK	Output which signals the start of a new word.		
S/CLOCK	When serial PCM data output mode is selected (DAS - H) this pin provides a 162 kHz signal which is synchronized with the output of the serial PCM data through the MSB, S0 pin. Each bit of the 12 bit PCM data will be valid before the positive edge of the 162 kHz signal.		

FIGURE 1

S1	S2	Sampling Frequency**
L	L	4 kHz
L	H	8 kHz
H	L	8 kHz
H	H	External**

Note: The 344 kHz oscillator that be used whether a 344 kHz, 8 kHz or external mode is selected.
 *External mode - Frequency may be from 0.1 to 10 kHz.
 **With 344 kHz oscillator. Other operating frequencies are possible and may proportionately modify the sample rate.

FIGURE 2



ABSOLUTE MAXIMUM RATINGS

Parameter	Symbol	Conditions	Rated value	Unit
Supply Voltage	V _{DD}	T _a 25°C	0.3 to 7	V
Input Voltage	V _{in}	T _a 25°C	0.3 to V _{DD}	V
Power Dissipation	P _{tot}	T _a 25°C	200 max	mW
Storage temperature	T _{stg}		-55 to +150	°C

Note: Stresses above those listed under ABSOLUTE MAXIMUM RATINGS may cause permanent damage to the device. This is a stress rating only and normal operation of the device at these or any other conditions above those specified in the operational sections of this specification is not implied. It is user's responsibility to determine the reliability of the device for the intended application.

OPERATING CONDITIONS

Parameter	Symbol	Conditions	Range	Unit
Supply voltage	V _{DD}	-	+3.0 to +6.0	V
Operating temperature	T _a	-	30 to +70	°C
External V _{CC} Frequency	f _{ext}	-	0 to 16	kHz

MSM5218RS ADPCM SPEECH ANALYSIS/SYNTHESIS IC

D.C./A.C. CHARACTERISTICS

(V_{CC} = 5V ± 10%, I_B = 0 to 70 μA unless otherwise noted.)

Parameter	Symbol	Conditions	Min	Typ	Max	Unit
Input High Voltage	V _{IH}	All input pins except D, E, F, G pins	3.0			V
Input Low Voltage	V _{IL}	All input pins except D, E, F, G pins			0.8	V
Input High Current	I _{IH}	V _{IH} = V _{CC}		1		μA
Input Low Current	I _{IL}	V _{IL} = 0V		1		μA
Output High Current	I _{OH}	S ₁ ON, V _{OL} = 5V, I _{OL} = 0.4 mA	50			μA
Output Low Current	I _{OL}	S ₁ ON, V _{OL} = 5V, I _{OL} = 0.4 mA	50			μA
Operating Frequency	f _o	Typical Oscillator		104	768	kHz
Operating Current	I _{CC}	f _o = 100 kHz		2	6	mA
Operating Current	I _{CC}	f _o = 100 kHz		0	12	mA
D/A Unit Output Impedance	Z _{out}			100		Ω
D/A Accuracy (Differential to Full Scale)	V _{LSB}	Full Scale		4		LSB
D/A Clock Frequency	f _{CLK}				560	kHz

FIGURE 3

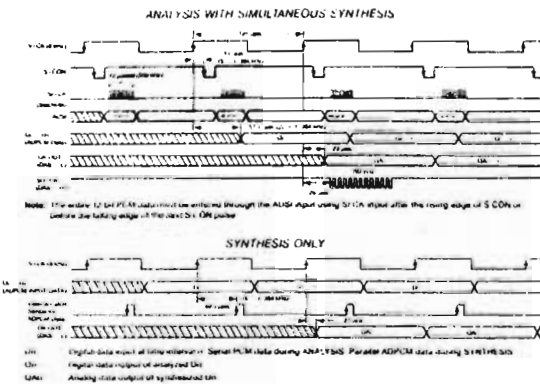


FIGURE 4 BLOCK DIAGRAM—ANALYZER



FIGURE 5 BLOCK DIAGRAM—SYNTHESIZER

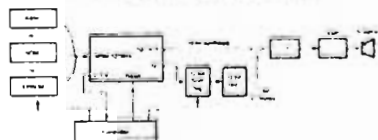
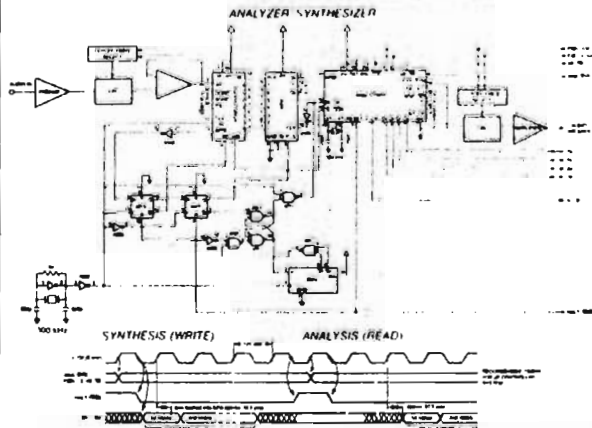


FIGURE 6 TYPICAL APPLICATION—ANALYZER/SYNTHESIZER



OKI SEMICONDUCTOR, INC. 650 N. MARY AVENUE, SUNNYVALE, CA 94086
 TELEPHONE (408) 720-1900 TELEX (25) 910-3380508

OKI Semiconductor reserves the right to make changes in specifications at any time and without notice. The information furnished by OKI Semiconductor in this publication is believed to be accurate and reliable. However, no responsibility is assumed by OKI Semiconductor for its use nor for any infringement of patents or other rights of third parties resulting from its use. No license is granted, under any patents or patent rights of OKI Semiconductor.

MSM5205RS ADPCM SPEECH SYNTHESIS IC

GENERAL DESCRIPTION

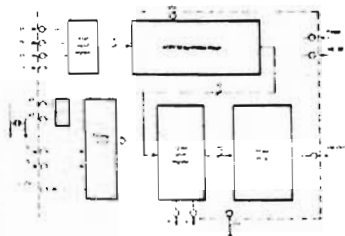
The MSM5205RS is a speech synthesis integrated circuit which includes compressed Adaptive Differential Pulse Code Modulation (ADPCM) data. The initial contents of a synthesis stage which requires the 3- or 4-bit ADPCM data to 12-bit Pulse Code Modulation (PCM) data and a D/A stage which reproduces analog signals from the PCM data.

The MSM5205RS is fabricated using OKI's advanced CMOS process which enables the power consumption. The single power supply requirement and its availability in 8-pin package (EP) make the MSM5205RS to be ideally suited for various applications.

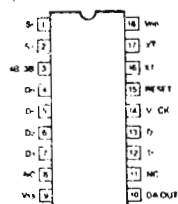
FEATURES

- 3- or 4-bit ADPCM system
- 12 to 52 kHz rate with INT/CLK
- 0.75 to 1.5 mW rate with EXT/CLK
- On-chip 10-bit D/A converter
- Low power consumption (18 mW typical)
- Single -5V supply
- Wide operating temperature (T_a : -30°C to $+70^{\circ}\text{C}$)
- 8-pin mounted DIP

FUNCTIONAL BLOCK DIAGRAM



PIN CONFIGURATION



PIN DESCRIPTION

- B1, B2** These inputs select the sampling frequency or the external mode according to figure 1
- EXT/CLK** Specifies whether 3- or 4-bit ADPCM data is to be processed
- D0-D3** ADPCM data inputs. For 3-bit ADPCM data, DO input is not used and should be connected to ground
- VCK** Ground (GND)
- DA-OUT** Output for synthesized analog signal. Peak-to-peak swing is proportional to VCK. Typical connection scheme is shown in figure 2
- T1, T2** IC lead pins used at the factory for testing purposes only. During normal operation, T1 is grounded and T2 is left open

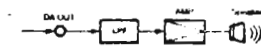
- VCC** An 8-pin pin. This pin outputs a signal whose frequency is equal to the sampling frequency generated by the EXT/CLK inputs. After the desired sampling frequency can be applied through the pin or an external mode is selected via the B1, B2 inputs
- Reset** An active low input which initiates the internal circuitry. Internally, the reset pin is a tri-state output with the VCK signal. Must be true for valid VCK cycle
- XT, XT'** Use these input and output for a 32 kHz crystal or ceramic resonator (figure 3)
- VDD** Power supply pin (nominally -5V)

FIGURE 1 FUNCTION TABLE

B1	B2	Sampling Frequency
L	L	4 kHz
L	H	8 kHz
H	L	0 kHz
H	H	External*

Note: The 32k Hz oscillator must be used whenever 4 kHz or 8 kHz or external mode is selected.
*External mode - Frequency may be from 0 to 10 kHz or 5V - 10V

FIGURE 2

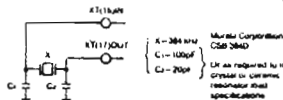


Out of frequency signal (LPF) should be filtered by the selected sampling frequency f_s sampled by $f_s/2$ sample/2 = 0.83

Sound quality is strongly dependent on the impedance of the low pass filter (80Ω) before the speaker load.

*If the 5205 is used in a system of ADPCM data rate is higher than 4 kHz, the DA-OUT pin should be connected to a load around 50 - 100 Ω.

FIGURE 3



XT and XT' (oscillator connection pins)

ABSOLUTE MAXIMUM RATINGS

Parameter	Symbol	Conditions	Rating	Units
Power supply voltage	V _{DD}	T _a = 25°C	0.3 to +7.0	V
Input voltage	V _{in}	T _a = 25°C	0.3 to V _{DD}	V
Power dissipation	P _D	T _a = 25°C	200 max	mW
Storage temperature	T _{stg}		-55 to +150	°C

Note: Stresses above those listed under ABSOLUTE MAXIMUM RATINGS may cause operational damage to the device. This is a stress rating only and not a normal operation of the device at these or all of other conditions unless those indicated in the operational section of this specification are not employed. If required in stress measurement using a combination of all included periods may affect device reliability.

MSM5205RS ADPCM SPEECH SYNTHESIS IC

OPERATING CONDITIONS

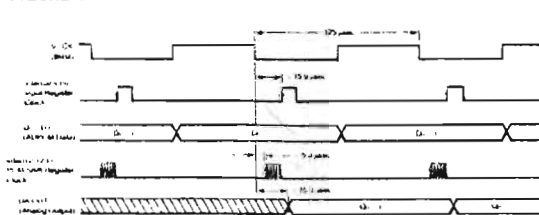
Parameter	Symbol	Conditions	Rating	Units
Power supply voltage	V _{CC}	---	3 to 5	V
Operating temperature	T _{OP}	---	30 to 70	°C
External V _{CK} frequency	f _{CK}	V _{CC} 5V ± 5%	0 to 10 kHz	KHz

D.C./A.C. CHARACTERISTICS

(V_{CC} 5V ± 5%, f_{CK} 30°C to 70°C, unless otherwise noted)

Parameter	Symbol	Conditions	Min.	Typ.	Max.	Units
Input High Voltage	V _{IH}	All inputs except I ₁ , I ₂ , XT and XT pins	3	---	V _{CC} + 0.3	V
Input Low Voltage	V _{IL}	All inputs except I ₁ , I ₂ , XT and XT pins	0	---	0.8	V
Input High Current	I _{IH}	V _{CC} V _{CC}	---	---	1	μA
Input Low Current	I _{IL}	V _{CC} 0V	---	---	1	μA
Output High Current	I _{OH}	V _{CK} pin V _{CC} = 4.2V	50	---	---	μA
Output Low Current	I _{OL}	V _{CK} pin V _{CC} = 0.4V	50	---	---	μA
Oscillator Frequency	f _{osc}	Specified Oscillator	---	384	768 kHz ± 5V ± 5%	KHz
Operating Current	I _{CC}	f _{osc} 384 kHz Min 5V	---	2	4	mA
D/A Accuracy (Internal 10-bit D/A)	V _I	Full Scale V _{CC} - 5V	---	± 4	---	LSB
D/A Output Impedance	V _{OH}	---	70	100	200	kΩ

FIGURE 4



A
P
E
N
D
I
C
E

C

APENDICE C

En este apéndice se presentan las propiedades y conceptos básicos de funcionamiento de algunos dispositivos electrónicos usados en el sistema de evaluación de voz presentado en el capítulo 6.

C.1.- AMPLIFICADOR OPERACIONAL (OPAM)

El OPAM es un amplificador cuya característica está determinada principalmente por la red de retroalimentación. En la figura C.1.a se muestra el símbolo de un OPAM. Las propiedades ideales del amplificador son:

$$\text{Ganancia} = \infty \quad (A = \infty)$$

$$V_0 = 0 \quad \text{cuando} \quad V_1 = V_2$$

$$\text{Impedancia de entrada} = \infty \quad (Z_i = \infty)$$

$$\text{Impedancia de salida} = 0 \quad (Z_0 = 0)$$

$$\text{Ancho de banda} = \infty \quad (\text{retraso de respuesta} = 0)$$

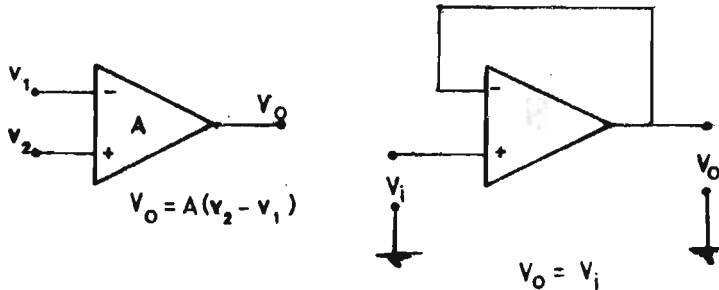


Fig. C.1.- a) Amplificador Operacional ideal.

b) Amplificador Operacional como Seguidor de Voltaje.

- SEGUIDOR DE VOLTAJE

El seguidor de voltaje es fundamentalmente un circuito inversor, donde la ganancia de malla cerrada es igual a la unidad y como su nombre lo indica funciona como un dispositivo el cual el voltaje de salida sigue exactamente el voltaje de entrada.

La ventaja del seguidor de voltaje y en general de los circuitos que no invierten, es la propiedad separadora de impedancia.

$$Z_i = \infty$$

$$Z_o = 0$$

Estos circuitos se usan mucho para proporcionar aislamiento de la fuente de la señal y la carga, impidiendo así interacciones perjudiciales o efectos de "carga".

0.2.- CRISTAL

El cristal es un dispositivo piezoeléctrico el cual transforma energía de voltaje en vibraciones mecánicas o en oscilaciones de voltaje. La frecuencia del cristal está en función de su espesor, con un cristal más delgado se produce una frecuencia más alta.

Los cristales generalmente están especificados como resonadores en serie o resonadores en paralelo. Cuando

un cristal se usa como resonador en serie su salida está en fase con la entrada, mientras que con un cristal como resonador en paralelo la salida está corrida en fase con respecto a la entrada .

Cuando el cristal es utilizado como oscilador, podemos tener:

- Osciladores no inversores que usan preferentemente cristales en serie por sus características.
- Osciladores inversores que usan preferentemente cristales en paralelo.

El circuito equivalente de un cristal, es como el mostrado en la figura C.2a. Cuando funciona como un resonador en serie el circuito equivalente será el de la figura C.2b. Cuando funciona en forma de un resonador en paralelo el circuito equivalente será como el de la figura C.2c.

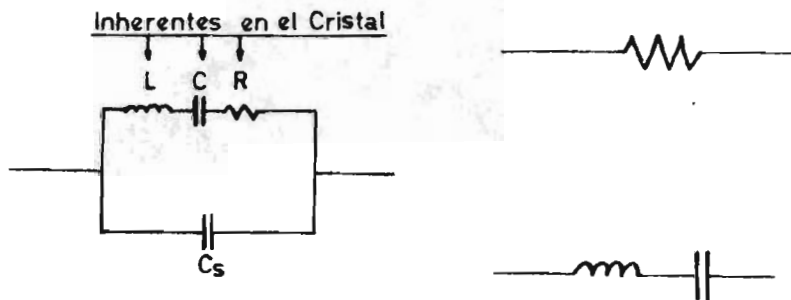


Fig. C.2.- Circuito equivalente del cristal.

0.3.- FLIP-FLOP

Un flip-flop es un circuito de conmutación secuencial con dos estados estables. El flip-flop, es un circuito que puede almacenar información pues posee memoria, pero solamente puede almacenar un bit de información, por que tiene solamente dos estados estables.

La operación de un flip-flop se puede expresar en forma matemática como:

$$Q_i = f(x_i, q_i)$$

donde Q_i es la salida en el tiempo t_i , x_i es la entrada en t_i , q_i es el estado de la memoria en t_i .

Hay solamente cuatro tipos básicos de flip-flops;

- a).- El flip-flop POSICION REPOSICION (RS)
- b).- El flip-flop AMO ESCLAVO (JK)
- c).- El flip-flop DISPARADOR (T)
- d).- El flip-flop de RETRASO (D)

Con excepción del flip-flop RS, todos los demás son circuitos síncronos. Esto provoca que solamente cambie su estado en tiempos predeterminados, usualmente persiguiendo el pulso de la señal de reloj, en cambio el flip-flop RS puede ser cambiado en cualquier tiempo.

En forma resumida de los flip-flops se puede comentar

- Los flip-flops son circuitos biestables y regenerativos
- Las entradas al flip-flop que sólo resultan efectivas con la presencia de un impulso de ritmo se denominan sincrónicas. Las entradas asíncronas o directas actúan directamente en los enclavamientos del flip-flop.
- El enclavamiento del flip-flop puede realizarse con acoplamientos cruzados de compuertas NAND y NOR.
- La lógica de un flip-flop queda contenida en la tabla característica o tabla de excitación del dispositivo. La tabla de excitación indica la excitación necesaria a la entrada para un estado lógico deseado.
- Es necesario comprobar la tabla característica de un flip-flop para determinar la posibilidad de salidas ambiguas.
- El flip-flop RS se efectúa con un circuito de enclavamiento y compuertas adicionales para el control de la entrada de impulsos de reloj o ritmo, de forma que los cambios a la salida del flip-flop están sincronizadas con los tiempos de dichos impulsos.
- El flip-flop JK es la versión en circuito integrado que realiza las funciones rítmicas de POSICION y REPOSICION y además no tiene ambigüedad para el estado de la salida.
- La mayoría de los flip-flops JK trabajan bajo el principio de amo-esclavo.
- Un elemento muy útil es el flip-flop T o circuito divi-

por dos.

- El flip-flop D es un elemento biestable muy empleado en los circuitos digitales.
- Un tipo de flip-flop puede convertirse fácilmente en otro empleando la lógica de control.

C.4.- BASES DE LOS CONVERTIDORES D/A (DIGITAL A ANALOGICO)

Un convertidor D/A es un circuito de decodificación, el cual traslada señales digitales a voltajes o corrientes analógicas. Acepta una señal digital D y una referencia analógica R como entrada, y genera una señal analógica A como salida. La relación entre la entrada y la salida es

$$A = RD$$

donde el número digital está representado por D que es siempre más pequeño que la unidad. Si D representa a un número binario;

$$D = a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n}$$

entonces la salida del convertidor D/A es;

$$A = R(a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n})$$

Esta última ecuación es así la función de transferencia para cualquier convertidor D/A binario.

Simbólicamente un convertidor D/A puede ser apreciado como un potenciómetro digital, donde R es el voltaje conectado a través del potenciómetro y θ es el ángulo a través del cual la flecha es rotada. La analogía entre un convertidor D/A y un potenciómetro se ilustra en la figura C.4a. La función de transferencia para este convertidor es la mostrada en la figura C.4b.

Como se aprecia, sólo se ha asumido que D , r y A son positivos, sin embargo generalmente en un proceso se tienen operaciones bipolares. Esta operación bipolar para un convertidor D/A se muestra en la figura C.4c. En la figura se pueden ver dos métodos posibles, una en la fig. C.4c.1, donde $V_o = 0$ y se encuentra a la mitad entre $+V_R$ y $-V_R$, y $\theta = 0$. Cuando $\theta > 0$, el V_o es positivo y cuando $\theta < 0$, el V_o es negativo. El segundo método (fig. C.4c.2), existen dos terminales, una para $+V_R$, y la otra para $-V_R$ con lo que tendremos un V_o positivo o negativo dependiendo de donde estemos colocados.

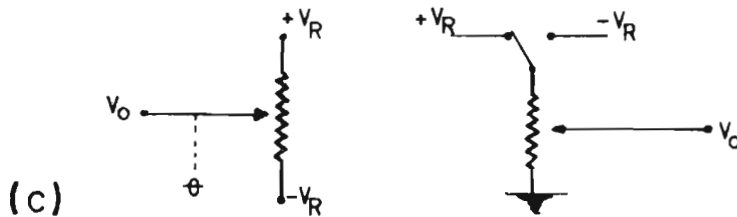
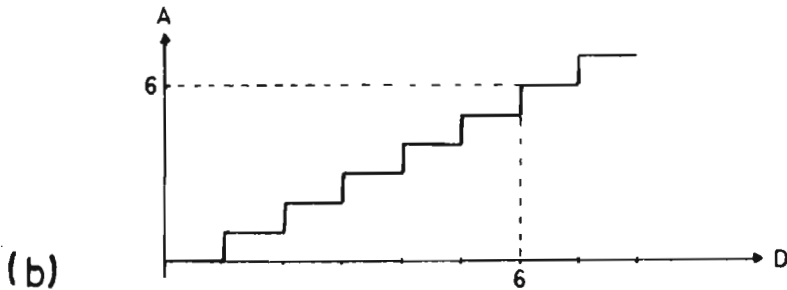
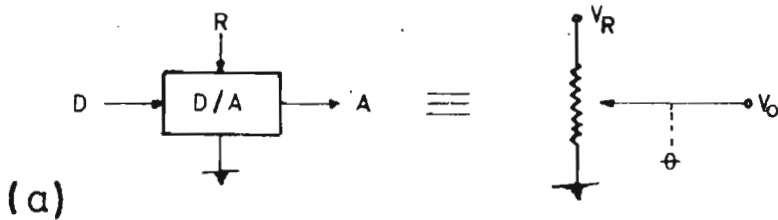


Fig. C.4.- Analogía del Convertidor D/A con un potenciómetro.

Una clasificación de las técnicas de conversión de señales digitales a analógicas queda establecida en forma sintética de la siguiente manera:

- Convertidores D/A Paralelo Digital a DC
- Convertidores D/A Serie.
- Convertidores Indirectos Digital a DC.
- Convertidores Digital a AC.
- Convertidores D/A de Tiempo Compartido.

Donde cada una tiene sus propias subdivisiones, con sus ventajas y utilidad. Aquí, solamente se mencionarán las subdivisiones de las dos primeras.

- Convertidores D/A Paralelo Digital a DC.
 - + El Convertidor D/A de Resistencia de Peso.
 - + El Convertidor D/A con Escalera de Resistencias ($R, 2R$)
 - + El Convertidor D/A con Voltaje de Peso.
 - + El Convertidor D/A con Escalera Invertida.
- Convertidores D/A Serie.
 - + El Convertidor D/A Serie con muestreo/Retención
 - + El Convertidor D/A Cíclico.
 - + El Convertidor D/A Igualando el Cambio.

G.5.- BASES DE LOS CONVERTIDORES A/D (ANALÓGICO A DIGITAL)

El propósito de los circuitos de conversión A/D es trasladar la palabra analógica real a una artificial en el dominio digital. Más específicamente, el convertidor A/D acepta voltajes o corrientes analógicas como entradas y proporciona salidas en forma digital.

Un convertidor A/D es un codificador. Acepta como entrada una señal analógica A y una referencia analógica R y proporciona como salida una señal digital D . En un convertidor ideal A/D, la señal de salida D está relacionada a R y A como:

$$D = \left[\frac{A}{R} \right]$$

donde D es la aproximación más cercana a la relación A/R dentro de la resolución de D .

Si D representa un número binario, tenemos:

$$D = a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n}$$

entonces

$$A \simeq R (a_1 2^{-1} + a_2 2^{-2} + \dots + a_n 2^{-n})$$

Esto es la forma en la cual se desarrolla la cuantización de una señal analógica para la mayoría de los convertidores A/D.

El proceso de conversión A/D tiene una analogía con

el proceso mecánico de poner peso, como se ilustra en la figura C.5a. Donde el peso desconocido W_x es análogo a la entrada analógica A ; el peso de referencia W_R es análogo a R , dividida por el valor máximo de D , mientras el número de pesos de referencia N requiere balancear la escala es análoga al número digital D . Si por ejemplo, W_x pesa 9.37 gramos y el peso de referencia tiene cada una un gramo, entonces se podrá determinar con $N = 9$ que será la más cercana a W_x , y el error debido a esta carencia de resolución es llamado error de cuantización que es en este caso de 0.37 gramos.

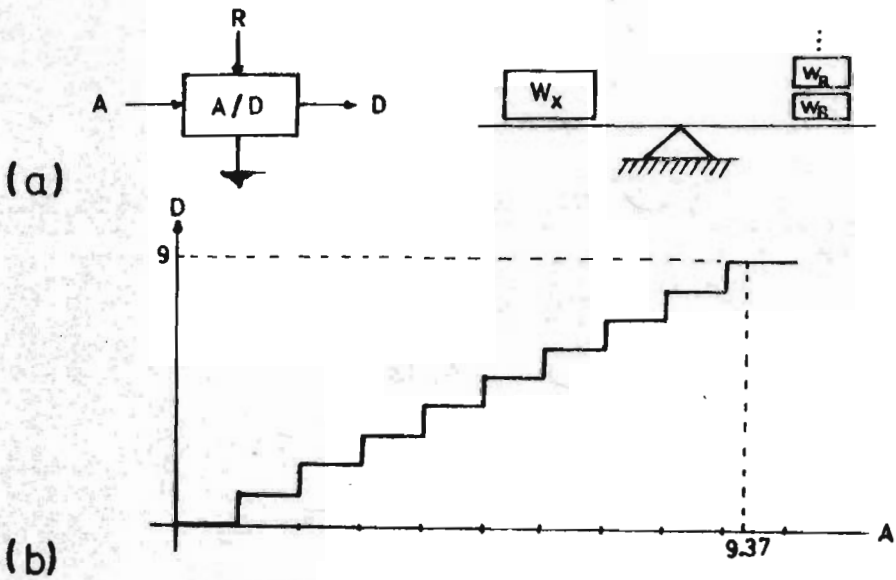


Fig. C.5.- Convertidor A/D.

En la figura C.5b se muestra la función de transferencia del convertidor A/D.

El error de cuantización ocurre no importando si la señal de entrada A es estática (DC) o cambiando con el tiempo (AC), por tanto es un error que está siempre presente. Otro tipo de error es el de muestreo, pero solamente ocurre cuando la entrada A cambia como una función del tiempo.

El proceso de conversión, no importando que tan rápido sea requiere de una cantidad finita de tiempo. Para obtener una conversión libre de ruido, las señales de entrada son comúnmente conmutadas dentro de un circuito de almacenaje analógico y se retiene constante durante el tiempo de conversión. Un circuito típico de conversión usa una señal de entrada específica solamente por un corto tiempo y entonces la ignora para un periodo relativamente largo, este proceso se llama muestreo.

Cuando la frecuencia de A es relativamente alta y el rango de muestreo pequeño ocurren errores considerables, los cuales se trataron en el capítulo 1.

De las ecuaciones aquí vistas, se puede ver que la salida digital D es siempre una aproximación del cociente A/R , esto significa que un convertidor A/D siempre divide la señal de entrada A por R . Esta característica es usada frecuentemente para normalizar señales de entrada, y algunas veces para desarrollar la operación aritmética de división.

Los convertidores A/D conectan los circuitos de computación y control digital con la palabra real. Una clasificación general de los convertidores A/D puede realizarse de la siguiente manera:

- Convertidores A/D con Realimentación en Paralelo.
- Convertidores A/D con Realimentación en Serie.
- Convertidores A/D Indirectos.
- Convertidores A/D de Alta Velocidad y Alta Precisión.
- Convertidores A/D con Velocidad Ultra-Alta.
- Convertidores de Voltaje a Digital Incremental.
- Técnicas de Conversión A/D misceláneas.

Esta clasificación tienen consecuentemente sus propias subdivisiones, como ejemplo podemos mencionar:

- Convertidores A/D con Realimentación en Paralelo
 - + Convertidor A/D Servo.
 - + Convertidor A/D de aproximación Sucesiva.
- Técnicas de Conversión A/D misceláneas.
 - + Convertidor A/D Logarítmico usando Descarga de Capacitor
 - + El Capcoder Logarítmico
 - + El Convertidor A/D con Tubo de Rayos Catódicos.

0.6.- TÉCNICA DE CONVERSION DE APROXIMACION SUCESIVA

El tipo de convertidor que es el más popular, pues aproximadamente el 90% de todos los convertidores A/D en el mercado usa la técnica de conversión de aproximación sucesiva, la cual en algunas ocasiones también es llamada como comparación sucesiva o la técnica de poner y tomar.

Puede ser entendida mejor la operación de un convertidor A/D de aproximación sucesiva haciendo una analogía como anteriormente se hizo con una balanza mecánica. En donde de un lado tenemos el peso desconocido W_x y del otro lado los pesos de referencia. Sin embargo, ahora los pesos son $1/2$, $1/4$, $1/8$, $1/16$, etc., y estos pesos de referencia están relacionados con cada uno con otro en forma binaria. El problema es determinar el peso W_x con el menor número de etapas posibles.

En el convertidor A/D de aproximación sucesiva, el voltaje de realimentación ($-V_F$) está para aproximar el voltaje de entrada V_x en una secuencia de etapas sucesivas, en donde cada etapa $-V_F$ es cambiada de acuerdo con los resultados de la comparación entre V_x y $-V_F$. La cantidad por la cual $-V_F$ es incrementada o decrementada es $V_R/2^{-i}$, donde i define la i -ésimas etapas en la operación

Volviendo a la balanza, si W_x es menor que la unidad (1 kg) pero es desconocido, podemos colocar el peso de $1/2$ kg, pero si $W_x < 1/2$ kg entonces ésta es reemplazada

por la de $1/4$ kg, sin embargo, si $w_x > 1/2$ kg el peso de $1/4$ kg se suma a la de $1/2$ kg. Este proceso se sigue hasta determinar el peso desconocido por aproximación sucesiva (figura 0.6).

Para visualizar mejor lo anterior tendremos un ejemplo específico, y los pesos los cambiamos por voltajes: Entonces, $V_R = -10$ V, $V_X = 8.3$ V y el voltaje de realimentación V_F debe aproximarse a V_X dentro del $\pm 1\%$ o ± 0.1 V, de este proceso tenemos las siguientes etapas:

- 1.- V_F como se explicó será $V_R/2 = -5$ V y se compara con $V_X = 8.3$ V. El voltaje de error V_E es positivo y de aquí el bit más significativo 2^{-1} , de la señal digital de la salida X es de "1" lógico.
- 2.- V_F se incrementa a $3V_R/4 = -7.5$ V. V_E es también positivo, y el bit 2^{-2} de X es también un "1" lógico.
- 3.- V_F se incrementa a $7V_R/8 = -8.75$ V. V_E es entonces negativo, y el bit 2^{-3} de X es por tanto un "0" lógico.
- 4.- V_F ahora se decrementa a $13V_R/16 = -8.125$ V. V_E es positivo, y el bit 2^{-4} de X es un "1" lógico.
- 5.- V_F se incrementa a $27V_R/32 = -8.4375$ V. V_E es entonces negativo y el bit 2^{-5} de X es un "0" lógico.

b.- V_F se decrementa a $53V_R/64 = - 8.27125 V$. V_E es positivo y el bit 2^{-6} es un "1" lógico.

De esta forma, V_F esta dentro de $\pm 0.1 V$ requerido a V_x . Por tanto seis etapas de comparación fueron necesarias para realizar esto. La salida digital así resultante es 110101, el cual representa el 53/64 de la escala total.

Básicamente, el funcionamiento del convertidor de aproximación sucesiva es un compromiso entre la precisión estática y la rapidez de conversión. Por lo que si se fija una alta velocidad de conversión se obtendrá del convertidor la precisión más baja. Así, un criterio excelente para un funcionamiento óptimo total del convertidor A/D es el producto velocidad-precisión.

En los convertidores A/D tratados aquí, el producto velocidad-precisión está limitado por un amplificador comparador, el cual se utiliza para captar la diferencia entre V_x y V_F y genera una señal de control digital. En la conversión se ocupa un tiempo muy pequeño, por lo que el tiempo para la realización de una etapa de comparación es mucho muy pequeño.

En el convertidor, el voltaje de realimentación cambia de 0 a $V_R/2$ en la primera etapa de comparación, de $V_R/2$ a $(1/2 \pm 1/4)V_R$ en la segunda etapa, de ésta a $(1/2 \pm 1/4 \pm 1/8)V_R$ en la tercera etapa, y así sucesivamente. Por tanto, la entrada al amplificador de comparación está

en función de la etapa. La salida del amplificador en respuesta a una entrada en función de la etapa puede ser por tanto, exponencial u oscilatoria.

Algunos de los convertidores más rápidos utilizan menos de 1 μ seg para una etapa de comparación. Durante este corto intervalo de tiempo debe desempeñar las siguientes operaciones:

- 1.- Debe fijarse el circuito de almacenamiento apropiado.
- 2.- Debe ser energizado apropiadamente el interruptor del convertidor D/A.
- 3.- Debe cambiar y determinar el voltaje de realimentación.
- 4.- El amplificador debe desviar de su nivel previo a un nuevo nivel.
- 5.- Debe ser repuesto el circuito de almacenamiento.
- 6.- Debe ser cortado el interruptor del convertidor D/A.
- 7.- Debe regresar a su nivel original el voltaje de realimentación.

Las principales razones para que la técnica de aproximación sucesiva sea usada casi universalmente en todos los sistemas de conversión de datos compatibles con la computadora son; la seguridad de la técnica de conversión y la inherente alta velocidad de conversión de datos. Además, la técnica de aproximación sucesiva proporciona una barata solución de velocidad promedio.

Los convertidores de aproximación sucesiva son principalmente usados donde una señal de entrada está variando continuamente con una relativa alta velocidad o donde una sucesión de entradas son multiplexadas a alta velocidad dentro del convertidor A/D.

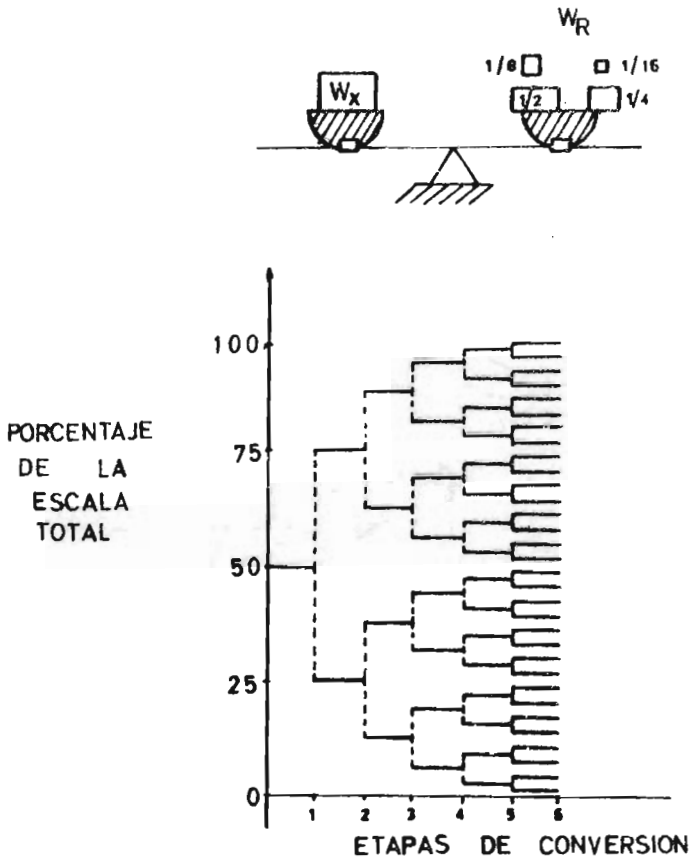


Fig. C.6.- Técnica de aproximación sucesiva.

C.7.- TECNICA DE CONVERSION DE LA ESCALERA DE RESISTENCIAS

Hay muchos tipos diferentes de mallas de resistencias, donde la más común es la escalera R-2R, es decir, es una malla en donde solamente se usan dos tipos de resistencias (figura C.7a).

Las características sobresalientes de esta escalera es que las impedancias en las ramas de cualquier nodo es igual, y que una corriente I fluyendo por el nodo y a través de una rama causa corrientes de $I/2$ para que ésta fluya a través de otras dos ramas. Por tanto esta corriente I causada cuando se cierra un interruptor genera un voltaje V_1 que es dividido entre dos cada vez que esta corriente fluye a través de un nodo como se aprecia en las figuras C.7 b y c.

Como se aprecia en la figura, los interruptores de la escalera de resistencias de S_1 a S_n pueden tener cualquiera de los dos estados V_R o cero. Las propiedades de una malla de escalera de resistencias cuando es usada en un convertidor D/A se ven mejor con unos ejemplos.

- 1.- Si S_1 se conecta a V_R es decir, $V_1 = V_R$ y los otros interruptores a tierra (figura C.7b) existirá un flujo de corriente que se divide en dos partes iguales, donde una $I/2$ fluye hacia el nodo N_2 y la otra $I/2$ fluye a través de la resistencia de carga $R_L = 2R$.

- 2.- Si S_1 está a tierra entonces $V_1 = 0$ y el S_2 se conecta a V_R , entonces $V_2 = V_R$ y los demás interruptores se conectan a tierra (figura 0.7c), esto causa una corriente I , esta corriente en el nodo N_2 es dividida, dando $I/2$ fluyendo hacia N_3 y otra $I/2$ fluyendo hacia N_1 . La corriente en el nodo N_1 se divide otra vez por lo que $I/4$ fluye a tierra e $I/4$ fluye a través de R_L .
- 3.- Si S_1 y S_2 son conectados a V_R , entonces V_1 y $V_2 = V_R$ por superposición la corriente que fluye a través de R_L es $I/2$ de V_1 e $I/4$ de V_2 por lo que la corriente total fluyendo en R_L es $3I/4$.

Todo esto puede ser extendido para varios interruptores. La corriente de salida de la malla de la escalera generalmente se conecta a un amplificador operacional, el cual la convierte a un voltaje.

El funcionamiento de los interruptores analógicos es dependiente del valor de las resistencias de la malla de escalera. Los errores de conmutación son pequeños cuando el valor de las resistencias en la escalera son grandes, o recíprocamente los errores son grandes cuando el valor de las resistencias son pequeños.

El valor de la escalera de resistencias no pueden ser incrementados a voluntad, porque las resistencias más grandes reducen la velocidad de conversión.

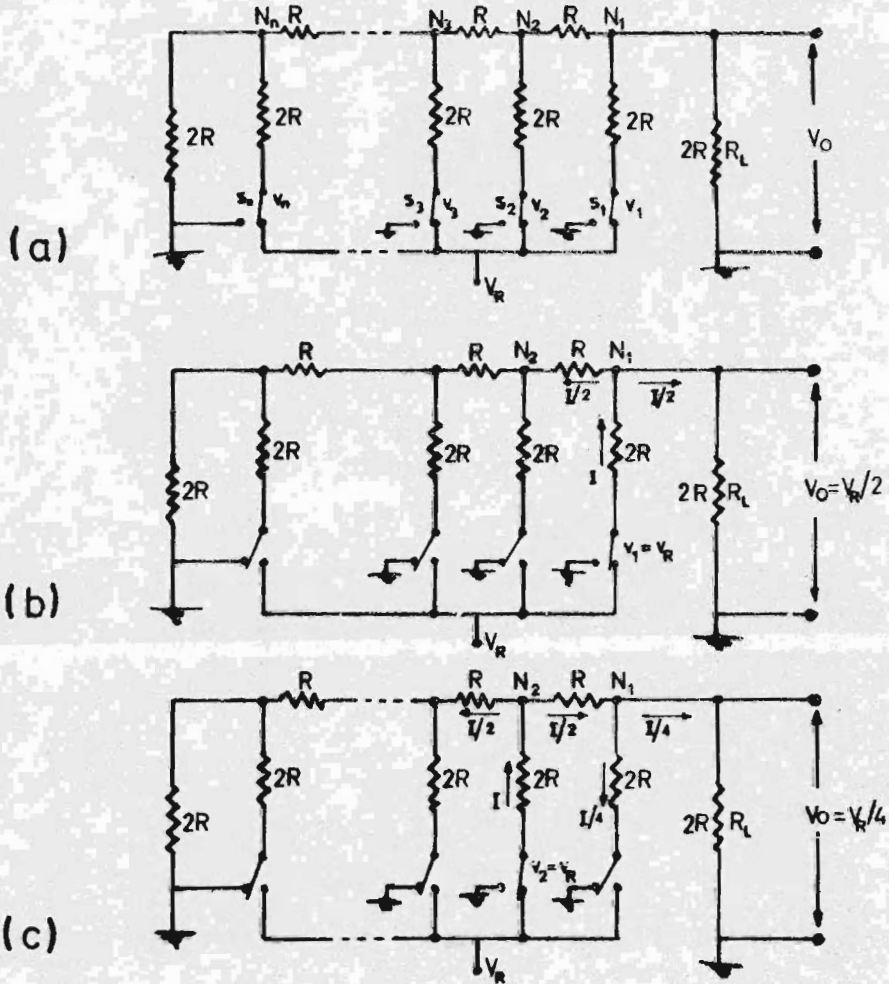


Fig. C.7.- Técnica de Conversión de la Escalera de Resistencias.

3.8.- TECNICAS MISCELANEAS DE CONVERSION A/D

La mayoría de las técnicas convencionales de conversión A/D requieren de circuitería analógica de alto funcionamiento y circuitería digital resultando circuitos híbridos, consistentes de uno o más chips analógicos y un chip digital. Sin embargo, con una nueva técnica totalmente MOS se realiza todo en un simple chip.

En un convertidor A/D de aproximación sucesiva utiliza en su conversión D/A la técnica R-2R, en donde como ya se explicó maneja un interruptor analógico que provoca cambios de corrientes. Sin embargo, si esto lo realizamos con tecnología MOS, el dispositivo ocupará una técnica de descarga de capacitor que se maneja por medio de un interruptor de cambio que tienen inherentemente un voltaje de compensación (off-set) cero y como amplificador tiene una resistencia de entrada muy alta.

Los capacitores son fácilmente fabricados en tecnología MOS, por tanto se usan capacitores en lugar de resistencias como componentes de precisión y usa cambio de voltaje en lugar de un cambio de corriente como medio de trabajo. Esta técnica de descarga de capacitor también es llamada redistribución de cambio o escalera de capacitores.

algunos problemas de control digital o de computación son considerablemente simplificados si las señales

digitales representan el logaritmo de la señal de entrada en lugar de la señal en si misma. Con lo que los convertidores A/D que realizan una función logarítmica tienen un uso limitado para solamente pocas aplicaciones.

- Multiplicación/División. Donde si las variables están disponibles en forma logarítmica, las operaciones se transforman en simples sumas y restas

$$\ln AB = \ln A + \ln B$$

$$\ln A/B = \ln A - \ln B$$

- Relación de Potencias.

$$\begin{aligned} \log P_o/P_i &= \log V_o^2/V_i^2 = 2 \log V_o/V_i \\ &= 2 \log V_o - 2 \log V_i \end{aligned}$$

- Comunicación de Voz Digital. El uso más grande, y eventualmente el potencial comercial más fuerte de los convertidores A/D logarítmicos es el que existe en la comunicación de voz digital. Donde señales analógicas son transformadas a señales digitales.

La operación de los convertidores logarítmicos está basada en la curva logarítmica de descarga de un capacitor. De la figura C.8a se puede ver que el tiempo requerido para descargar un capacitor desde el voltaje de referencia V_R a un punto A, es logarítmico relacionado al voltaje V_A , y en términos matemáticos la descarga de un capacitor se expresa como:

$$V(t) = V_R$$

donde $V(t)$ es el voltaje en el capacitor en cualquier tiempo. Si $V(t)$ es comparado con el voltaje de entrada V_x , y si el tiempo es igual $t = t_x$ tenemos:

$$V_x = V_R$$

Despejando t_x nos da

$$t_x = RC \ln (V_x/V_R)$$

donde si $t_x = 0$, $V_x = V_R$, y si $t_x = \infty$, $V_x = 0$.

La técnica de conversión A/D con un arreglo de capacitores puede ilustrarse con una versión conceptual de un convertidor de 5 bits. Este consiste de un arreglo de capacitores de peso binarios, más un capacitor de peso adicional que corresponde al bit menos significativo (LSB), de un comparador y de interruptores los cuales conectan las placas de los capacitores a ciertos voltajes. El proceso de conversión se realiza en una secuencia de tres operaciones.

1.- **MODO DE MUESTREO.**- Mostrado en la figura C.8b, en donde el interruptor superior conecta las placas superiores de los capacitores a tierra y los interruptores inferiores conectan las placas al voltaje de entrada. Con lo que resulta un cambio sobre lo que existía almacenado en las placas superiores y será proporcional

al voltaje de entrada V_{in} .

2.- MODO DE RETENCION.- Mostrado en la figura C.8c, donde el interruptor superior se conserva aterrizado, y con los interruptores inferiores conectamos las placas a tierra. Como no existió ningún cambio sobre las placas superiores, lo que existía en ellas se conserva y el potencial será $-V_{in}$.

3.- MODO DE REDISTRIBUCION.- Mostrado en la figura C.8d, donde se empieza por probar el bit más significativo (MSB), ésto se realiza conectando el interruptor S_1 y con lo cual conectamos la placa inferior del capacitor al voltaje de referencia V_{ref} del capacitor más grande. Con ésto el voltaje es dividido en dos capacitancias iguales, y el V_x que en la etapa previa era $-V_{in}$, ahora se ve incrementado por $V_{ref}/2$ como resultado de la prueba, entonces:

$$V_x = -V_{in} + V_{ref}/2$$

al realizarse lo anterior, tenemos el signo de V_x , entonces:

$$\text{si } V_x < 0 \quad \text{por tanto} \quad V_{in} > V_{ref}/2$$

y el MSB = 1, pero

$$\text{si } V_x > 0 \quad \text{por tanto} \quad V_{in} < V_{ref}/2$$

el MSB = 0, a la salida del comparador del bit que ha sido probado. Cuando el MSB (b_4) es cero solamente así el interruptor S_1 se regresa a tierra. De una ma

nera similar de como se determinó este primer bit, así se determinarán los siguientes bits continuando progresivamente con el siguiente capacitor más grande es decir, el siguiente capacitor más grande se comparará con V_{ref} probando la división del voltaje con lo que resultará:

$$V_x = -V_{in} + b_4 \frac{V_{ref}}{2} + \frac{V_{ref}}{4}$$

Como se aprecia en la ecuación anterior, V_x converge de nuevo hacia cero - su valor inicial - cuando se completa todo el proceso de conversión.

En la figura 3.8e se muestra la configuración final para la salida digital 01001. Como se puede ver en la figura, todos los capacitores que corresponden a un bit "0" están completamente descargados. El cambio total original sobre las placas superiores ahora han sido redistribuidos en forma binaria solamente sobre los capacitores que les corresponde un "1". Para la resolución de una conversión de N bits, se requieren N redistribuciones.

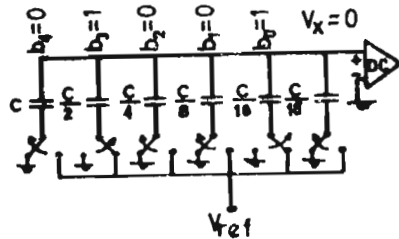
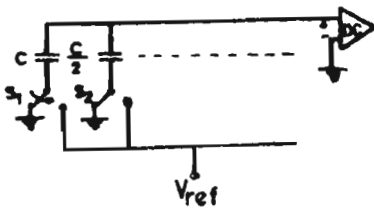
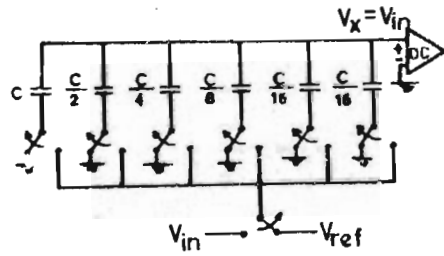
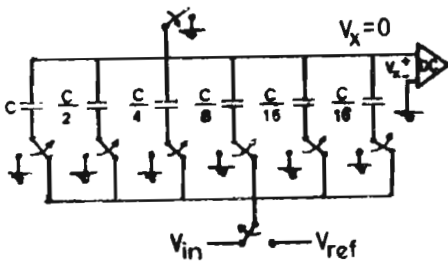
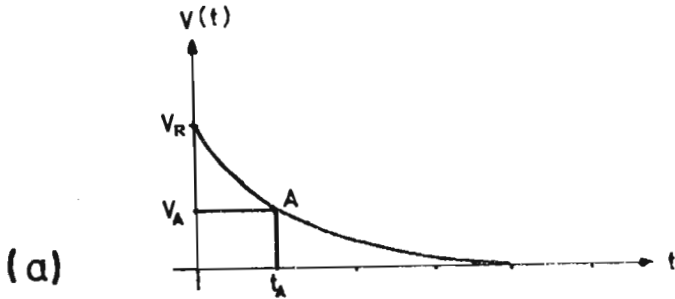


Fig. C.8.- Técnica de Conversión mediante un Arreglo de Capacitores.

U.9.- MUESTREO Y RETENCION

Los circuitos de muestreo y retención juegan un papel muy importante en el equipo de conversión y de interfase. Estos circuitos, muestrean en un punto en particular el valor de una señal analógica variable en el tiempo y almacenan este valor hasta que sea tomada la muestra siguiente o bien hasta que el procesamiento de datos sobre la señal esté completo.

Cualquier circuito muestreador y retenedor consiste de un conmutador o interruptor analógico, un almacenaje intermedio (generalmente un capacitor) y un amplificador. La mayoría de los circuitos de muestreo y retención pueden ser adaptados o categorizados dentro de dos tipos:

TIPO I. Como se ilustra en la figura U.9a, en donde el almacenaje intermedio se encuentra antes del amplificador no inversor.

TIPO II. Como se ilustra en la figura U.9b, en donde el almacenaje intermedio está conectado en la malla de realimentación de un amplificador inversor.

En un circuito ideal de muestreo y retención, el interruptor no tiene compensación, opera instantáneamente y tiene impedancia infinita cuando se abre. Además, el amplificador no tiene ni compensación ni corriente perjudicial y tiene impedancia de entrada infinita. Finalmente,

el almacenamiento intermedio puede seguir a la entrada (V_A) sin retraso, y puede retener el dato sin distorsión para cualquier longitud de tiempo.

La operación del circuito de muestreo y retención ideal se puede expresar matemáticamente como:

$$V_o(t) = V_A(t) \quad \text{si } D = 0$$

$$V_o(t) = V_A(t_c) \quad \text{si } D = 1$$

donde t_c es el instante donde D cambia de un nivel bajo (0) a un nivel alto (1), y donde D es la señal de control digital.

Sin embargo, a pesar de lo que parezca el circuito muestreador y retenedor es un dispositivo relativamente complejo, por lo que hay muchos parámetros que se necesitan especificar.

- Impedancia de Entrada
- Ganancia de Voltaje
- Exactitud de la Ganancia de Voltaje
- Tiempo de Adquisición
- Tiempo de Encendido
- Compensación

El circuito de muestreo y retención del tipo I tiene la ventaja de un tiempo de adquisición muy corto debido a que el capacitor cambia directamente pues se encuentra conectado directamente al interruptor y no al amplificador.

además, no requiere de una resistencia de precisión cuando se desea una ganancia unitaria y no invierten la polaridad.

El circuito del tipo II tiene la ventaja de cambios muy pequeños en la impedancia de entrada.

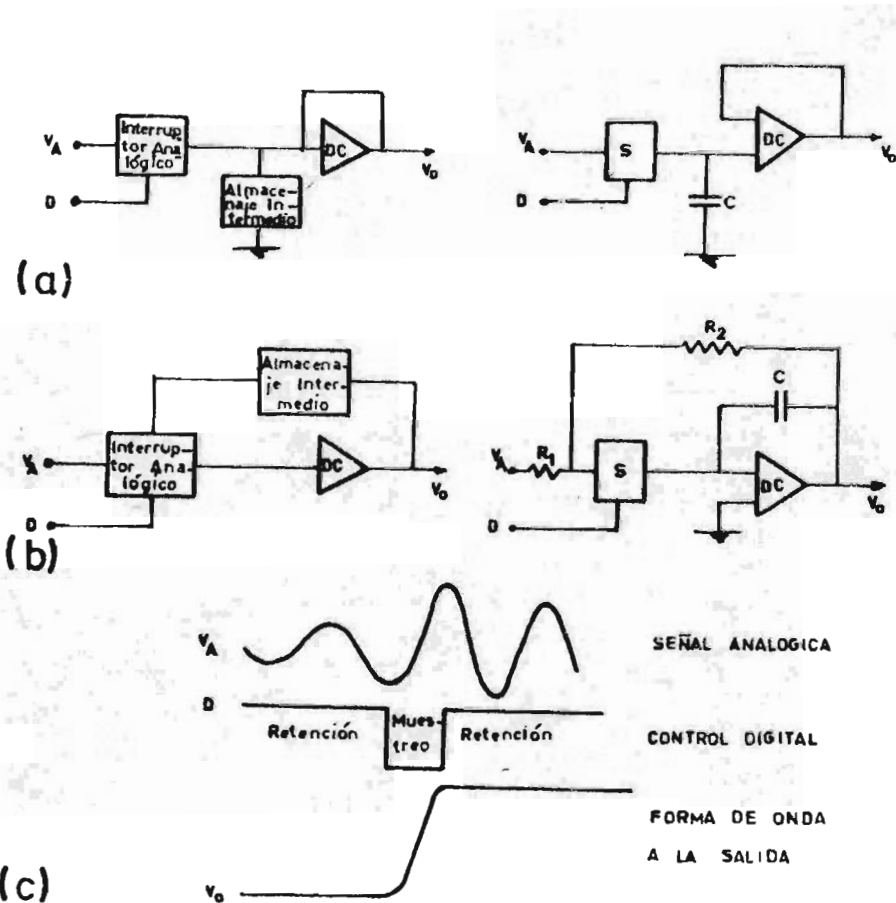


Fig. C.9.- Circuito muestreador y reteneor.

C.10.- COMPADAJOR

Los comparadores se usan en la construcción de dispositivos analógico/digitales, ya que la señal de salida digital, es simplemente la respuesta a la pregunta: ¿es la señal de entrada analógica, mayor o menor que la señal de referencia analógica?. Las señales de referencia y de entrada pueden provenir de fuentes de voltaje o de corriente o de una combinación de ambas.

Los comparadores de voltajes son usados para comparar las magnitudes de dos niveles de voltaje y determinar cual es más grande. Los voltajes serán variables, como es el caso de la entrada analógica de un convertidor A/D.

Dentro de las aplicaciones del comparador es que sirve como compuerta para el convertidor D/A, para incrementar y decrementar el nivel de voltaje de salida dependiendo del resultado de la comparación entre la señal de entrada y la de referencia. También como un circuito que toma decisiones para los convertidores A/D.

Las características importantes del comparador y las condiciones de operación que deben ser consideradas y conocidas para el diseño son:

- Precisión Total
- Sensitividad o Resolución
- Error de Compensación Inicial

- Coeficiente de Temperatura del Error del Voltaje de Compensación
- Rango del Voltaje de Entrada
- Rechazo de modo Común de la Entrada
- Corriente de Entrada
- Impedancia de Entrada
- Histéresis en la Detección
- Recuperación de la Diferencia de Señales de Entrada Grandes
- Tiempo de Encendido
- Tiempo de Apagado
- Rechazo de Variaciones de Voltaje de la Fuente de Poder a la Entrada
- Rango de Temperatura de Operación y del Medio Ambiente Físico en General.

Existen diversas construcciones para un comparador, pero solamente explicaremos como ejemplo el funcionamiento de un comparador simple de voltaje.

En la figura C.10 se muestra el símbolo del comparador y un circuito simple detector de magnitudes relativas de los voltajes de las señales de entrada V_{A1} y V_{A2} . En el circuito, el elemento que determina el umbral del voltaje de entrada es el voltaje base emisor V_{BE} del transistor Q_1 . Si V_{A1} es más positivo que V_{A2} el diodo base-emisor de Q_1 es encendido y la corriente en la base I_B fluye en Q_1 , esta corriente permite el flujo de la corriente al colector con lo cual existe un cambio de voltaje a la sa-

lida (ΔV_o). Cuando la diferencia entre V_{A1} y V_{A2} se incrementa, el ΔV_o incrementa y acerca el valor que es detectado por el circuito conectado a la entrada como una indicación que V_{A1} es más positivo que V_{A2} . El valor de V_o es la diferencia entre V_{A1} y V_{A2} .

Las tres regiones de operación del circuito detector son: una cuando la región de V_{A1} no es más grande o lo suficientemente grande que V_{A2} para que resulte un cambio de voltaje en la salida; otra es la región de incertidumbre donde la salida en el tiempo da una indicación de que V_{A1} es más positiva que V_{A2} ; y la región de encendido, donde el nivel de la salida ha cambiado a un valor definitivo indicando que V_{A1} es más positivo que V_{A2} por lo menos la cantidad necesaria para encender el circuito.

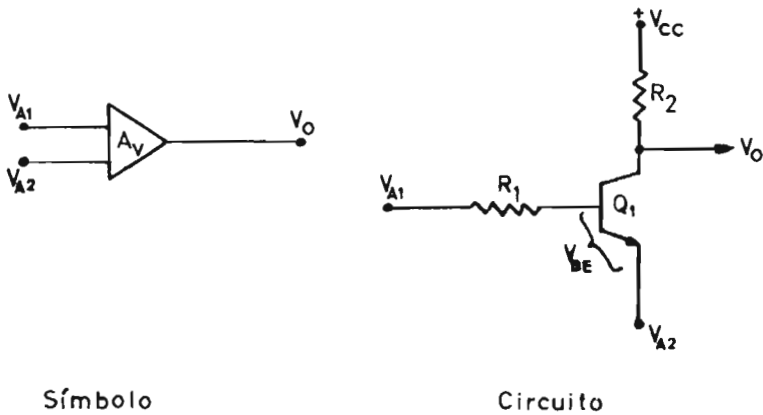


Fig. C.10.- Circuito Comparador

B
I
B
L
I
O
G
R
A
F
I
A

BIBLIOGRAFIA

- 1.- ADPCM COMES OF AGE: ADPCM PRODUCES AN ACCEPTABLE LOW
BIT RATE VOICE ENCODING SCHEME
M.J.Narasimha y S.R. Foster
Telephone Engineer and Management
Septiembre 15, 1984
p.p. 92, 94, 99, 100

- 2.- APPLICATIONS OF DIGITAL SIGNAL PROCESSING
A.V.Oppenheim
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA 1978

- 3.- AUTOMATIC SPEECH SYNTHESIS
D. O'Shaughnessy
IEEE Communications Magazine
Diciembre 1983
p.p. 26 - 34

- 4.- CURRENT PERSPECTIVES IN DIGITAL SPEECH
R.E. Crochiere y J.L. Flanagan
IEEE Communications Magazine
Enero 1983
p.p. 32 - 40

- 5.- DIGITAL FILTERS AND THE FAST FOURIER TRANSFORM
B. Liu
Princeton University
Stroudsburg, Pennsylvania, USA

- 6.- DIGITAL PROCESSING OF SPEECH SIGNALS
L.R. Rabiner y R.W. Schafer
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA 1978
- 7.- DIGITAL SIGNAL PROCESSING
L.R. Rabiner y Ch.M. Rader
IEEE Press
New York, N.Y., USA 1972
- 8.- ELECTRONIC SPEECH SYNTHESIS: TECHNIQUES, TECHNOLOGY AND APPLICATIONS
G. Bristow
M.C Weaw Hill Book, Co.
R.R. Donelley and Sons, Co.
New York, USA 1984
p.p. 346
- 9.- HOW COMPUTERS TALK - AND WHY
J.L. Flanagan
Telephony
Julio 13, 1981
p.p. 25, 28, 29, 31
- 10.- INFORMATION TRANSMISSION, MODULATION AND NOISE
M. Schwartz
International Student
McGraw Hill, Tosho Pritting, Co. LTD
Tokyo, Japan 1981
- 11.- INGENIERIA DE CONTROL MODERNA
K. Ogata
Prentice Hall, Inc.
México D.F., México 1980
p.p. 902

- 12.- LBRV TAKES THE "BYTE" OUT OF T1 EXPANSION
B.W. Berry y S.R.Foster
Telecommunications North American Edition
Octubre 1984
p.p. 80, 83, 85, 86, 89.
- 13.- MULTIRATE DIGITAL SIGNAL PROCESSING
R.E. Crochiere y L.R. Rabiner
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA
- 14.- NEW ALGORITHMS, CHIPS BESTOW HUMAN QUALITIES ON SYNTHESIZED SPEECH
D.Bursky
Electronic Design
Mayo 16, 1985
p.p. 113, 120, 122, 124, 126, 129.
- 15.- THE PHILOSOPHY OF PCM
B.M. Oliver, J.R. Pierce y C.E. Shannon
Proc. IRE, Vol 36, No. 11
Noviembre 1948
p.p. 1324 - 1331
- 16.✓ PRINCIPLES OF COMPUTER SPEECH
J.H. Witten
Academic Press
London, Great Britain 1982
p.p. 286
- 17.- PRINCIPLES OF PULSE CODE MODULATION
K.W. Cattermole
Iliffe Book, Co. LTD
Hungria 1969.

18.- SPEECH ANALYSIS

R.W. Shafer y J.D. Markel
IEEE Press
New York, USA 1979

19.- SPEECH PROCESSING: HEARING BETTER, TALKING MORE

R. Rosenberg
Electronics
Abril 21, 1986
p.p. 26 - 30

20.- SPEECH RECOGNITION: PACKAGING MORE CAPACITY IN SMALLER SYSTEM

S.Chiba
JEE, Junio 1984
p.p. 79 - 82

21.- SPEECH SYNTHESIS

J.N. Holmes
M. & B. Monograph, Mills and Boon Limited
London, Great Britain 1972
p.p. 65

22.- SPEECH SYNTHESIS - FROM TEXT TO SOUND

Telephony
Julio 13, 1981

23.- TECHNOLOGY CONTINUES TO ADVANCE IN SYNTHESIS OF HUMAN SPEECH

S. Hiraoka
JEE, Mayo 1984
p.p. 41 - 45.

- 24.- THEORY AND APPLICATION OF DIGITAL SIGNAL PROCESSING
L.R. Rabiner y B. Gold
Prentice Hall, Inc.
Englewood Cliffs, New Jersey, USA 1975
- 25.- TUTORIAL ON SIGNAL PROCESSING FOR COMMUNICATIONS: PART
I.- DIGITAL TECHNIQUES FOR COMMUNICATIONS SIGNAL PRO-
CESSING
K.V. Mina, V.B. Lawrence y J.J. Werner
IEEE Communications Society Magazine
Enero 1978
p.p. 18 - 22
- 26.- USE ADPCM FOR HIGHLY INTELLIGIBLE SPEECH SYNTHESIS
S. Ciarcia
Byte Publications, Inc.
Junio 1983
p.p. 35 - 41, 44, 46, 48, 49.

BIBLIOGRAFIA COMPLEMENTARIA

- 1.- ALL-MOS CHARGE REDISTRIBUTION ANALOG-TO-DIGITAL CONVERSION TECHNIQUES. PART I
J.L. McCreary y P.R. Gray
IEEE Journal of Solid -State Circuits
Vol. 10, No. 6, Diciembre 1975
p.p. 371 - 179.
- 2.- ALL-MOS CHARGE REDISTRIBUTION ANALOG-TO-DIGITAL CONVERSION TECHNIQUES. PART II
R.E. Suárez, P.R. Gray y D.A. Hodges
IEEE Journal of Solid-State Circuits
Vol. 10, No. 6, Diciembre 1975
p.p. 379 - 385.
- 3.-AMPLIFICADORES OPERATIVOS; DISEÑO Y APLICACION
G.E. Tobey, J.G. Graeme y LN.P. Huelsman
Edit. Diana
México D.F., México 1978
p.p. 548
- 4.- ANALOG TO DIGITAL / DIGITAL TO ANALOG CONVERSION TECHNIQUES
D.F. Hoeschele Jr.
John Weley and Sons, Inc.
New York, USA 1968
p.p. 455
- 5.- A CHARGE -TRANSFER MULTIPLYING DIGITAL - TO ANALOG CONVERTER
J.F. Albarrán y D.A. Hodges
IEEE Journal of Solid-State Circuits
Vol. 11, No. 6, Diciembre 1976
p.p. 772 - 779.

6.- ELECTRONIC ANALOG / DIGITAL CONVERSIONS

H. Schmid

Van Nostrand Reinhold Co.

New York, USA 1970

p.p. 527

7.- THE FAST FOURIER TRANSFORM

E.O. Brigham

Prentice Hall, Inc.

Englewood Cliffs, New Jersey, USA 1974

p.p. 252.