



UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO

FACULTAD DE INGENIERÍA

ANÁLISIS DE CODIFICACIONES DE VOZ
UTILIZADAS EN TELEFONÍA CELULAR
GSM

T E S I S

QUE PARA OBTENER EL TÍTULO DE
INGENIERA EN TELECOMUNICACIONES
P R E S E N T A:
JUDITH VALDÉS CORDERO



DIRECTOR DE TESIS: DR. JOSÉ ABEL HERRERA CAMACHO

CIUDAD UNIVERSITARIA

MARZO DE 2004



Universidad Nacional
Autónoma de México

Dirección General de Bibliotecas de la UNAM

Biblioteca Central



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.

**ESTA TESIS NO SALE
DE LA BIBLIOTECA**

UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO
INSTITUTO DE INVESTIGACIONES Y ENSEÑANZA DE LA QUÍMICA
CARRERA DE QUÍMICA

TRABAJO DE GRADUACIÓN
DEL PROGRAMA DE QUÍMICA
EN QUÍMICA

ALUMNO: [Faint Name]
CARRERA DE QUÍMICA
CARRERA DE QUÍMICA

TRABAJO DE GRADUACIÓN
DEL PROGRAMA DE QUÍMICA
EN QUÍMICA

DEDICATORIA

A la Universidad Nacional Autónoma de México,
y en especial, a la Facultad de Ingeniería.

A mis padres con todo mi amor y agradecimiento, a mis hermanos Ernesto y Rosa María por su continuo apoyo, ejemplo y cariño, y a mi sobrino Ray por hacerme la vida más divertida.

A mi tía Judith Cordero con infinito amor y gratitud, por su ejemplo y su ayuda y a Judy Campos por crecer y compartir conmigo.

A mi Maestro, Abel Herrera Camacho por toda su enseñanza, paciencia y tiempo dedicado a este trabajo.

A Diego, Enrique, Francisco, Goyo, Jorge, José Luis, Rita y Ximena por tiempos, experiencias y enseñanzas, por su amistad.

A Gilberto por el amor, apoyo y admiración compartida y a su hermosa familia por toda su convivencia.

Autorizo a la Dirección General de Bibliotecas de la UNAM a difundir en formato electrónico e impreso el contenido de mi trabajo recepcional

NOMBRE: JUDITH VADES CERDEPO

FECHA: 18/marzo/2009

FIRMA: [Firma]

Índice.

| | |
|--|----|
| Objetivos | 1 |
| Introducción | 2 |
| Capítulo 1. Telefonía Celular | 3 |
| 1.1 Sistema de telefonía celular | 3 |
| 1.1.1 Configuración de un sistema celular | 4 |
| 1.1.2 Reutilización de frecuencias | 6 |
| 1.2.3 Asignación de canales | 7 |
| 1.2.4 Interfaz de radio | 8 |
| 1.2.5 Localización | 9 |
| 1.2.6 Función de traspaso (<i>Handover</i>) | 9 |
| 1.2.7 Función de seguimiento (<i>Roaming</i>) | 10 |
| 1.2.8 Interferencia de multi – trayectoria | 10 |
| 1.2.9 Control de potencia | 11 |
| 1.2.10 División de celdas | 11 |
| 1.2 Estándares de primera generación | 13 |
| 1.2.1 AMPS (<i>Advances Mobile Phone Service</i>) | 13 |
| 1.2.2 NMT-450 (<i>Nordic Mobile Telephone 450 MHz</i>) | 14 |
| 1.2.3 TACS (<i>Total Access Cellular System</i>) | 15 |
| 1.2.4 NMT-900 | 16 |
| 1.2.5 El estándar japonés NTT | 16 |
| 1.2.6 C-450 o "C-NETZ" (<i>German Digital System</i>) | 17 |
| 1.2.7 El estándar francés Radiocom 2000 (RC-2000) | 18 |
| 1.2.8 RMTS (<i>Italian System</i>) | 18 |
| 1.2.9 MATS-E | 18 |
| 1.3 Estándares de Segunda generación | 20 |
| 1.3.1 GSM (<i>Global System for Mobile communications</i>) | 20 |
| 1.3.2 DCS-1800 | 24 |
| 1.3.3 PCS-1900 | 25 |
| 1.3.4 D-AMPS | 26 |
| 1.3.5 CDMA (IS-35) | 27 |

| | | |
|---|---|-----------|
| 1.3.6 | PDC | 29 |
| 1.4 | Evolución de los estándares y estado actual | 31 |
| Capítulo 2. Sistema GSM de telefonía celular | | 35 |
| 2.1 | Arquitectura del sistema | 35 |
| 2.1.1 | Subsistema de Terminal Móvil (MS) | 36 |
| 2.1.2 | Subsistema de Estación Base (BSS: <i>Base Station Subsystem</i>) | 36 |
| 2.1.3 | Subsistema de red y conmutación (NSS) | 37 |
| 2.1.4 | Subsistema de Operación (OSS: <i>Operation Subsystem</i>) | 38 |
| 2.2 | Realización Técnica | 40 |
| 2.2.1 | Aspectos técnicos de la comunicación de voz | 40 |
| 2.2.2 | Transmisión de datos: Modo transparente | 42 |
| 2.2.3 | Transmisión de datos: Modo no transparente | 43 |
| 2.2.4 | Realización técnica del Servicio de Mensajes Cortos | 43 |
| 2.3 | Aspectos Técnicos de Seguridad | 45 |
| 2.4 | Interconexión con otras redes | 48 |
| 2.5 | Servicios | 52 |
| 2.6 | Codificación de voz | 54 |
| 2.7 | Nuevos desarrollos en GSM | 56 |
| Capítulo 3. Codificación de voz | | 60 |
| 3.1 | Historia | 60 |
| 3.2 | Tipos de codificación | 62 |
| 3.2.1 | Codificadores de forma de onda | 62 |
| 3.2.2 | Vocoders | 69 |
| 3.2.3 | Codificadores híbridos | 79 |
| 3.3 | Simulación de codificadores | 95 |
| 3.3.1 | DPCM | 95 |
| 3.3.2 | ADPCM | 96 |
| 3.3.3 | Vocoder LPC | 96 |
| 3.3.4 | Vocoder LPC10 | 97 |
| 3.3.5 | Codificación RELP | 97 |
| 3.3.6 | Multipulso | 99 |
| 3.3.7 | RPE-LTP | 99 |

| | |
|--|-----|
| Capítulo 4. Evaluación de los codificadores seleccionados | 101 |
| 4.1 Sistemas de Prueba | 101 |
| 4.1.1 Medidas Objetivas | 101 |
| 4.1.2 Medidas Subjetivas | 102 |
| 4.2 Resultados | 106 |
| | |
| Conclusiones | 111 |
| | |
| Bibliografía. | 113 |

Objetivo

Se marcan como objetivos los siguientes:

1. Analizar el ambiente de la telefonía celular, la importancia de los sistemas digitales y en especial GSM.
2. Analizar los diferentes esquemas de codificación de voz utilizadas en GSM.
3. Simular por medio del programa Matlab los algoritmos de codificación de voz seleccionados.
4. Determinar la calidad de las codificaciones seleccionadas.
5. Determinar las características necesarias que debe tener una codificación de voz aplicada a la telefonía celular GSM, para en un trabajo futuro tener las bases y poder diseñar y desarrollar una mejora de dicha codificación.

Introducción

La necesidad del ser humano de comunicarse ha llevado al desarrollo de grandes tecnologías en el ámbito de las telecomunicaciones y en especial en la telefonía celular debido a la gran movilidad de las personas en el mundo. Estos grandes avances realmente se han dado en un tiempo considerablemente corto, por lo que no es de extrañarse que en el tiempo en que se desarrolló este trabajo, nuevas tecnologías pudieron ser creadas.

Este trabajo tiene la finalidad dentro de un panorama general de analizar la importancia y los diferentes tipos de codificación de voz utilizada en un sistema específico de telefonía celular. Para lograr dicha finalidad, este trabajo está dividido en capítulos cada uno con un fin específico.

Así, el capítulo uno presenta un panorama general de los sistemas de telefonía celular actuales, desde los analógicos hasta los digitales, resaltando la importancia de los últimos y en especial *GSM* con base en la demanda y características técnicas mostrados de dichos sistemas.

En el capítulo dos se presenta a detalle el funcionamiento de un sistema de telefonía celular *GSM*, desde la arquitectura del sistema, realización técnica, seguridad, servicios, interconexión con otras redes y por supuesto en este caso, haciendo énfasis en la importancia de la codificación de voz y los diferentes tipos de codificación que se utilizan en dicho sistema y la repercusión de sus características en un espectro de frecuencias limitado y saturado como el de la telefonía celular *GSM*.

En el capítulo tres se presenta un panorama histórico sobre la codificación, los tipos de codificación más importantes usados en telefonía celular, sus características y algoritmos de una manera evolutiva hasta llegar a los algoritmos en los que se centra este trabajo además de la simulación de estos codificadores con ayuda de una *PC* y del programa *Matlab*, obteniendo ya ciertas características importantes para este trabajo.

Por último, el capítulo cuatro, se describirá los diferentes métodos de evaluación que son usados para determinar la calidad de las codificaciones, tanto medidas objetivas como subjetivas y por supuesto aplicarlas a las codificaciones seleccionadas, presentando los resultados de dicho procedimiento.

Capítulo 1

Telefonía celular

1.1 Sistema de telefonía celular

El desarrollo de la telefonía celular ha tenido un gran auge en los últimos años debido a la gran necesidad creada por el hombre para comunicarse en cualquier momento y en cualquier lugar de nuestra vida cotidiana.

Los primeros desarrollos empiezan cuando el Departamento de Policía de Detroit a inicios de los 20's realizaron los primeros experimentos de transmisión unidireccional en la banda de 2 [MHz] con el fin de tener comunicación entre la estación base y los coches de policía y en el año de 1931 se puso en funcionamiento el primer sistema de comunicaciones bidireccional extendiendo este sistema a los demás servicios públicos.

En 1935 Edwin H. Armstrong publica un importante trabajo a cerca de las señales de frecuencia modulada (*FM*), con importantes aportaciones, debido a que este tipo de modulación la información se transporta en una frecuencia instantánea, por lo que resulta más resistente a los desvanecimientos rápidos que sufre la señal en entornos móviles, teniendo como ventaja que se requiere menos potencia de transmisión. Esto generó que a finales de los años 30 la policía de Connecticut desarrollara el primer sistema de comunicaciones móviles en *FM*.

Posteriormente los sistemas de comunicaciones móviles se extendieron a la milicia, donde Laboratorios *Bell* y *Western Electric* iniciaron el desarrollo de sistemas de comunicaciones móviles para vehículos militares, logrando grandes avances en cuanto a la fiabilidad, reducción de tamaño y disminución de costos, lo que daría la capacidad tecnológica para desarrollar sistemas basados en *FM* comerciales.

El sistema predecesor de los sistemas celulares fue *IMTS* (*Improved Mobile Telephone System*), el cual incluía el concepto de *trunking* o encadenamiento automático de llamada, que permite un total acceso de todos los usuarios a todos los canales de la banda, incrementando la capacidad del sistema y la eficiencia espectral, un servicio bidireccional (*full-duplex*) y una marcación directa. Sin embargo este incremento en la capacidad del sistema y eficiencia espectral no fueron suficientes ya que cada estación base disponía de 25 canales con radios de cobertura entre 50 y 70 [Km], con lo que para zonas muy pobladas esto no era suficiente, teniendo como consecuencia que los Laboratorios *Bell* realizaran trabajos con el concepto de arquitectura celular. Este planteaba el uso de un gran número de estaciones base de cobertura denominada célula. El sistema permite la

reutilización de frecuencias en células suficientemente separadas sin que la interferencia sea un problema, multiplicando la capacidad de la banda de frecuencias utilizada.

El sistema *IMTS* fue insuficiente para la aplicación del concepto celular, ya que es necesario una disponibilidad de una banda de frecuencia definida en tamaño y el desarrollo de sistemas de señalización y control complejos para tener el seguimiento de las llamadas, con lo que con los pocos canales disponibles en este sistema era imposible.

Fue en el año de 1973 cuando en Estados Unidos, Martin Cooper, quien en ese tiempo trabajaba para *Motorola*, introdujo el primer radioteléfono, sin embargo los primeros sistemas comerciales fueron desarrollados por la compañía Japonesa *NTT* seis años más tarde y en 1981, los países nórdicos (Noruega, Suecia, Finlandia y Dinamarca) introdujeron un sistema celular similar a *AMPS* (*Advanced Mobile Phone System*). Por otro lado, en Estados Unidos, gracias a que la entidad reguladora de ese país adoptó reglas para la creación de un servicio comercial de telefonía celular, en 1983 se puso en operación el primer sistema comercial en la ciudad de Chicago.

1.1.1 Configuración de un sistema celular

Un sistema celular se constituye de varios bloques que se ensamblan para formar una red de comunicaciones. Los cuales son:

1. Estación móvil, *MS* (*Mobile Station*): que es un dispositivo usado para comunicarse en una red celular.
2. Estación transceptora base, *BST* (*Base Station Transceiver*): un transmisor / receptor usado para transmitir / recibir señales de la sección de radio de la red.
3. Centro conmutado móvil, *MSC* (*Mobile Switching Center*): establece y mantiene las llamadas que hacen en la red.
4. Controlador de estación base, *BSC* (*Base Station Controller*): controla las comunicaciones entre un grupo de estaciones base y un único centro conmutador móvil.
5. Red de telefonía pública conmutada, *PSTN* (*Public Switched Telephone Network*): se encarga de la sección terrestre de la red.

La figura 1.1 ilustra cómo estos bloques se relacionan unos con otros dentro de la red. Las estaciones base y su controlador a menudo se refieren colectivamente como el subsistema estación base, *BSS* (*Base Station Subsystem*). Debido a que la topología celular se desarrolló por el limitado espectro de radio, se reutilizan las mismas frecuencias en celdas no adyacentes. Una región geográfica se divide en celdas, cada celda tiene una estación base que transmite a través de un vínculo de radio a las estaciones móviles dentro de la celda. Un grupo de controladores de

estaciones base se encuentran conectadas a centro conmutador móvil a través de vínculos de microondas y líneas telefónicas. El centro conmutador móvil se conecta a la red de telefonía pública conmutada, la cual deriva llamadas a otras estaciones móviles o teléfonos terrestres.

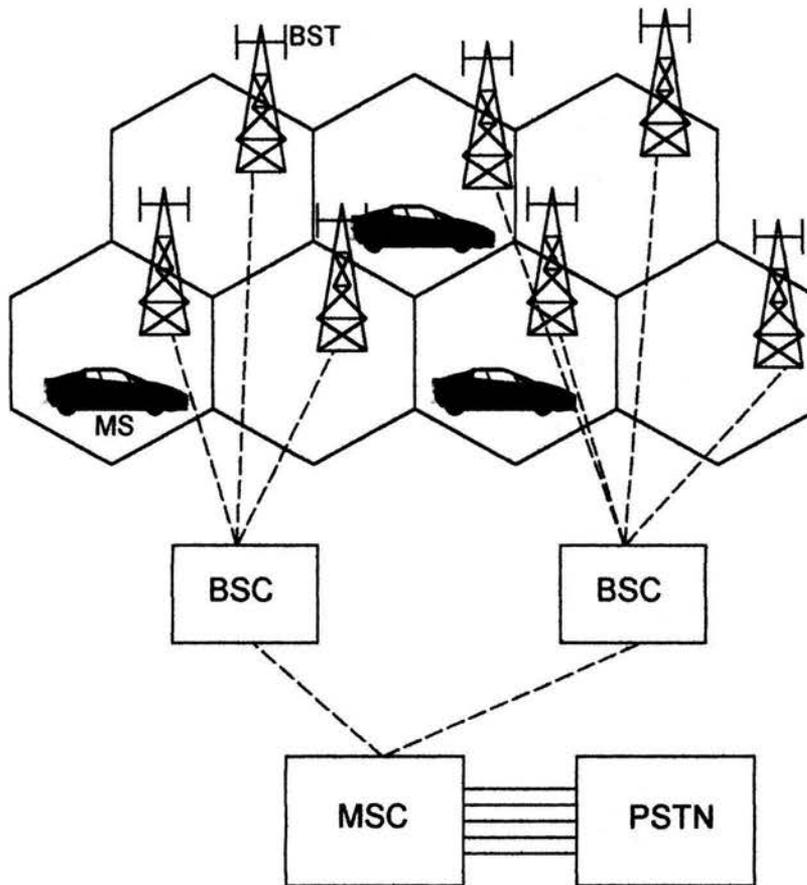


Figura 1.1 Un sistema celular.¹

La conmutación entre estaciones base y el equipo telefónico celular se realiza mediante transmisores de potencia relativamente baja, a una frecuencia de 800 [MHz]. La potencia de salida de una estación base no puede exceder los 100 [W] de potencia efectiva radiada (ERP).

El principal objetivo de un sistema celular es proporcionar una cobertura geográfica continua dentro de los que se conoce como Áreas de Servicio Geográfico CGSA (*Cell Geographic Service Area*). La localización y el número de las estaciones geográficas esta determinada por el tráfico específico y las necesidades geográficas.

¹ <http://www.acm.org/crossroads/espanol/xrds7-2/cellular.html>

1.1.2 Reutilización de frecuencias

Cada estación base tiene su propio equipo para transmitir y recibir hacia y desde los móviles. Aunque un sistema celular podría usar células rectangulares (las cuales no tienen traslapes), se han escogido células de forma hexagonal porque permite cubrir un área mayor con un número menor de estaciones base. La estación base puede colocarse en el centro de la celda o en la esquina de los hexágonos. En el primer caso, la estación celular utiliza antenas omnidireccionales, mientras que en el segundo se utilizan antenas direccionales con un ancho del lóbulo de radiación de 120° , que permiten cubrir parte de tres células adyacentes.

A cada celda se le asigna un grupo de canales, los mismos conjuntos de frecuencias se vuelven a utilizar en celdas no adyacentes separadas por una distancia adecuada, esto es a lo que se conoce como reutilización de frecuencias y determina una de las características esenciales de los sistemas celulares. En vez de cubrir el área mediante un transmisor de gran potencia en un sitio elevado, se puede proporcionar el servicio mediante varios transmisores de potencia moderada distribuidos de manera adecuada en toda el área, lo que permite un incremento en la capacidad del sistema. En la figura 1.2 se muestra gráficamente un ejemplo del área de servicio cubierta por un grupo de células.

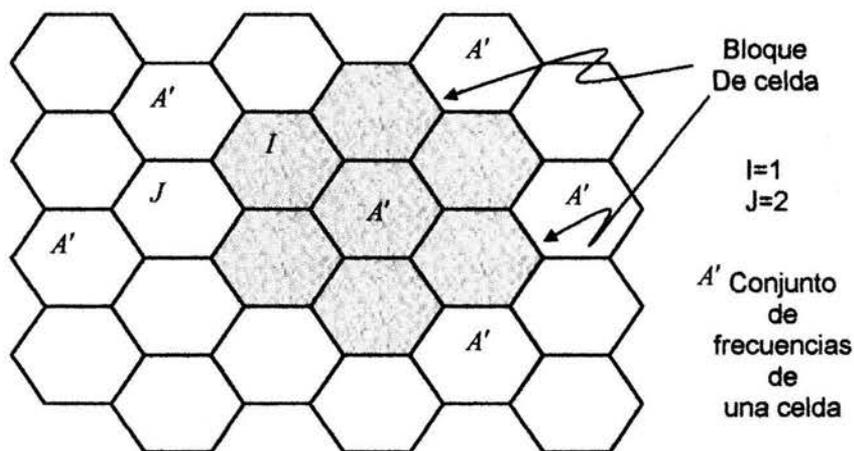


Figura 1.2 Área de servicio cubierta con un grupo de células.

La reutilización de frecuencias ocasiona una interferencia entre canales que utilizan la misma frecuencia. Para controlar este efecto se define el parámetro "D" como la distancia mínima entre celdas que utiliza el mismo grupo de frecuencias, obviamente a valores grandes de D menor interferencia entre canales que utilizan la misma frecuencia.

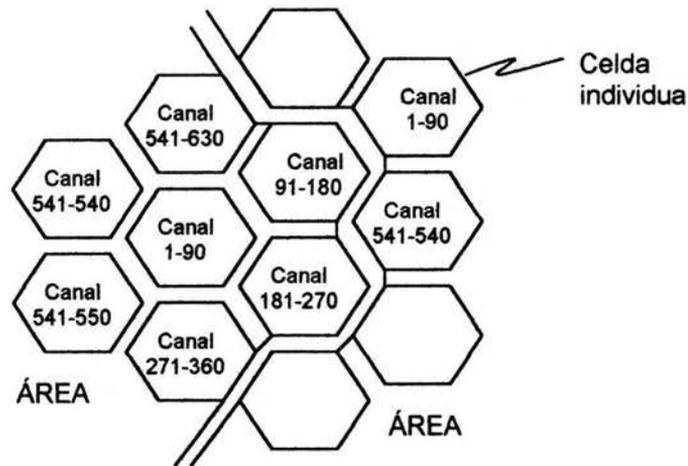


Figura 1.3 Distribución de canales entre celdas.

1.1.3 Asignación de canales

Para la utilización eficiente del espectro de radio, se requiere un sistema de reutilización de frecuencias que aumente la capacidad y minimice las interferencias. Se han desarrollado una gran variedad de estrategias de asignación de canales para llevar a cabo estos objetivos. Las estrategias de asignación de canales se pueden clasificar en fijas o dinámicas. La elección de la estrategia de asignación de canales va a imponer las características del sistema, particularmente, en cómo se gestionan las llamadas cuando un usuario pasa de una celda a otra (*handover*).

En una estrategia de asignación de canales fija, a cada celda se le asigna un conjunto predeterminado de canales. Cualquier llamada producida dentro de la celda, sólo puede ser atendida por los canales inutilizados dentro de esa celda en particular. Si todos los canales de esa celda están ocupados, la llamada se bloquea y el usuario no recibe servicio. Existen algunas variantes de ésta estrategia. Una de ellas permite que una celda vecina le preste canales si tiene todos sus canales ocupados. El Centro de Conmutación Móvil (*MSC*) supervisa que estos mecanismos de presta no interfieran con ninguna de las llamadas en progreso de la celda donadora.

En una estrategia de asignación de canales dinámica, los canales no se colocan en diferentes celdas permanentemente. En su lugar, cada vez que se produce un requerimiento de llamada la estación base servidora pide un canal al *MSC*. Éste entonces coloca un canal en la celda que lo pidió siguiendo un algoritmo que tiene en cuenta diversos factores como son la frecuencia del canal a pasar, su distancia de reutilización y otras funciones de costo.

Las estrategias de asignación dinámicas aumentan las prestaciones del sistema, pero requieren por parte del MSC una gran cantidad de cómputo en tiempo real.

1.1.4 Interfaz de radio

Dado que el número de canales de radio es mucho menor que el número total de usuarios potenciales, los canales bidireccionales sólo se asignan si se necesitan. Esta es la principal diferencia con la telefonía estándar, donde cada terminal está continuamente unido a un conmutador haya o no haya llamada en progreso.

En una red móvil como *GSM (Global System for Mobile communications)*, los canales de radio se asignan dinámicamente. En *GSM*, así como en otros sistemas de telefonía celular, el usuario que está en espera permanece atento a las posibles llamadas que se puedan producir escuchando un canal específico. Este canal transporta mensajes llamados mensajes de búsqueda: su función es la de advertir que un usuario móvil está siendo llamado. Este canal es emitido en todas las celdas, y el problema de la red es determinar en qué celdas llamar a un móvil cuando se le necesite.

El establecimiento de cualquier llamada, ya sea el móvil origen o destino de la llamada, requiere medios específicos por los cuales la estación móvil pueda acceder al sistema para obtener un canal. En *GSM*, este procedimiento de acceso se realiza sobre un canal específico del móvil a la base. Este canal, que envía además de otra información, los mensajes de búsqueda, es conocido en *GSM* como canal común dado que lleva información hacia y desde el móvil al mismo tiempo. Los canales asignados durante un periodo de tiempo a un móvil se les llaman canales dedicados. Basados en esta distinción se pueden definir dos macro-estados:

1. Modo desocupado ("*idle*"), en el que el móvil escucha; la estación móvil no tiene ningún canal para sí misma.
2. Modo dedicado, en el que se asigna un canal bidireccional a la estación móvil para sus necesidades de comunicación, permitiéndole a éste intercambiar información punto a punto en ambas direcciones.

El procedimiento de acceso es una función particular que permite a la estación móvil alcanzar el modo dedicado desde un modo desocupado.

1.1.5 Localización

La movilidad de los usuarios en un sistema celular es la fuente de mayores diferencias con la telefonía fija, en particular con las llamadas recibidas. Una red puede encaminar una llamada hacia un usuario fijo simplemente sabiendo su dirección de red, dado que el conmutador local, al cual se conecta directamente la línea del abonado, no cambia. Sin embargo en un sistema celular la celda en la que se debe establecer el contacto con el usuario cambia cuando éste se mueve. Para recibir llamadas, primero se debe localizar al usuario móvil, y después el sistema debe determinar en qué celda está actualmente.

En la práctica se usan tres métodos diferentes para tener este conocimiento. En el primer método, la estación móvil indica cada cambio de celda a la red. Se le llama actualización sistemática de la localización al nivel de celda. Cuando llega una llamada, se necesita enviar un mensaje de búsqueda sólo a la celda donde está el móvil, ya que ésta es conocida. Un segundo método sería enviar un mensaje de página a todas las celdas de la red cuando llega una llamada, evitándonos así la necesidad de que el móvil esté continuamente avisando a la red de su posición. El tercer método es un compromiso entre los dos primeros introduciendo el concepto de área de localización. La identidad del área de localización a la que una celda pertenece se les envía a través de un canal de difusión (*broadcast*), permitiendo a las estaciones móviles saber el área de localización en la que están en cada momento. Cuando una estación móvil cambia de celda se pueden dar dos casos: ambas celdas están en la misma área de localización: la estación móvil no envía ninguna información a la red o las celdas pertenecen a diferentes áreas de localización: la estación móvil informa a la red de su cambio de área de localización.

Cuando llega una llamada solamente se necesita enviar un mensaje a aquellas celdas que pertenecen al área de localización que se actualizó la última vez. *GSM* realiza este método.

1.1.6 Función de traspaso (*Handover*)

En el apartado anterior se trataron las consecuencias de la movilidad en el modo *idle*. En el modo dedicado, y en particular cuando una llamada está en progreso, la movilidad del usuario puede inducir a la necesidad de cambiar de celda servidora, en particular cuando la calidad de la transmisión cae por debajo de un umbral. Con un sistema basado en células grandes, la probabilidad de que ocurra esto es pequeña y la pérdida de una llamada podría ser aceptable. Sin embargo, si queremos lograr grandes capacidades tenemos que reducir el tamaño de la celda, con lo que el mantenimiento de las llamadas es una tarea esencial para evitar un alto grado de insatisfacción en los clientes.

Al proceso de la transferencia automática de una comunicación (de voz o datos) en progreso de una celda a otra para evitar los efectos adversos de los movimientos del usuario se le llama "handover" (o "handoff"). Este proceso requiere primero algunos medios para detectar la necesidad de cambiar de celda mientras estamos en el modo dedicado (preparación del *handover*), y después se requieren los medios para conmutar una comunicación de un canal en una celda dada a otro canal en otra celda, de una forma que no sea apreciable por el usuario.

1.1.7 Función de seguimiento (*Roaming*)

Para que los teléfonos celulares sean comercialmente exitosos, necesitan ser capaces de operar en cualquier ciudad con servicio celular. La mayoría de las empresas tienen acuerdos con otras para proporcionar a sus clientes servicio cuando viajen a otras ciudades. Cuando un cliente celular viaja a otra ciudad y usa el servicio de una empresa diferente, el proceso es llamado *roaming*. Mientras el suscriptor está en la ciudad foránea, la empresa anfitriona efectúa la tarificación del *roaming* y la envía a la empresa del cliente. En las áreas geográficas donde hay una gran cantidad de tráfico *roaming* usan una técnica llamada *verificación dinámica*. Los portadores cambian la información del *ESN* para asegurar que los clientes viajeros son suscriptores válidos.

1.1.8 Interferencia de multi - trayectoria

Una de las fuentes de interferencia más importantes, es un fenómeno denominado Interferencia de Multi - trayectoria. Esta interferencia es similar en causa al eco acústico. En la figura 1.4 vemos un vehículo recibiendo una señal directa y una que es reflejada por un edificio. La señal directa toma un cierto tiempo para llegar a la unidad móvil. Este tiempo es función de la distancia desde la antena de la estación celular a la antena del vehículo. La señal reflejada tiene una trayectoria más larga y toma más tiempo para llegar a la antena del vehículo. Esto crea un eco que puede causar problemas significativos en radio transmisión digital.

Regularmente las multi-trayectorias no son un problema en sistemas de una portadora por canal, donde la tasa de transmisión digital es baja. En sistemas de comunicaciones usando multicanalización por división de tiempo con modulación digital, donde se requieren altas velocidades, el eco causará errores, resultando en una limitación en la velocidad de señalización. La tasa de error no puede reducirse mejorando la relación portadora – ruido o la relación portadora – interferencia.

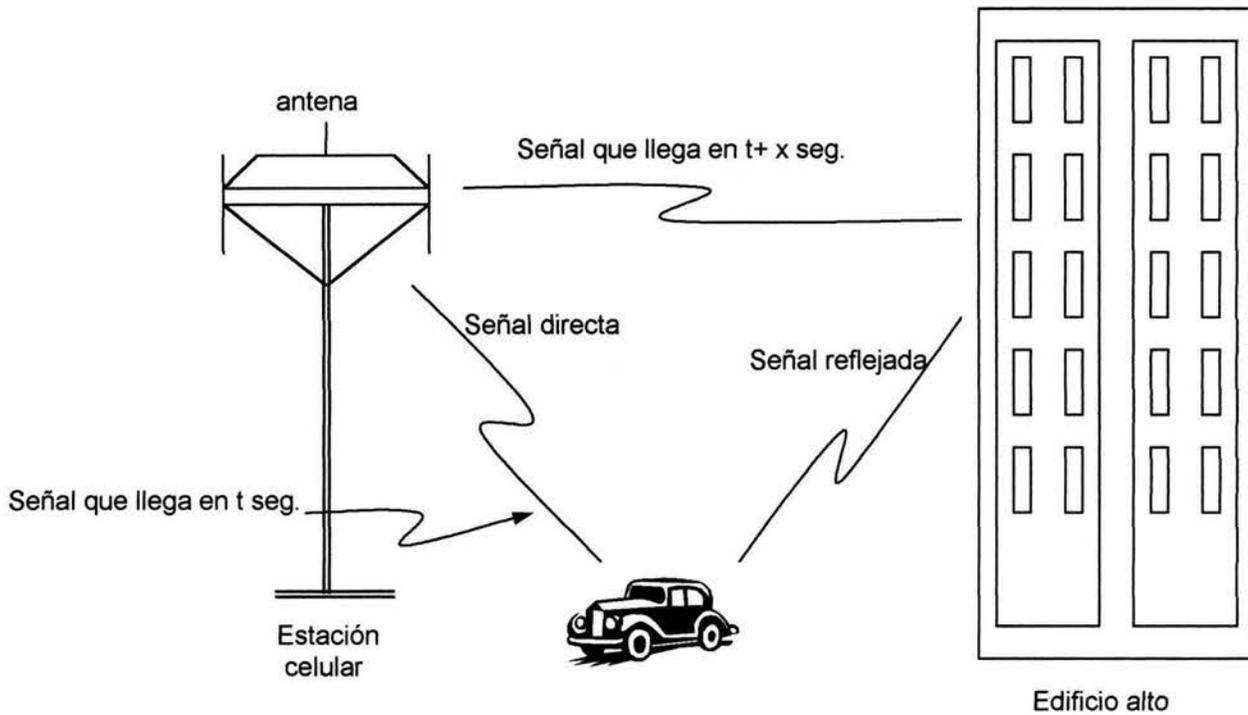


Figura 1.4 Interferencia multi – trayectoria.

1.1.9 Control de potencia

En los sistemas celulares de radio, los niveles de potencia transmitida por cada unidad de los clientes, están bajo un control constante por las estaciones base servidoras. Esto se hace para asegurar que cada móvil transmite la potencia más baja necesaria. El control de potencia no sólo hace que dure más la batería, sino que también reduce mucho la S/I del canal inverso.

1.1.10 División de celdas

El "*splitting*" es el proceso de subdividir una celda congestionada en celdas más pequeñas, cada una con su propia estación base y la correspondiente reducción en la altura de la antena y de la potencia de transmisión. El "*splitting*" incrementa la capacidad de un sistema celular dado que incrementa el número de veces que se reutilizan los canales. Definiendo nuevas celdas que tengan un radio más pequeño que las celdas originales instalando estas pequeñas celdas entre las celdas existentes, se incrementa la capacidad debido al incremento de canales por unidad de área.

Con este punto de partida, la telefonía celular se optó como una alternativa a la telefonía inalámbrica existente en varios países y no tuvo que pasar mucho tiempo para que se empezara a saturar el servicio obligando a desarrollar e implantar nuevas formas de acceso múltiple al canal y

convirtiendo los sistemas analógicos a digitales. Para diferenciar una etapa de otra, la telefonía celular se ha caracterizado por contar con diferentes generaciones. A continuación se describen cada una de ellas.

1.2 Estándares de primera generación

Los sistemas de telefonía móvil de la primera generación tienen muchas similitudes, como una transmisión analógica y un mecanismo de acceso *FDMA* (*Frequency Division Multiple Access*) y como características generales están:

1. Funciones típicas de un teléfono; siempre llamadas punto a punto.
2. Más servicios de datos.
3. Circuitos de conmutación: relativamente económicos para conexiones largas, pero relativamente costosos para conexiones cortas.
4. *Roaming* internacional para voz en Europa.
5. Adquisición relativamente económica debido al gran mercado.
6. Cargos de llamada altos que se deben a las altas inversiones en redes.

1.2.1 *AMPS (Advances Mobile Phone Service)*

Este sistema que tuvo sus inicios en Estados Unidos por los años 70's no logró un mercado comercial hasta 1983, pero con un gran éxito ya que más de 50 millones de sistemas *AMPS* están en uso a través del mundo e implementado en aproximadamente 250 redes.

AMPS fue estandarizado como *EIA/TIA 553* y hecha obligatoria para las redes de telefonía móvil en los Estados Unidos por la *FCC* (*Federal Communications Commission*). El punto de partida de esto fue que el *roaming* debe ser posible entre redes, aunque esto no fue posible verdaderamente hasta los 90's.

Este sistema trabaja en la banda de frecuencia de 800 [MHz], donde para la transmisión de la estación móvil a la estación base utiliza frecuencias de 824 [MHz] a 849 [MHz] y de 869 [MHz] a 894 [MHz] en sentido inverso.²

El espectro localizado para *AMPS* es compartido por dos portadoras celulares en cada área o región. Cada portadora divide el espectro en canales, utilizados para comunicar desde las estaciones base en las celdas hasta los dispositivos móviles, y canales utilizados para la comunicación entre los dispositivos móviles y las estaciones base. Los canales son divididos en

² BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.151-152

canales de voz de 30 [KHz] que emplean Modulación de Frecuencia para transmitir la voz con una potencia de la terminal móvil de 3 [W].³

Narrowbandamps (NAMPS) es una variante de *AMPS* desarrollado por *Motorola* a los inicios de los 90's y la cual fue estandarizada por la *TIA* como *IS-88*, *IS-89* y *IS-90*. En este tipo de sistema el canal de voz es reducido de 30 [KHz] a 10 [KHz] con lo que se incrementa la eficiencia en frecuencia por un factor de 3. Además es incorporado el servicio de recepción de mensajes digitales derivado del sistema *GSM's SMS*. Este estándar es compatible con *AMPS* por lo que es posible el continuo uso de la infraestructura de la red incluyendo las licencias por el uso del espectro. Sin embargo el sistema *NAMPS* puso altos criterios para la realización de las terminales en términos de estabilidad y un receptor selectivo, lo que resulto en terminales más caras. El éxito de *NAMPS* fue limitado, particularmente porque los sistemas digitales alcanzaron el mercado que daba el mismo incremento de eficiencia que *NAMPS*, como *DAMPS (Digital Advanced Mobile Phone System)*, del cual se hablara más adelante. Diez redes *NAMPS* se encontraban en uso a través del mundo a los finales de 1996, con un total aproximado de 1.5 millones de subscriptores. En realidad la mayoría de esas redes son combinaciones de redes *AMPS/NAMPS* en las cuales solo una pequeña porción de los canales ha sido actualmente convertido a *NAMPS*.

Actualmente el sistema *AMPS* se encuentra evolucionando paulatinamente al sistema *DMAPS* el cual coexiste con el *AMPS*, en modo dual y funciona con acceso *TDMA (Time Division Multiple Access)*.

1.2.2 NMT-450 (Nordic Mobile Telephone 450 MHz)

La iniciativa para la creación del sistema *NMT* fue tomada por el Consejo Nórdico con la permanente cooperación de los gobiernos Danés, Finlandés, Islandés, Noruego y Sueco. Las especificaciones fueron presentadas en 1975 durante la Conferencia de Telecomunicaciones Nórdicas. Todos los fabricantes estaban permitiendo el diseño y fabricación de sistemas que cumplían con el estándar que incluía descripciones de la interfaz de radio y muchas otras interfaces incluyendo el enlace entre la estación base y un conmutador.

Debido a que *NMT* estaba siendo introducida en todos los países Escandinavos, el *roaming* era uno de los puntos de salida para el diseño. En Estocolmo durante 1978, los diseñadores demostraron un modelo inicial trabajando. La red nórdica *NMT* fue oficialmente puesta en uso el 1 de octubre de 1981. Además un poco después en Septiembre del mismo año, la primera red mundial *NMT* había sido puesta en operación en Arabia Saudita. Lo que quiere decir que el

³ MEDINA N., Guadalupe, Telefonía Digital, Ingenierías, Abril – Junio 2001 Vol. IV No.11 p.53

estándar había sido exportado antes de llegar a ser posible en Escandinavia. Esta tendencia se extendió en muchos países europeos que adoptaron el sistema, incluyendo Los países bajos (1985), Luxemburgo (1985) y Bélgica (1987). *NMT-450* fue también construida en Francia y España y durante los 90's en muchos países europeos.

Sin embargo muchos estándares más modernos están ya disponibles, los equipos están aun siendo cambiados para expandir las redes existentes y hasta construir unas nuevas en países en expansión.

El estándar *NMT-450* es un sistema celular analógico que usa red de arquitectura jerárquica. Pares de frecuencias son usadas para la comunicación entre los teléfonos móviles y las estaciones base. Las terminales móviles transmiten en la banda de frecuencia de 450 [MHz] y recibe en 460 [MHz], con una separación duplex de 10 [MHz]. Dos secciones del espectro de 4.5 [MHz] son usadas en estas redes, divididas en canales de 25 [KHz]. Celdas relativamente largas son utilizadas con diámetros mayores de 50 [Km]. Comparado con el estándar americano *AMPS*, el sistema *NMT* era más avanzado desde una perspectiva técnica, por ejemplo, estaba basado en una moderna tecnología de conmutación digital y con estaciones base relativamente inteligentes diferentes a los utilizados en los sistemas *AMPS*, las cuales casi todas las funciones eran computadas y concentradas en un conmutador central, lo cual produce que muchos procesos pueden tomar lugar inmediatamente. Mientras que cada acción en *AMPS* requiera contacto con los controles centrales, un método que demanda muchas más transmisiones dentro de la red. ⁴

1.2.3 TACS (*Total Access Cellular System*)

La telefonía móvil fue introducida en el Reino Unido después que en los países escandinavos. Un grupo de expertos que se encontraban reunidos en 1981 tomaron la decisión de que debían ser dos proveedores de servicios móviles: *British Telecom*, el cual llamaba a su sistema *CELLNET*, mas un operador del sector privado, *Recal/Milicom* (ahora conocido como *Vodafone*) los que dieran un nuevo sistema para el Reino Unido. Varios estándares eran considerados incluyendo una propuesta alemana y una francesa, tan buena como el *NMT-450*, pero presiones realizadas por *Vodafone* resultaron en una selección de una variante del sistema americano *AMPS* y le dieron el nombre de *Total Acces Communications System*. Lo que significa un beneficio en el costo. Sin embargo, el sistema *AMPS* usa una frecuencia mucho más alta que los sistemas Europeos en ese tiempo. Por lo tanto la banda de los 900 [MHz] fue seleccionada por los británicos como una variante, y este espacio fue reservado para la telefonía móvil. Una de las estrictas condiciones en la licencia concedida a ambos operadores de redes, 15 [MHz] del espectro, fue que las redes

tendrían que estar comercialmente disponibles desde Abril de 1985. Bajo las presiones de la competición, ambas redes abrieron en enero de 1985, tres meses antes de la fecha a la cual estaban obligados para la licencia. Posteriormente Irlanda siguió con su red TACS.

Desde una perspectiva técnica, el cambio más grande esta en el ancho del canal, el cual fue reducido de 30 [KHz] en AMPS a 25 [KHz] para TACS, mientras que la banda de frecuencia cambio de 800 [MHz] a 900 [MHz], con una frecuencia de transmisión de la terminal móvil de 890 [MHz] - 915 [MHz] y una frecuencia de transmisión de la estación base de 935 [MHz] – 960 [MHz].⁵

1.2.4 NMT-900

Sucesor del estándar NMT-450, el primer sistema llegó a operar en Escandinavia a finales de 1986. En Europa, ha sido implementado en más países que solo en Los países bajos y Suiza y el *roaming* fue ofrecido entre esas redes. La mayoría de los parámetros cambiaron a los de su predecesor: dos secciones de 25 [MHz] son usadas y la separación dúplex es incrementada a 45 [MHz], el espaciado del canal es de 12.5 [KHz], logrando colocar más usuarios en el espectro limitado. El sistema también usa celdas más pequeñas, con un diámetro aproximado que van de 0.5 a 3 [Km]. Los teléfonos de bolsillo fueron posibles por primera vez teniendo una potencia de transmisión de 1 [W].⁶

1.2.5 El estándar japonés NTT

Japón fue de los primeros países en tener en operación sistemas de telefonía celular, esto en el año de 1979. Su verdadero nombre es MSCL1, sin embargo es mejor conocido por NNT debido a que su elaborador y operador es *Nippon Telegraph and Telephone Corporation*. El sistema usa dos porciones de 15 [MHz] en la banda de 800 [MHz] y un espaciamiento de canal de 25 [KHz]. Cuando la capacidad llegó al tope, lo cual se vio a mediados de 80's, NTT desarrollo una nueva versión (MSCL2), con una eficiencia del espectro más grande el cual, sin embargo, fue también diseñado para operar en el mismo espectro de 2 x 15 [MHz]. Esta segunda versión aun no permitía teléfonos de bolsillo. El ancho del canal fue reducido a 12.5 [KHz] y una reducción de interferencia, *interleaving* es aplicado también entre celulas individuales, asignando frecuencias en pasos de 6.25 [KHz] en cambio de 12.5 [KHz], como se muestra en la figura 1.5

⁴ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.156-149

⁵ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.150-151

⁶ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. p.149

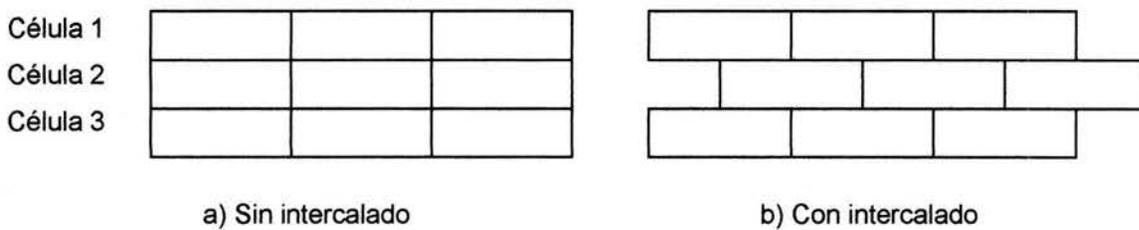


Figura 1.5 Intercalado del espaciado de canales.

Un problema con este tipo de sistemas es que necesitan una estabilidad en la frecuencia de síntesis. Un valor común de esta estabilidad es una desviación de 1 ppm, lo cual significa para una frecuencia de transmisión de 800 [MHz], que la desviación será aproximadamente de 800 [Hz]. Esto es más del 10% de espaciado del canal de 6.25 [kHz]. El problema de estabilidad de la portadora apenas juega un papel en los sistemas digitales.

El sistema *NNT* no está en uso fuera de Japón. Con la segunda generación de los sistemas de celular Japón de nuevo no tendrá éxito en exportar su estándar *PDC*.⁷

1.2.6 C-450 o "C-NETZ" (German Digital System)

A los inicios de los años 80's, se empezaron trabajos en Alemania para el desarrollo del estándar *C-450*, un estándar de la comunicación celular en la banda de los 450 [MHz]. Este sistema es también conocido con el nombre dado por su operador Deutsche Bundespost a su red: *C-NETZ*. Esta red fue abierta al público en 1985 y *Siemens* fue el proveedor del servicio. El verdadero éxito en vender el *C-450* a redes de otros países no fue venidero. Por el limitado tamaño solo ha sido construido en Portugal y Sudáfrica.

Este estándar fue creado con una alta capacidad y un área densamente poblada en mente. El dueño asumió una capacidad menor a 300,000 suscriptores en el oeste de Alemania, el cual resultó en un sistema con radios de las células de 20 [Km] promedio y un mínimo de 5.4 [Km] y agrupamientos de siete células. Ya a finales de 1996, con sus 605,00 suscriptores, esta red, es por mucho más larga que otras redes trabajando en la banda de los 450 [MHz]. Esto se traduce en un éxito con un limitado espectro asignado. De dos 4.5 [MHz], su flexibilidad fue un punto importante, el espaciado de canales de 10 y 25 [KHz] fue posible y el intercalado era una opción.

⁷ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. p.153-154

Una especial característica de este estándar es que cuando fue desarrollado, se asumió que podría estar disponible un servicio de comunicación de datos de 4.8 [Kbit/s]. Haciendo que el C-450 fuera el primer sistema de comunicaciones móviles con este tipo de servicios integrales, por cierto, este servicio nunca fue implementado.⁸

1.2.7 El estándar francés *Radiocom 2000 (RC-2000)*

El estándar fue desarrollado por *MATRA* y se ha usado en Francia desde 1981. Este es un híbrido formado de los sistemas de radiocomunicaciones y de telefonía móvil. La banda de frecuencias utilizada es de 200 [MHz] para comunicaciones locales. Los 900 [MHz] fue adherida más tarde y por último este sistema no esta en uso fuera de Francia.⁹

1.2.8 *RMTS (Italian System)*

El primer sistema de telefonía celular italiano fue de acuerdo al estándar *RMTS*. En 1983 la red ya daba servicio al público con tecnología desarrollada por *Italtel*, un gran proveedor italiano de comunicaciones. El sistema usaba la banda de los 450 [MHz] y tenía una excelente calidad de transmisión. Sin embargo tenía una capacidad máxima de 70,000 subscriptores. Las terminales móviles eran fabricadas por *Italtel* y *Italian Telletra*, y tenían que ser arrendadas desde el operador de red. *RTMS* no permitió el uso de teléfonos de bolsillo y no fue utilizada fuera de Italia.¹⁰

1.2.9 *MATS-E*

Este estándar es una unión desarrollada por varias compañías y el grupo *Alcatel*. El éxito de este estándar fue muy limitado y la tecnología fue usada solo en Francia y Kuwait. *MATS-E* trabaja tanto en la banda de frecuencia de 450 [MHz] como en la banda de 900 [MHz].¹¹

Para terminar con la primera generación a continuación se expone la tabla 1.1 comparativa de los parámetros técnicos de los sistemas de la primera generación de la telefonía celular.

⁸BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. p.155

⁹BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. p. 156

¹⁰ idem

¹¹ idem

| | AMPS | NMT-450 | TACS | NMT-900 | NTT | C-450 | RADIOCO M 2000 | RMTS |
|--|---------------|--------------------|---------------|--------------------|-----------------|-------------------------------|----------------------|-------------|
| País de origen | U.S.A | Escandinavia | Reino Unido | Escandinavia | Japón | Alemania | Francia | Italia |
| Primer red comercial | 1983 | 1981 | 1985 | 1986 | 1979 | 1986 | 1981 | 1983 |
| Banda de frecuencia | 800 [MHz] | 450 [MHz] | 900 [MHz] | 900 [MHz] | 800 [MHz] | 450 [MHz] | 200, 420 y 900 [MHz] | 450 [MHz] |
| Frecuencia de transmisión de la terminal móvil | 824-849 [MHz] | Banda de 450 [MHz] | 890-915 [MHz] | Banda de 890 [MHz] | 860-885 [MHz] | 461-465 [MHz] | | |
| Frecuencia de transmisión de la estación base | 869-894 [MHz] | Banda de 460 [MHz] | 935-960 [MHz] | Banda de 935 [MHz] | 915-940 [MHz] | 451-455 [MHz] | | |
| Ancho de banda disponible | 2 x 25 [MHz] | 2 x 4.5 [MHz] | 2 x 15 [MHz] | 2 x 15 [MHz] | 2 x 15 [MHz] | 2 x 4.5 [MHz] | 28 [MHz] | |
| Separación Dúplex | 45 [MHz] | 10 [MHz] | 45 [MHz] | 45 [MHz] | 55 [MHz] | 10 [MHz] | | |
| Ancho del canal | 30 [kHz] | 25 [kHz] | 25 [kHz] | 12.5 [kHz] | 25 y 12.5 [kHz] | 20 [kHz], 10 [kHz] y 25 [kHz] | 12.5 [kHz] | |
| Potencia de la terminal portátil | 3 [W] | 15 [W] | 4 o 10 [W] | 6 [W] | 5 [W] | 15 [W] | | |
| Potencia del teléfono de bolsillo | 0.6 [W] | No fue posible | 0.6 [W] | 1 [W] | No fue posible | No fue posible | No fue posible | No fue pos. |

Tabla 1.1 Comparación de parámetros técnicos de los estándares analógicos.

1.3 Estándares de Segunda generación

A los sistemas de telefonía digital, es a lo que se le conoce como segunda generación y su ingreso fue en el año de 1990, estos sistemas utilizan protocolos de codificación más sofisticados y se emplea en los sistemas de telefonía celular actuales, a continuación se nombran algunos avances generales de este tipo de sistemas contra los primeros.

1. Una eficiencia más alta en el uso del espectro, esto debido a la utilización de codificación de voz y un reutilización más agresivo de frecuencias.
2. Sistemas con capacidad más alta con relación al costo.
3. Un costo más bajo en la infraestructura y terminales por:
 - a. Aplicación de circuitos integrados.
 - b. Espacios de tiempo permitiendo a los usuarios compartir el canal.
 - c. Criterios más bajos en la relación señal a ruido para la transmisión digital.
4. Buena integración en la red digitalizada *ISDN*.
5. Nuevos servicios, como comunicación de datos, fax y *SMS (Short Message Service)*.
6. Mejor privacidad y facilidades en la encriptación de datos.

1.3.1 GSM (*Global System for Mobile communications*)

A principio de los 80's, después de que el sistema *NMT*, funcionara con éxito, fue obvio para varios países europeos que los sistemas analógicos existentes, tenían limitaciones. Primero, el potencial demandado de servicios móviles fue mayor a la capacidad esperada de las existentes redes analógicas. Segundo, las diferentes formas de operación no ofrecían compatibilidad para los usuarios de móviles: una terminal no podía acceder dentro de una red *NMT*, y viceversa. Además, el diseño de un nuevo sistema de telefonía celular requiere tal cantidad de investigación que ningún país europeo podía afrontarlo de forma individual, lo que provocó la idea de un nuevo sistema, desarrollado entre varios países.

Ya dado que el requisito principal para un sistema de radio, es decir, el ancho de banda ya estaba resuelto con la reserva de la banda de frecuencias de 900 ± 25 [MHz] para comunicaciones móviles en 1978, solo quedaba organizar el trabajo.

Todos estos factores, llevaron a la creación en 1982 de un nuevo cuerpo de estandarización dentro del *CEPT (Conférence Européenne des Postes et Télécommunications)*, cuya tarea era especificar un único sistema de radiocomunicaciones para Europa a 900 [MHz]. El recién Nacido *Groupe*

Special Mobile (GSM) tuvo su primer encuentro en Diciembre de 1982 en Estocolmo, bajo la presidencia de Thomas Haug de la administración sueca. Treinta y una personas de once países estuvieron presentes en este primer encuentro. En 1990, por requerimiento del Reino Unido, se añadió al grupo de estandarización la especificación de una versión de *GSM* a la banda de frecuencia de 1800 ± 75 [MHz], a esta variante se le llamó *DCS1800 (Digital Cellular System 1800)*. El significado actual de las siglas *GSM* ha cambiado y en la actualidad estas corresponden a *Global System for Mobile communications*.

Algunos de los propósitos del sistema estaban claros desde el principio: uno de ellos era que el sistema debía permitir la libre circulación de los clientes en Europa, el ya mencionado *roaming*. Y también estaba claro también que la capacidad ofrecida por el sistema debería ser mejor que las existentes redes analógicas.

En 1982, los requerimientos básicos para *GSM*, estaban establecidos. Éstos fueron revisados ligeramente en 1985, quedando principalmente como siguen:

Servicios:

- El sistema será diseñado de forma que las estaciones móviles se puedan usar en todos los países participantes.
- El sistema debe permitir una máxima flexibilidad para otros tipos de servicios, como los servicios relacionados con la *RDSI*.
- Los servicios ofrecidos en las redes *PSTN* e *ISDN*, así como otras redes públicas deben ser posibles, en la medida de las posibilidades, en el sistema móvil.
- Debe ser posible la utilización de las estaciones móviles pertenecientes al sistema a bordo de barcos, como extensión del servicio móvil terrestre. Se debe prohibir el uso aeronáutico de las estaciones móviles *GSM*.
- En lo referente a las estaciones, a parte de las montadas en vehículos, el sistema debe ser capaz de suministrar estaciones de mano así como otras categorías de estaciones móviles.

Calidad de los servicios y seguridad:

- Desde el punto de vista del cliente, la calidad de voz telefónica en el sistema *GSM* debe ser al menos tan buena como la que tenía la primera generación de sistemas analógicos a 900 [MHz].

- El sistema debe ser capaz de ofrecer encriptación de la información del usuario pero debe permitir la posibilidad de que esto no influya en el costo de aquellos clientes que no requieran este servicio.

Utilización de la radio frecuencia:

- El sistema permitirá un gran nivel de eficiencia espectral así como la posibilidad de servicios para el cliente a un costo razonable, teniendo en cuenta tanto las áreas urbanas como rurales y el desarrollo de nuevos servicios.
- El sistema permitirá la operación en el rango de frecuencias comprendido entre los 890-915 [MHz] y entre los 935-960 [MHz].
- El nuevo sistema de 900 [MHz] para comunicaciones móviles del *CEPT*, debe coexistir con los anteriores sistemas en la misma banda de frecuencias.

Aspectos de Red:

- La numeración del plan estará basada en la recomendación correspondiente del *CCITT*.
- El diseño del sistema debe permitir diferentes estructuras de carga y velocidades para su utilización en diferentes redes.
- Para la interconexión de los centros de conmutación y los registros de localización, se usará un sistema de señalización internacionalmente estandarizado.
- No se debe requerir ninguna modificación significativa de las redes públicas fijas.
- El sistema *GSM* debe habilitar la implementación de la cobertura común de las redes públicas móviles terrestres *PLMN* (*Public Land Mobile Network*).
- La protección de la información y el control de la información de la red debe ser proporcionada por el sistema.

Costos:

- Los parámetros del sistema deben ser escogidos teniendo en cuenta un costo límite del sistema completo, principalmente el de las unidades móviles.

Desde el principio quedó claro, de forma extraoficial, que el sistema debía estar basado en una transmisión digital, y que la voz debería estar representada por una cadena digital a una velocidad de 16 [kbit/s]. La decisión oficial, sin embargo, no fue tomada hasta 1987.

Desde 1984 hasta 1986, el GSM se encargó de comparar las diferentes posibilidades técnicas para la transmisión (digital o analógica), en particular en sus respectivas eficiencias espectrales (cuál aprovechaba mejor el espectro de frecuencias). Se decidió comparar varias propuestas técnicas de prototipos permitidos en la actual transmisión de radio. En 1985, las administraciones francesas y alemanas de Correos y Telégrafos unieron sus esfuerzos para realizar cuatro estudios que condujeran a otros tantos prototipos. La prueba comparativa de ocho prototipos, se realizó en Diciembre de 1986 en los laboratorios del CNET (*Centre National d'Etudes des Télécommunications*) cerca de París, bajo el control del Núcleo Permanente. Todos estos prototipos hicieron uso de la transmisión digital, y la mayoría fueron propuestos por compañías de teléfonos.

Los resultados de las comparaciones se publicaron a comienzos de 1987. Hubo grandes discusiones sobre el modelo que iba a ser escogido, debido a que podía afectar al prestigio de la empresa que lo propuso. Sólo se decidieron las características del método de transmisión, las cuales fueron las siguientes:

Tamaño medio de la banda de transmisión (200 [KHz] de separación de portadoras), en comparación con los sistemas de banda estrecha (12.5 ó 25 [KHz] que existían en los sistemas analógicos) o con los sistemas de banda ancha (uno de los candidatos propuso una separación de portadoras de 6 [MHz]); Transmisión digital de voz a una velocidad no superior a 16 [kbit/s]; Multiplexación en el tiempo de orden 8, con una evolución en el futuro hacia la multiplexación de orden 16 cuando se defina un codificador de voz a la mitad de velocidad; *Hopping* de frecuencias lento, que consiste en cambiar la frecuencia usada por un canal a intervalos regulares de tiempo. En GSM la frecuencia de transmisión permanece constante durante la transmisión de una trama completa. Esta técnica procede de los sistemas de transmisión militares, y se decidió incluirla en las principales características de la transmisión de radio de GSM, además de utilizarla por motivos de seguridad, también para conseguir una mayor diversidad de frecuencias, y para paliar los efectos de los desvanecimientos de tipo *Rayleigh*.¹²

La red GSM tiene una ruta digital de radio, se transfieren bits a través del interfaz Aire. Debido a este formato de bits, la información transferida emplea un espectro de frecuencias considerablemente ancho: cuanto mayor sea la velocidad de bits usada, se necesitará un espectro más ancho. Si el espectro es ancho se reduce la capacidad del interfaz Aire.

¹²http://ceres.ugr.es/~alumnos/c_avila/gsm0.htm

Esta es la razón por la que el sistema de codificación de la voz que se utiliza en las redes *PSTN/ISDN* no es adecuada como tal a la interfaz Aire de *GSM*. Las redes *PSTN/ISDN* utilizan el formato de codificación *Ley-A* generando 64 [kbits/s] de flujo de información digital por cada canal de tráfico. Si este flujo de bits se transfiriese través del Interfaz Aire, el espectro de frecuencias sería tan ancho que no sería razonable.

Para mantener la capacidad del interfaz Aire, el sistema *GSM* introduce otra forma de codificar y decodificar la voz. El sistema de codificación utilizado se llama *RPE-LTP (Regular Pulse Excitation - Long Term Prediction)* que es una forma más efectiva de convertir voz en bits. Utilizando este sistema de codificación, la codificación de voz ocupa 13 [kbit/s] en lugar de 64 [kbit/s].¹³

El estándar *GSM* será estudiado y detallado en el siguiente capítulo.

1.3.2 *DCS-1800*

Este estándar fue desarrollado en 1990 a petición de Reino Unido y es un derivado de *GSM*. El punto más importante de salida cuando consideraron un nuevo estándar fue un mercado masivo, es decir, que pudiera ser usado en áreas altamente pobladas. Se pretendió ofrecer alta capacidad y una alta calidad a un precio económico y la introducción de teléfonos más ligeros en el mercado.

Para la capacidad requerida, la nueva red debe usar una banda entre 1.7 y 2.3 [GHz]. La red tuvo que usar células mucho más pequeñas, sobretodo en áreas densamente pobladas. Después de haber comparado varias opciones tecnológicas, Reino Unido a petición de *ETSI* desarrollaron una variante de *GSM*, así solo cambios mínimos fueron seleccionados, los cuales consistieron principalmente en limitaciones en la interfaz de radio. Las otras interfaces permanecieron idénticas al sistema *GSM*, lo cual significa, entre otras cosas, que existiendo conmutadores *GSM* este sistema puede ser usado. Solo cuatro redes *DCS-1800* fueron construidas entre 1993y 1997, dos de esas en el Reino Unido, una en Alemania y la otra en Francia.

Si se compara *DCS-1800* con *GSM*, los cambios técnicos son:

Una nueva banda de frecuencia (1800 [MHz]), de la cual el estándar usa dos secciones de 75 [MHz] del espectro, tres veces más que en *GSM*, ya que solo tiene dos secciones de 25 [MHz] en la banda de 900 [MHz]. Estas frecuencias son de 1710 a 1785 [MHz] y 1805 a 1888 [MHz], con la

¹³ <http://neutron.ing.ucv.ve/revista-e/No2/Gsmseuly.htm>

separación dúplex de 95 [MHz]. Al igual que GSM los canales son de 200 [kHz] de ancho, obteniendo 374 canales disponibles. El salto de frecuencias es también implementado ligeramente diferente.

Para este sistema ETSI después de analizar escenarios operacionales opto por manejar dos clases de potencia de transmisión de la terminal dependiendo del tamaño de la célula, así se escoge una potencia de 1[W] para células de 0.7 a 19 [Km] y de 0.25 para células de 0.5 a 13 Km].

El *roaming* internacional y el nacional ahora son posibles y no esta limitado a escoger una entrada de red, pero tampoco a seleccionar parte de una red. De esta manera, dos operadores nacionales de redes de DCS-1800 pueden compartir los costos de la infraestructura y construir cada uno solo parte de la red. Esto no es posible en GSM. Además tanto el *roaming* nacional como el internacional son solo posibles si los dos operadores de red están de acuerdo con ello.

Lo anterior es solo una fase del sistema que corre paralelamente con GSM. Sin embargo importantes componentes de la segunda fase son adheridas: Una extensión a los servicios de mensajes cortos (SMS), más facilidades para un *roaming* nacional, más algoritmos de encriptación para facilitar la exportación fuera de Europa, nuevos servicios como: llamada en espera, identificación de llamadas, tres a la vez, etc.

Con lo que respecta a la codificación, este sistema puede utilizar el código *half-rate* de GSM. Esto debido a las pequeñas células y el gran espectro disponible. Sin embargo aumentando a una codificación *full-rate* hay una mayor probabilidad de éxito.¹⁴

1.3.3 PCS-1900

Este estándar también es derivado de GSM y fue el principal objetivo de los operadores de redes que ganaron la licencia en la subasta de los PCS en los Estados Unidos, su estandarización es propuesta como TAG5. Muchos cambios han sido hechos con respecto al estándar DCS-1800 para adaptar el sistema al mercado de los Estados Unidos. De esta manera cumple con el criterio de transmisión y señalización y muchos estándares internacionales ahora tienen variantes norteamericanos. El uso de muchos espacios de tiempo en una comunicación simple ha sido incluida en esta estándar, de la misma manera han sido los algoritmos de compresión de voz que soporta. Sin embargo aparecería el algoritmo EFR desarrollado específicamente para el mercado de este estándar, pero también ha sido estandarizado para GSM y DCS-1800.

1.3.4 D-AMPS

A mediados de los 80's los operadores de las redes americanas se vieron en la necesidad de incrementar la capacidad de sus redes *AMPS*, lo que dio como resultado el estándar *D-AMPS*, entre otras cosas. Las primeras investigaciones las empezó la *CTIA (Cellular Telecommunication Industry Association)*, el cual publicó los primeros parámetros en 1988, posteriormente el desarrollo fue transferido a *TIA (Telecommunication Industry Association)*, una organización responsable de la estandarización en el campo de las telecomunicaciones y la cual concluyó el primer estándar en 1991 bajo el nombre de *IS-54*. Realmente este estándar es implementado dentro de un sistema *AMPS*, en el cual un cierto número de canales son digitalizados y de hecho como muchos de los usuarios aun tienen terminales analógicas, el estándar desarrolla un modo dual (*AMPS/D-AMPS*) en terminales obligatorio para que las llamadas puedan también ser hechas en células o redes sin canales digitales.

La meta original del estándar *D-AMPS* era incrementar la eficiencia de la frecuencia en un factor de tres. Para lograr este fin, los canales en *AMPS* de 30 [KHz] son divididos en tres espacios de tiempo en los cuales tres diferentes terminales pueden transmitir voz digitalizada, sin olvidar que la versión original, *IS-54* aun utiliza la señalización existente en los sistemas analógicos *AMPS*. Cuando una conexión es establecida, la terminal indica si prefiere una conexión digital. *D-AMPS* también reconoce otras clases de potencia, como por ejemplo 3 [W], 1.2 [W] y 0.8 [W], siendo el valor de 1.2 [W] el más común.

Debido a la división en tres de los canales de voz, este sistema solo permite una relativamente baja velocidad de datos, por lo que utiliza una muy eficiente codificación de voz, esta es una codificación de predicción lineal *Vectorsum (VSELP)*, el cual utiliza una tasa de bits de 7.95 [kbit/s]. Una compresión rígida resulta, por lo tanto, en una considerable baja calidad de voz comparada con un sistema analógico. Además esta compresión no permitía el uso de modems. Esta desventaja fue dirigida en una revisión posterior, *IS-136*, terminada en 1994. Esta revisión usa un nuevo sistema de señalización e integra servicios de datos y fax a 22.8 [kbits/s], además de un ahorro de energía, introduciendo el *sleep mode*, en la cual una terminal no necesita un monitoreo constante en el canal de señalización y hay un servicio de mensajes cortos como en *GSM*. Otra nueva característica, es un seguro soporte para macro, micro y picocélulas para una estructura jerárquica de células. Ahora no hay una obligación de *handover* cuando una terminal entra o deja una célula, como demanda un sistema convencional sin una estructura jerárquica de células.

¹⁴ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. p 193-196

Un canal adaptable tiene una asignación automática seleccionando la mejor frecuencia para una célula y por lo tanto hace que la planificación de frecuencias sea innecesaria. Un *mobile-assisted hand-off* es usado, lo cual significa que una terminal continuamente mide la fuerza del campo de la célula adyacente y pasa los resultados de esas medidas a la red. La red decide si inicia o no una entrega con base en los datos colectados. Esto puede también tomar lugar entre diferentes modos o bandas, por lo que es posible por lo tanto para conmutadores entre 800 [MHz] de *AMPS*, 800 [MHz] de *D-AMPS* y 1900 [MHz] de *D-AMPS* durante una llamada.

1.3.5 CDMA (IS-35)

CDMA es una técnica de acceso múltiple especificada por la *TIA* como *IS-95*. En marzo de 1992, la *TIA* estableció el subcomité *TR 45.5* con la finalidad de desarrollar un estándar de telefonía celular digital con espectro extendido. En julio de 1993, la *TIA* dio su aprobación al estándar *CDMA IS-95*.

Los sistemas *IS-95* dividen el espectro en portadoras de 1.25 [MHz]. Unos de los aspectos únicos de *CDMA* son que a pesar de que existe un número fijo de llamadas telefónicas que pueden ser manipuladas por un proveedor de servicios de telefonía (*carrier*), éste no es un número fijo. La capacidad del sistema va a depender de muchos factores.

Cada dispositivo que utiliza *CDMA* está programado con un pseudocódigo, el cual es usado para extender una señal de baja potencia sobre un espectro de frecuencia amplio.

La estación base utiliza el mismo código en forma invertida (todos los ceros son unos y los unos ceros) para des-extender y reconstruir la señal original. Todos los otros códigos permanecen extendidos, indistinguibles del ruido de fondo.

Hoy en día existen muchas variantes, pero el *CDMA* original es conocido como *cdmaOne* bajo una marca registrada de *Qualcomm*. A *CDMA* se le caracteriza por su alta capacidad y celdas de radio pequeño, que emplea espectro extendido y un esquema de codificación especial y lo mejor de todo es muy eficiente en potencia.

El sistema *CDMA IS-95* fue desarrollado para la banda de 800 [MHz], pero existen variantes para 1.9 y 2.1 [GHz]. Tiene canales de 1.8 [MHz] de ancho, estos comprenden un canal de voz de 1.25 [MHz] más dos bandas de guardia de 275 [KHz] cada una. Varios canales son usados para la salida y llegada de rutas, es decir, una frecuencia duplex. La separación duplex es de al menos 45 [MHz] y utiliza una separación de 80 o 50 [MHz] para las versiones de 1.8 y 2.1 [GHz] respectivamente.

En este estándar un canal puede manejar entre 25 y 40 llamadas simultáneamente y un avanzado control de potencia es requerido, las terminales continuamente miden la fuerza del campo de la estación base, pero también reciben dos potencias de corrección de la red cada 1.25 [ms]. Una favorable causa – efecto de este tipo de control de potencia es que una terminal nunca usa más potencia que la estrictamente necesaria, con lo cual se hace un efecto de ahorro de energía, con base en una prueba, la cual mostró que una terminal móvil transmite menos que 1 [mW] el 98% del tiempo.

Otro aspecto especial de este estándar es la aplicación de un receptor inclinado. La señal recibida en este tipo de receptores, la cual contiene muchos componentes de tiempo retrasado como consecuencia de la propagación multi - trayectorias es analizada. Para el retraso de la señal recibida primero, varios componentes de la señal pueden contribuir a la señal, en vez de degradarla. Por lo tanto algunos especialistas cuestionan la posibilidad de incluir receptores inclinados en productos de un mercado de masas.

El número de usuarios por canal no es un arreglo estricto. El mayor número de usuarios, el más grande promedio de interferencias por otras terminales y por lo tanto la más grande probabilidad de transmisiones de error, es responsabilidad de un operador de red, quien esta familiarizado con la relación entre el número de errores de transmisión y la calidad de la voz. Para seleccionar un máximo número de usuarios por canal y por lo tanto una mínima calidad. La codificación utilizada es *QSELP (Qualcomm code excited linear predictive)*. Otra codificación de voz disponible en el mercado para calidades de voz extremadamente alta es "*Purevoice*", operando a 13 [kbit/s]. Este estándar hace posible definir una máxima codificación por usuario (13, 8, 6 y 4 [kbit/s]), de esta manera, servicios con diferentes calidades de voz pueden ser asignados con diferentes precios.

Además de la voz, *CDMA IS-95* también ofrece transmisión de datos, fax, mensajes cortos. Estos soportan una velocidad de transmisión arriba a 14.4 [kbit/s] y aun lo más notable es que estos servicios pueden ser utilizados simultáneamente con comunicaciones de voz. Una interfaz de radio es desarrollada con canales de 2.5 [MHz] para soportar servicios de datos arriba a 76.8 [kbit/s].

Es importante poner en claro que *CDMA IS-95* solo describe una interfaz de radio, por lo tanto no es un completo sistema de comunicaciones como *GSM*.¹⁵

1.3.6 PDC

En Japón a inicios de los 90's el sucesor del sistema analógico *NTT* era desarrollado, este sistema inicialmente fue referenciado a un sistema japonés de telefonía celular (*JDC, Japanese Digital Cellular*), el cual lo nombraron posteriormente (*PDC, Pacific Digital Cellular*) y algunas veces también ha sido referido como *Personal Digital Cellular*. Este estándar fue desarrollado por *NTT* en cooperación con fabricantes como *NEC, Lucent, Ericsson* y *Motorola*. Estos fabricantes trabajaron juntos para implantar este estándar en otros países asiáticos, sin embargo en la transmisión de datos, *PDC* no ha sido implementado fuera de Japón y Tailandia es el único país con planes conocidos para la licencia de una red *PDC*.

Este estándar envuelve a un sistema *TDMA* con tres espacios de tiempo, como en *D-AMPS*. La codificación de voz utilizada es *VSELP* con un *bite-rate* de 11.2 [kbit/s] incluyendo la codificación de canal. La encriptación también se aplica. Los tres espacios de tiempo resultan en una modulación de 42 [kbit/s] y un ancho de canal de 25 [kHz]. Esto proporciona una eficiencia alta en el espectro y de acuerdo con algunos especialistas no japoneses, es mejor por más de dos a *GSM*. Debido a que el estándar no obliga a una determinada clase de antena, tanto las estaciones base como la unidad móvil utilizan diversos tipos de antenas. Los servicios de datos y fax son posibles con velocidades de 9.6 [kbit/s].

Dos bandas de frecuencia son usadas: la de 800 [MHz] y la de 1.5 [GHz]. Dos secciones de 16 [MHz] están disponibles en la primera banda, y una frecuencia duplex de 130 [MHz] es usada. Para la banda de 1.5 [GHz] dos áreas son usadas de 24 [MHz] y la frecuencia duplex es de 48 [MHz].

Con lo que se refiere a potencia, este estándar cuenta con cuatro clases de potencia, 3 [W], 2[W], 0.8 [W] y 0.3 [W], siendo la típica para teléfonos de bolsillo la de 0.8 [W].

La codificación de voz *Half-rate* es contemplada en un futuro, en la cual seis espacios de tiempos serán usados, lo que se logrará doblar la capacidad del sistema. Por último servicios de datos más rápidos (64 [kbit/s]) están también contemplados.¹⁶

En la tabla 1.2 podemos observar una comparación de los parámetros técnicos de los estándares digitales mencionados anteriormente.

¹⁵ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. p. 202-205

¹⁶ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. p. 206-207

| ESTÁNDAR | GSM / DCS-1800 / PCS-1900 | D-AMPS | CDMA IS-95 | PDC |
|--|---|--|---|---------------------------------|
| Banda de frecuencia | 900 / 1800 / 1900 | 800/1900 | 800 / 1900 | 800 / 1500 |
| Móvil a estación base | 890-915 / 1710-1785 / 1850-1890 | 824-849 / 1850-1890 | 824-849 / 1850-1890 | 940-956 / 1429-1441 y 1453-1465 |
| Estación base a móvil | 935-960 / 1805-1880 / 1930-1970 | 869-894 / 1930-1970 | 869-894 / 1930-1970 | 810-826 / 1477-1489 y 1501-1513 |
| Separación duplex | 45 / 95 / 80 | 45 / 80 | 45 / 80 | 130 / 48 |
| Método de acceso | TDMA, 8 espacios de tiempo | TDMA, 3 espacios de tiempo | CDMA | TDMA, 3 espacios de tiempo |
| Ancho de canal | 200 kHz | 30 kHz | 1.25 MHz | 25 kHz |
| Espectro disponible | 2 x 25 MHz / 2 x 75 MHz / 2 x 80 MHz | 2 x 25 MHz / 2 x 80 MHz | 2 x 25 MHz / 2 x 80 MHz | 2 x 16 MHz / 2 x 24 MHz |
| Velocidad de modulación | 271 kbit/s | 48.6 kbit/s | 1.2 a 14.4 kbit/s | 42 kbit/s |
| Codificación de voz | LPC, 13 kbit/s | VSELP, 7.95 kbit/s | QSELP | VSELP 11.2 kbit/s |
| Potencia de transmisión del teléfono de bolsillo | 2 W / 1 W / 1 W | 1.2 W | 0.2 W ambas bandas | 0.8 W |
| Detalles especiales | | ACA: Planeación de frecuencia no requerida | No clusters: Planeación de frecuencia no requerida | |
| Técnica de Handover | Asistencia móvil | Asistencia móvil | Soft handover | Asistencia móvil |
| Ecuador adaptivo | Obligatoria | Obligatoria | No requerido | Opcional |
| Técnicas especiales aplicadas | Corrección en el tiempo de sincronización | — | Receptores inclinados y un control de potencia avanzado | — |

Tabla 1.2 Comparación de parámetros técnicos de los estándares digitales.

1.4 Evolución de los estándares y estado actual

Como hemos visto, los sistemas anteriores, aunque ya con algunas innovaciones hacia servicios dirigidos a datos, se han desarrollado fundamentalmente para telefonía. Con la primera y segunda generación actualmente nos encontramos con un mundo compartido por las tecnologías analógica y digital, pero con una clara evolución hacia redes totalmente digitales.

En consecuencia surgen las llamadas generaciones 2.5 y 3 de celulares, la primera se caracteriza por ofrecer propiedades extendidas, ya que cuenta con más capacidades adicionales que los sistemas de la segunda generación, como: *GPRS (General Packet Radio System)*, *HSCSD (High Speed Circuit Switched)*, *EDGE (Enhanced Data Rates for Global Evolution)*, *IS-136B* e *IS-95Bm* entre otros, además de ser un puente más económico para que las redes evolucionen a la tercera generación.

La tercera generación se caracteriza por contener a la convergencia de voz y datos con acceso inalámbrico a Internet, en otras palabras, es apta para aplicaciones multimedia y altas transmisiones de datos. Los protocolos empleados en los sistemas de la tercera generación soportan altas velocidades de información y están enfocados para aplicaciones más allá de la voz como audio (*mp3*), video en movimiento, videoconferencia y acceso rápido a Internet, sólo por nombrar algunos. Se prevé que estos sistemas alcancen velocidades de hasta 384 [kbit/s], permitiendo una movilidad total a usuarios, viajando a 120 [km/h] en ambientes exteriores. También alcanzará una velocidad máxima de 2 [Mbit/s], permitiendo una movilidad limitada a usuarios, caminando a menos de 10 [km/h] en ambientes estacionarios de corto alcance o en interiores.

El primer paso hacia un sistema de la tercera generación se dio con la reserva en el ámbito mundial de la banda de frecuencia para *IMT-2000* adoptada por la Asamblea Mundial de Radiocomunicaciones en 1992.

Las iniciativas de la normalización de la tercera generación se han desarrollado principalmente a *IMT-2000* y *UMTS* la norma global en fase de desarrollo de la *UIT*, contempla tres grupos de velocidades 144 [kbit/s], 384 [kbit/s] y 2 [Mbit/s] y las bandas de frecuencia de 1885-2005 [MHz] y 2170-2200 [MHz], así como 1980-2010 [MHz] y 2170-2200 [MHz] para satélites.

En Japón, los estándares de tercera generación están siendo desarrollados por *ARIB* (radio) y *TTC* (redes). *NTT DoCoMo*, el mayor operador inalámbrico de este país está defendiendo la

interfaz de radio de acceso por división de código de banda ancha *WCDMA* como norma nacional japonesa.

Con respecto a Estados Unidos, este país ha reservado la banda dúplex inferior al espectro de *IMT-2000* de la *UIT* para operadores *PCS* de segunda generación, imposibilitando habilitar los sistemas globales de tercera generación. La adaptación de Estados Unidos de la reserva de frecuencia de *GSM*, *GSM1900*, beneficiará la evolución de *GSM* así como al desarrollo de la segunda generación.

UMTS se perfila como el más prometedor de los sistemas de tercera generación proporcionando transmisión de voz y datos a altas velocidades, rompiendo con las barreras entre telecomunicaciones fijas y móviles y promoviendo la convergencia entre ambas.

Por otro lado, *GSM* esta evolucionando, obteniendo funciones multimedia con lo que soportara hasta 384 [Kbit/s], con movilidad restringida y hasta 115 [Kbit/s] con movilidad total. Debido a la creciente demanda de servicios de datos de valor añadido para teléfonos móviles *ETSI* ha desarrollado nuevas tecnologías de datos sobre redes *GSM* basada en conmutación de circuitos (*HSCSD*) y conmutación de paquetes (*GPRS*).

Para la tercera generación la importancia de las técnicas de conmutación han sido importantes, lo que ha dado lugar al desarrollo de *GPRS* (*General Packet Radio Service*), la cual es una técnica de conmutación de paquetes que emplea una codificación reducida del canal para alcanzar velocidades de 14.4 [kbit/s] por espacio de tiempo consiguiendo una velocidad máxima de 115 [kbit/s]. Esta técnica esta basada en una nueva codificación de los canales de radio, en la asignación de canales distintos para transmisión y recepción. Por lo tanto, *GPRS* es una técnica adecuada para manejar el tráfico de la información como se da en Internet o en redes de área local. Tiene la capacidad para suministrar datos directamente a la terminal del usuario. Si este se encuentra apagado o fuera de cobertura, inmediatamente que le usuario se presente en la red, el sistema automáticamente le indicara que tiene un mensaje en espera y le envía tanto el texto como las imágenes que contiene dicho mensaje. Al tratarse de una técnica de conmutación de paquetes solamente ocupa ancho de banda cuando se envían datos, lo que permite una utilización más eficiente del espectro al permitir compartir un canal entre distintos usuarios en el proceso conocido como multiplexación estadística.

Partiendo de *GPRS* y como un paso intermedio hacia los sistemas de banda ancha de tercera generación, se encuentra *EDGE* (*Enhanced Data Rates for GSM Evolution*) también llamado *GSM348*. Introduce nuevos métodos en la capa física, incluyendo el esquema de modulación *8PSK* y diferentes formas de codificación de datos para la protección de errores. El resultado es

que *EDGE* alcanza velocidades de transferencia y datos de hasta 384 [kbit/s], es decir, 48 [kbit/s] por intervalo temporal de la portadora de 200 [kHz] propia de *GSM*, lo que permite el acceso a Internet de alta velocidad e incluso para videoconferencia. La importancia de *EDGE* radica en el hecho de que en ella podrían converger tanto *GSM* como *TDMA IS-136*.

Con lo que respecta a la arquitectura, en los sistemas de tercera generación, el servicio en cualquier lugar y tiempo representa una parte fundamental de la funcionalidad. A ésta puede añadirse la necesidad de ampliar el área de servicio, unificar diversas tecnologías, integrar redes fijas y móviles, etc. Se puede esperar, entonces, que la diversidad de servicios que ofrecen los sistemas de tercera generación, represente un desafío para la arquitectura de los sistemas regulados por *ITU*. Ningún sistema en el pasado había pretendido proveer tan amplia gama de servicios de operación global y tan elevado nivel de capacidad y desempeño. En la figura 1.6 se simplifican los alcances de los sistemas de tercera generación, donde se pueden ver distintos sistemas operando simultáneamente, desde un ambiente de alta capacidad en picoceldas y servicios de voz, hasta un ambiente de cobertura satelital y servicios multimedia.

Con relación a las predicciones sobre la cantidad de usuarios que podría albergar la tercera generación, *The Yankee Group* anticipa que en este año (2004) habrá más de 1,150 millones en el mundo, comparados con los 700 millones que hubo en el 2000. Dichas cifras nos anticipan un gran número de capital involucrado en la telefonía inalámbrica, lo que con mayor razón las compañías fabricantes de tecnología, así como los proveedores de servicios de telecomunicaciones estarán dispuestos a invertir su capital en este nuevo desarrollo.¹⁷

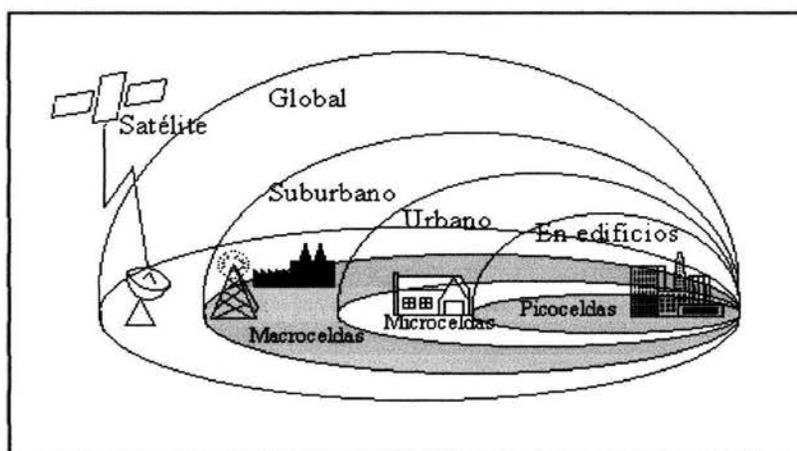


Figura 1.6 Sistemas operando simultáneamente en la tercera generación.

¹⁷ <http://www.yucatan.com.mx/especiales/celular/3g.asp>

| ASPECTOS | SEGUNDA GENERACIÓN | TERCERA GENERACIÓN |
|-----------------------------------|--|---|
| Tecnología digital. | Tecnología digital para modulación de voz, codificación de canal e implementación de canales de control y tráfico. | Incremento en el uso de tecnologías digitales incluyendo radios programables. |
| Operación en distintos ambientes. | Están optimizados para operar en algunos ambientes, como vehicular y pedestre en redes fijas y móviles | Se busca la operación multiambientes. |
| Bandas de frecuencia | 800 [MHz], 900 [MHz], 1.5 [GHz] y 1.8 [GHz]. | Uso de una banda de frecuencia global para sistemas terrestres y satelitales. |
| Servicios proporcionados | <ul style="list-style-type: none"> • Llamadas telefónicas/fax. • Correo de voz • Servicios de datos velocidades menores a 100[kbit/s]. • Roaming limitado a una región específica. | <ul style="list-style-type: none"> • Combinaciones de un teléfono móvil, laptop PC y televisión. • Altas velocidades web y servicio de datos en conmutación de circuitos o paquetes. • Videoconferencias • TV • Agenda electrónica de recordatorio de citas • Velocidades de 144 [kbit/s] a 2 [Mbit/s] • Roaming con Búsqueda Global con la disponibilidad de cobertura satelital. |

Tabla 1.3 Comparación de Sistemas de Segunda y tercera generaciones¹⁸

¹⁸ <http://www.revista.unam.mx/vol.2/num2/art3/>

Capítulo 2

Sistema GSM de telefonía celular

La demanda por parte de los usuarios de comunicaciones móviles para poderse mover a través de edificios, ciudades o países, ha llevado en buena parte al desarrollo de nuevas redes de comunicaciones móviles.

Un sistema de telefonía celular es responsable de proporcionar cobertura a través de un territorio particular, llamado región de cobertura o mercado. La interconexión de muchos de estos sistemas define una red inalámbrica capaz de proporcionar servicios a los usuarios a través de un país o continente.

Para proporcionar comunicaciones inalámbricas en una región en particular, se debe emplear una red integrada de estaciones base (*MS: Mobile Station*) para proporcionar una suficiente cobertura de radio a todos los usuarios. Las estaciones base, a su vez, deben estar conectadas a un eje central llamado Centro de Conmutación Móvil (*MSC: Mobile Switching Center*). El *MSC* proporciona conectividad entre la Red Telefónica de Conmutación Pública (*PSTN: Public Switching Telephone Network*) y las numerosas estaciones base, y por último, entre todos los abonados de un sistema. La *PSTN* forma la red de telecomunicaciones global que interconecta los centros de conmutación de telefonía convencional, llamados oficinas centrales, con los *MSCs* de todo el mundo.

Para conectar a los abonados con las estaciones base, se establecen enlaces de radio usando un protocolo de comunicaciones cuidadosamente definido, llamado la interfaz de radio. La interfaz de radio (*RI: Radio Interface*) debe asegurar una gran fiabilidad en el canal para asegurar que los datos se envían y se reciben correctamente entre el móvil y la estación base, y es por ello por lo que se realizan una codificación de la voz y una codificación del canal.

En la estación base, los datos de señalización y sincronización se descartan, y el resto de información de voz o datos se pasan a través del *MSC* hasta las redes fijas. Mientras que cada estación base puede realizar unas 50 llamadas simultáneas, una *MSC* típica es responsable de conectar hasta 100 estaciones base a la *PSTN* (hasta 5000 llamadas a la vez), y es por eso que la interfaz entre el *MSC* y la *PSTN* requiere una gran capacidad en cualquier instante de tiempo. A continuación se describirán individualmente los conceptos antes mencionados.¹

¹ http://ceres.ugr.es/~alumnos/c_avila/gsm0.htm

2.1 Arquitectura del sistema

El sistema tiene una arquitectura jerárquica y fue diseñada para volúmenes de tráfico razonablemente grandes. A diferencia de los sistemas analógicos la eficiencia del espectro es mucho mejor y una importante razón de esto es la utilización de poderosos compresores de voz.

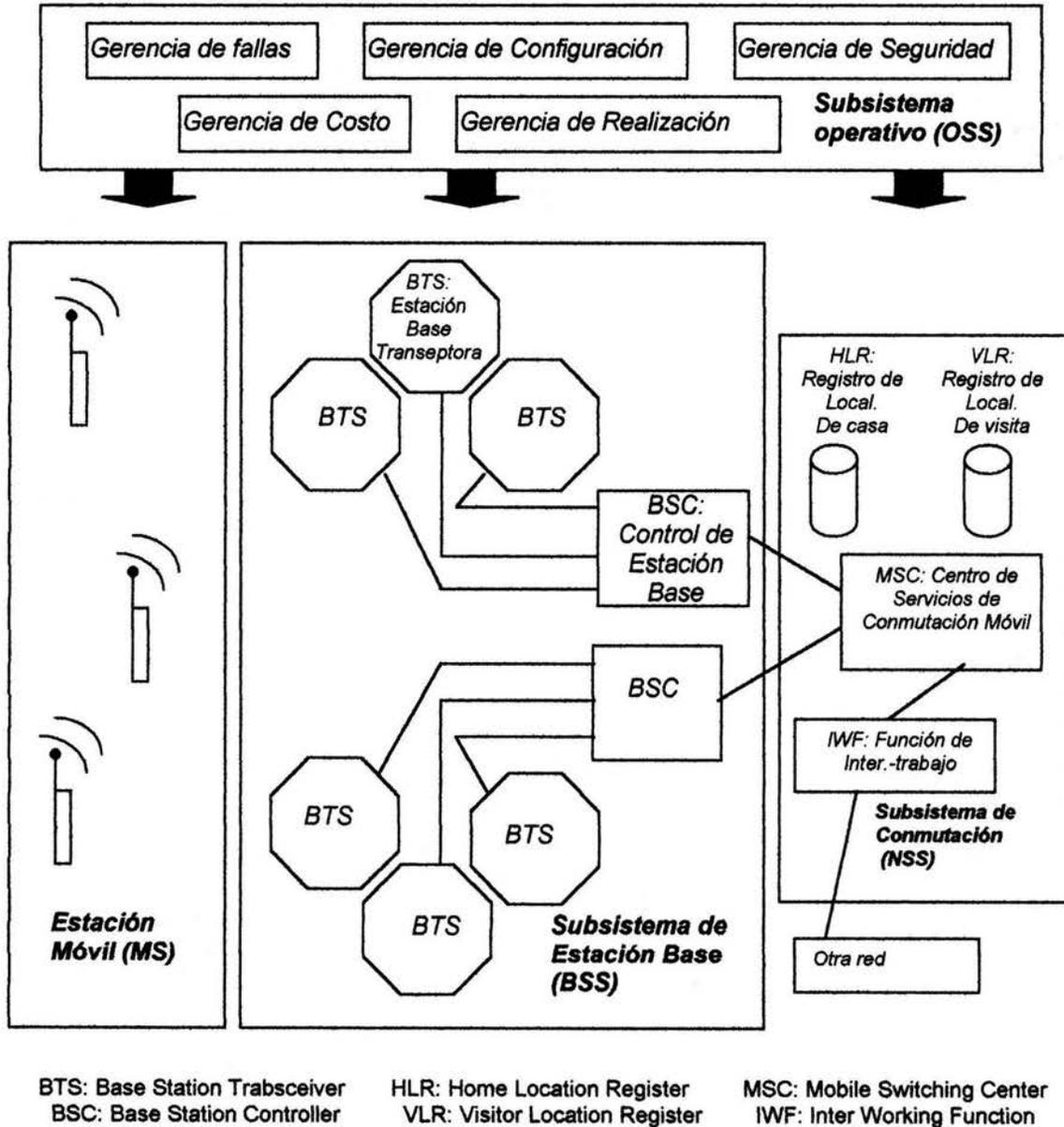


Figura 2.1 Arquitectura del sistema GSM.²

² BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.162

Las Estaciones Móviles (*MS*), el Subsistema de Estación Base (*BSS*), el Subsistema de Conmutación (*NSS: Network and Swirching Subsystem*) y el Subsistema de Operación (*OSS: Operation Subsystem*) y Mantenimiento son mostrados en la figura 2.1.

2.1.1 Subsistema de Terminal Móvil (MS)

La estación móvil representa normalmente la única parte del sistema completo que el usuario ve. Existen estaciones móviles de muchos tipos, pero quizás los más desarrollados sean los equipos de mano, pero cualquiera que sea la terminal tienen tres funciones básicas.

1. Funciones relacionadas con el servicio usado, por ejemplo las comunicaciones de voz (micrófono, altavoz y teclado) o para fax (*scanner* e impresora).
2. Funciones relacionadas con la transmisión de servicios en *GSM*.
3. Los adaptadores de terminales que pueden ser requeridos para realizar un seguimiento entre varias interfaces de las primeras dos funciones.

En *GSM*, el usuario es identificado por una pequeña "tarjeta inteligente" contenida en la terminal, esta tarjeta es llamada *SIM card (Subscriber-Identity Module)* y se usa para separar la identificación del usuario con la del equipo. Aparte de excepciones, como llamadas de emergencia, el acceso a la red es imposible sin la *SIM card*.

2.1.2 Subsistema de Estación Base (BSS: Base Station Subsystem)

Este subsistema esta constituido por dos partes:

- Transceptor de la Estación Base (*BTS: Base Transceiver Station*).
- Controlador de la Estación Base (*BSC: Base Station Controller*).

EL *BTS* esta compuesto por transmisores, receptores, equipo específico a la interfaz de radio para poder establecer contacto con las estaciones móviles y se pueden considerar como complejos *modems* de radio, con otras pequeñas funciones. Un *BTS* típico de la primera generación consistía en unos pequeños armarios (de 2 [m] de alto y 80 [cm] de ancho), que contenían todos los dispositivos electrónicos para las funciones de transmisión y recepción. Las antenas tenían generalmente unas pocas decenas de metros, y los armarios se conectaban a ellas por cables de conexión. Un *BTS* de este tipo era capaz de mantener simultáneamente 3 ó 5 portadoras de radio,

permitiendo entre 20 y 40 comunicaciones simultáneas. Actualmente el volumen de los *BTS* se ha reducido mucho, esperándose un gran avance en este campo dentro de *GSM*.

Una parte importante del *BTS* es la Unidad *Transcoder* y Adaptadora de Velocidad (*TRAU: Transcoder /Rate Adapter Unit*) que realiza la codificación y decodificación de la voz, y el control en las transmisiones de datos. El controlador de la estación base esta a cargo de las funciones de conmutación en el *BSS*. El *BSC* puede conectarse a varios *BTS*'s, estas comunicaciones se hacen empleando el protocolo *ISDN* por medio de la *Interface A-Bis*.³

El segundo componente del *BSS* es el *BSC*. Está encargado de toda la gestión de la interfaz de radio a través de comandos remotos sobre el *BTS* y principalmente de la gestión de localización de los canales de tráfico y de la gestión del "*handover*". El *BSC* está conectado por un lado a varios *BTS*'s y por otro al *NSS* (más específicamente a un *MSC*).

Un *BSC* es en definitiva un pequeño conmutador con una gran capacidad de cómputo. Sus funciones principales, como ya se ha mencionado son la gestión de los canales de radio y de los *handovers*. Un *BSC* típico consiste en uno o dos armarios, y puede gestionar hasta algunas decenas de *BTS*'s, dependiendo de su capacidad de tráfico.⁴

2.1.3 Subsistema de red y conmutación (*NSS: Network and Switching subsystem*)

El *NSS* incluye las principales funciones de conmutación en *GSM*, así como las bases de datos necesarias para los datos de los abonados y para la gestión de la movilidad. La función principal del *NSS* es gestionar las comunicaciones entre los usuarios *GSM* y los usuarios de otras redes de telecomunicaciones. Dentro del *NSS*, las funciones básicas de conmutación son realizadas por el *MSC*, cuya función principal es coordinar el establecimiento de llamadas hacia y desde los usuarios *GSM*. El *MSC* tiene interfaces con el *BSS* por un lado, y con redes externas por el otro. La interfaz con las redes externas requiere un "*gateway*" para la adaptación, cuya función es más o menos importante dependiendo del tipo de datos y de la red a la que se accede.

El *NSS* también necesita conectarse a las redes externas para hacer uso de su capacidad de transportar datos o señalización entre entidades *GSM*. En particular el *NSS* hace uso de un tipo de señalización parcialmente externo a *GSM*, que sigue el sistema de señalización del *CCITT N° 7* (que usualmente se conoce como la red *SS7*); esta red de señalización habilita el trabajo interactivo entre máquinas del *NSS* dentro de una o varias redes *GSM*.

³ http://agamenon.uniandes.edu.co/~revista/articulos/redes_moviles/rm.html

⁴ http://ceres.ugr.es/~alumnos/c_avila/gsm0.htm

Como parte del equipo, un *MSC* controla unos cuantos *BSCs* y es normalmente bastante grande. Un *MSC* incluye cerca de media docena de armarios de conmutación. Además de los *MSCs*, el *NSS* incluye las bases de datos. La información del abonado relativa al suministro de los servicios de telecomunicación está situada en el Registro de Posiciones Base (*HLR: Home Location Register*), independientemente de la posición actual del abonado. El *HLR* también incluye alguna información relacionada con la posición actual del abonado. Como una máquina física, un *HLR* es típicamente una computadora independiente, sin capacidades de conmutación, y capaz de manejar a cientos o miles de abonados. Una subdivisión funcional del *HLR* es el Centro de Autenticación (*AuC: Authentication Center*), cuya función se limita a la gestión de la seguridad de los datos de los abonados.

La segunda función de bases de datos identificada en *GSM* es el *VLR (Visitors Location Register)*, asociado a uno o más *MSCs*, y encargado del almacenamiento temporal de los datos para aquellos abonados situados en el área de servicio del correspondiente *MSC*, así como de mantener los datos de su posición de una forma más precisa que el *MSC*.⁵

En resumen, el Subsistema de red se ocupa de:

- Control de la llamada.
- Interfuncionamiento de redes.
- Datos del abonado y gestión de los servicios.
- Tarifación.
- Gestión de la movilidad.
- Gestión de la seguridad.
- Señalización de Interfaz A y *PSTN*.
- Control del *BSS*.

2.1.4 Subsistema de Operación (*OSS: Operation Subsystem*)

El modelo *OSI* describe las áreas en funcionamiento de la gestión de fallas, costos, configuración, realización y seguridad. Para realizar estas tareas el *OSS* esta directamente o indirectamente conectada a todos los otros sistemas en la red. Un *OSS* es también requerido para registrar nuevos suscriptores de una manera sencilla. Sin embargo, el control de suscriptores no forma parte de las especificaciones *GSM*.

⁵ http://ceres.ugr.es/~alumnos/c_avila/gsm0.htm

Estrictamente hablando, las especificaciones describen algunas interfaces entre componentes en la arquitectura. Estas son:

- La interfaz de radio, el cual describe los intercambios de datos entre el Subsistema de Terminal Móvil y el Subsistema de Estación Base.
- La interfaz *Abis*, la cual describe la comunicación entre la Estación Radio-Base y el Controlador de Estación Base.
- La interfaz *A* que describe el intercambio de datos entre el Subsistema de Estación Base y el Subsistema de Conmutación.

2.2 Realización Técnica

Los puntos centrales de la realización técnica se pueden definir primero en la interfaz de radio, que como ya se había mencionado, describe los intercambios de datos entre el Subsistema de Terminal Móvil y el Subsistema de Estación Base, y en donde se aborda el tema canal de radio como medio para establecer y mantener comunicaciones fiables, para posteriormente describir la transmisión de datos.

2.2.1 Aspectos técnicos de la comunicación de voz

La realización de una comunicación de voz empieza digitalizando la señal de voz usando un código *PCM* adaptivo. La codificación de voz comprende compresión de la señal digital por el código de predicción lineal (*LPC: Linear Predictive Coding*), que da como resultado un flujo de datos de 13 [Kbit/s], sin embargo con la introducción de códigos más eficientes *half-rate*, será reducido a más de 6.5 [kbit/s]. Por ahora a los 13 [Kbit/s] se le suma un flujo de datos de 9.8 [Kbit/s] de la codificación de canal debida a la detección de errores, corrección de errores y sincronización. Se utiliza un intercalado que asegura que los datos no ocurran como bits consecutivos en la interfaz de radio, pero son distribuidos sobretiempos. Esto es una ventaja desde que la probabilidad de que dos bits consecutivos serán transmitidos incorrectamente es mucho mejor que la probabilidad de que dos bits separados por tiempo contengan errores. Y si solo un bit es incorrecto todo el tiempo, esto puede ser resuelto con un algoritmo de corrección de errores. Después de la codificación de canal, la velocidad de los datos es $13+9.8=22.8$ [kbit/s], que es lo que conocemos como *bursts*.

Debido a que *GSM* es un sistema de radio truncado, lo que significa que existen menos canales que usuarios y que tal vez todos estos usuarios quieran usar el sistema al mismo tiempo, el espectro de radio se encuentra compartido por todos los usuarios, lo que hace necesario implementar un método para dividir el ancho de banda entre tantos usuarios como sea posible. En *GSM* una combinación de Acceso Múltiple por División de Frecuencia (*FDMA: Fecuecy Division Multiple Acces*) y Acceso Múltiple por División en el Tiempo (*TDMA: Time Division Multiple Acces*) es usada.

FDMA es la manera más común de acceso truncado. Con *FDMA*, se asigna a los usuarios un canal de un conjunto limitado de canales ordenados en el dominio de la frecuencia. La Unión de Telecomunicaciones Internacional (*ITU*), asignó las bandas de 890-915 [MHz] para el enlace ascendente y 935-960 [MHz] para el enlace descendente. Ya que este rango fue usado en la

década de los 80's por los sistemas analógicos del día, el CEPT tuvo la previsión de reservar los últimos 10 [MHz] de cada banda para las redes GSM que estaban siendo desarrolladas.

TDMA es común en los sistemas de telefonía fija. Las últimas tecnologías en los sistemas de radio son la codificación de la voz y la compresión de datos, que eliminan redundancia y periodos de silencio y decrecientan el tiempo necesario en representar un periodo de voz. Los usuarios acceden a un canal de acuerdo con un esquema temporal. Aunque no hay ningún requerimiento técnico para ello, los sistemas celulares, que emplean técnicas TDMA, siempre usan TDMA sobre una estructura FDMA. Un sistema puro TDMA tendría sólo una frecuencia de operación, y no sería un sistema útil.

TDMA implica el uso de técnicas de compresión de voz digitales, que permite a múltiples usuarios compartir un canal común utilizando un orden temporal. La codificación de voz, reduce mucho el tiempo que se lleva en transmitir mensajes de voz, eliminando la mayoría de la redundancia y periodos de silencio en las comunicaciones de voz. Otros usuarios pueden compartir el mismo canal durante los periodos en que éste no se utiliza.

Los usuarios comparten un canal físico en un sistema TDMA, donde están asignado unos espacios de tiempo (*slots*). A todos los usuarios que comparten la misma frecuencia se les asigna un *slot* de tiempo, que se repite dentro de un grupo de *slots* que se llama trama.

Así para GSM, la parte FDMA envuelve la división de frecuencia del ancho total de 25 [MHz] en 124 frecuencias portadoras de ancho de banda de 200 [KHz]. Una o más frecuencias portadoras están entonces asignadas a cada estación base. Cada una de estas frecuencias portadoras es entonces dividida en el tiempo (usando un esquema TDMA) en ocho canales de tiempo, un *slot* GSM es de 577 [μ s], y cada usuario tiene uso del canal (mediante su *slot*) cada 4.615 [ms]. Un canal de tiempo es usado para transmitir por el móvil y uno para recibir.⁶

La figura 2.2 ilustra el mecanismo de acceso. En la figura una terminal móvil esta comunicando en el *slot* de tiempo 2 y en un canal de los 890 – 915 [MHz]. La comunicación de la estación base a la terminal móvil siempre ocurre en un *slot* de tiempo diferente, pero en el mismo número de canal en la banda de 935 – 960 [MHz] y podemos ver un cambio en la frecuencia del canal 94 al canal 96, debido a que un canal adyacente nunca es usado en la misma célula.

⁶ <http://neutron.ing.ucv.ve/revista-e/No2/Gsmseuly.htm>

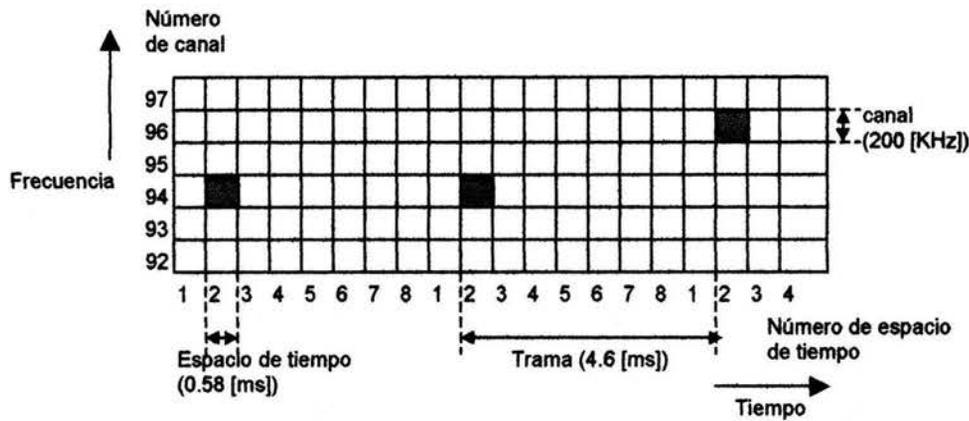


Figura 2.2 Mecanismo de acceso para GSM.⁷

2.2.2 Transmisión de datos: Modo transparente

Este es el primer modo que realmente usa el mismo protocolo de radio, esto significa que la velocidad de datos es constante, pero los datos libres de error no esta garantizado. La capacidad del canal de 22.8 [Kbit/s] debe ser distribuida sobre los datos de usuario y el error de corrección. La tabla 2.1 muestra varias velocidades de los servicios usado, así como su atraso asociado y la calidad de conexión o servicio.

| SERVICIO | TIPO DE CANAL | CALIDAD DE LA CONEXIÓN | RETRASO |
|--------------|---------------|------------------------|----------|
| 9.6 [Kbit/s] | Full rate | Baja | 330 [ms] |
| 4.8 [Kbit/s] | Full rate | Media | 330 [ms] |
| 4.8 [Kbit/s] | Half rate | Baja | 600 [ms] |
| 2.4 [Kbit/s] | Full rate | Media | 200 [ms] |
| 2.4 [Kbit/s] | Half rate | Media | 600 [ms] |

Tabla 2.1 Velocidad y retraso para el Modo transparente⁸

⁷ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.166

⁸ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.169

2.2.3 Transmisión de datos: Modo no transparente

Debido a la fluctuante calidad de las conexiones en un sistema móvil, es obvio que este tipo de redes debe tener una detección y corrección de errores. Esto se logra con este tipo de modo, el cual garantiza una transmisión libre de errores, pero con una velocidad de transmisión variable.

Un protocolo de radio diferente llamado (*RLP: Radio Link Protocol*) es usado en este modo y es una derivación del estándar *HDLC*. Los paquetes de datos con un tamaño de 60 [bits] son proporcionados con código de corrección de error audaz. En el caso de una recepción rectificable o libre de error, los paquetes continúan siguiendo uno a uno. Pero si los paquetes contienen errores no rectificables después de la recepción, una petición es transmitida para repetir la transmisión (*ARQ: Auto Repetition Request*). Esto hace posible una transmisión libre de errores, aunque la velocidad de transmisión no este garantizada.⁹

Las características de este tipo de modo pueden ser observadas en la tabla 2.2.

| SERVICIO | TIPO DE CANAL | CALIDAD DEL SERVICIO | RETARDO |
|--------------|---------------|----------------------|-----------|
| 9.4 [Kbit/s] | Full Rate | Alta | >330 [ms] |
| 4.8 [Kbit/s] | Full Rate | Alta | >600 [ms] |

Tabla 2.2 Velocidad, calidad y retardo para el Modo No transparente.¹⁰

2.2.4 Realización técnica del Servicio de Mensajes Cortos

GSM tiene un servicio de mensajes cortos (*SMS*), estos son mensajes individuales de hasta 160 caracteres y pueden ser recibidas o enviadas por una terminal móvil (terminal móvil que soporte este servicio). La velocidad de transmisión es aproximadamente 600 [bit/s]. Arriba de 93 caracteres pueden ser enviados a todos los usuarios en un área geográfica definida. Como estos grupos de mensajes no son transmitidos vía canal de señalización, estos mensajes sólo pueden ser recibidos sí la terminal móvil no esta en una sesión de comunicación (es decir llamada).

⁹ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.169

¹⁰ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.170

2.3 Aspectos Técnicos de Seguridad

Los aspectos de seguridad para GSM deben:

1. Proveer seguridad en la identificación de la terminal móvil.
2. Prevenir llamadas de escuchas escondidos a terceros.

En el caso GSM, las funciones que hacen esto posible son: autenticación, encriptación y reemplazamiento de la identidad de un móvil por un alias temporal para que sea imposible ver desde la radio – transmisor cual es el usuario que esta generalmente comunicando. Por otra parte, los operadores de la red pueden seleccionar una opción en la cual el Módulo de Identidad del Subscriber demanda un número de identificación personal (*PIN: Personal Identification Number*) del usuario.

Por lo tanto se debe disponer de la Autenticación o Cifrado de la información de señal y datos con una clave de cifrado diferente para cada llamada. Cada estación móvil tiene su propia clave de cifrado (*Ki*), el cual es un número secreto para cada usuario presente en el Modulo de Identidad del Subscriber de la terminal móvil y en el Centro de Autenticación de Estaciones (*AuC*) que tiene arriba de 128 [bits] de longitud. En el momento en que una estación móvil es captada por una *VLR* (*Visitor Location Register*), el *IMSI* (*International Mobile Subscriber Identification*) de la estación es transferido según el procedimiento especificado por el GSM.

Una vez que se conoce el *IMSI*, se sabe cual es el *HLR* (*Home Location Register*) del que depende la estación móvil a través de la señalización número 7, que sirve como canal de comunicación entre el *HLR* correspondiente y *VLR*. A partir de este momento, entra en funcionamiento el centro de autenticación que recupera el *IMSI* depositado ahora en el *HLR* su *Ki*, es decir la clave de cifrado.

El *AuC* selecciona un *RAND*, es decir un número aleatorio de 128 bits, el cual puede tener por lo tanto 2^{128} valores, con la clave *Ki* y un algoritmo (Algoritmo de autenticación), elabora una firma o respuesta firmada denominada *SRES* de 32 bits. En la figura 2.3 contiene una representación gráfica de la autenticación en GSM.

Con un algoritmo *A8* (Algoritmo para el cálculo del número *Kc* requerido para la encriptación) se genera una clave para cifrado que denominaremos *Kc*. Estos tres valores [*RAND*, *SRES* y *Kc*] forman un trío que es específico para cada estación móvil, y sólo es utilizado una vez para cada

comunicación, después es destruida. Normalmente cada *AuC* calcula para cada estación móvil una serie de tríos, con la intención de disponer de una cierta reserva y poder utilizarla en cualquier momento. Los algoritmos *A5* (Algoritmo de encriptación) y *A8* no permiten revelar nada acerca de la clave K_i , suponiendo que conocemos el trío de valores. El *VLR* autentifica la estación móvil mediante la comprobación de que el *SRES* calculado en el *VLR* es el mismo que el obtenido en la estación móvil. Para ello, se dispone en la estación móvil de los algoritmos *A3* y *A8*, y la propia clave de la estación móvil K_i , una vez que recibe la *RAND* del *VLR*. El cifrado de la información, tanto de voz y datos como de la señalización, se efectuará atendiendo al criterio de que la información generada en la estación móvil se pasa a través del algoritmo *A5* usando K_c como clave para el tráfico de datos.

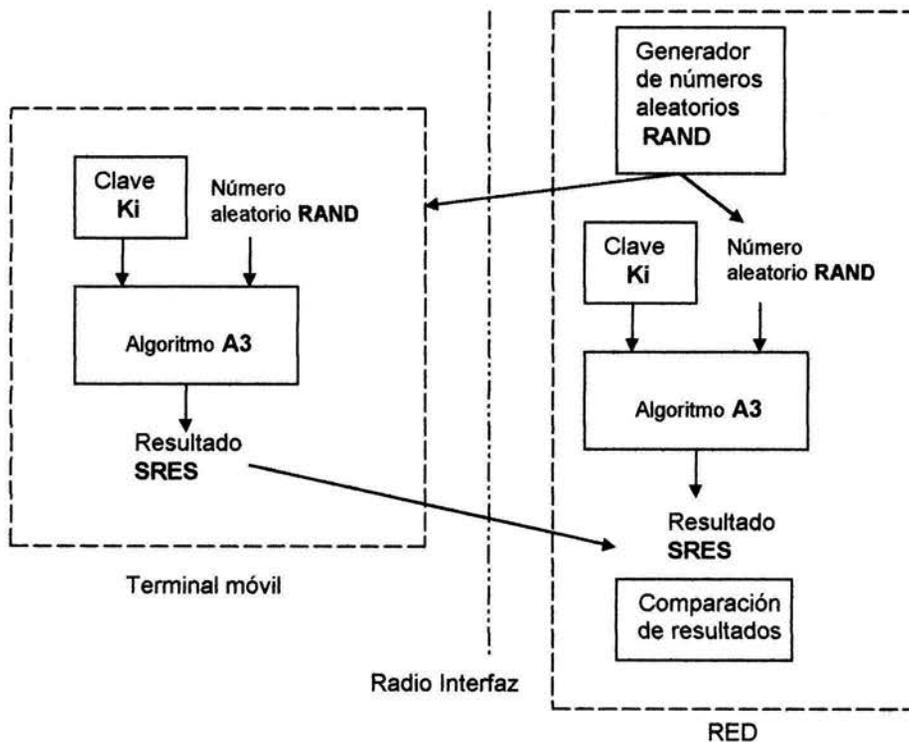


Figura 2.3 Autenticación para GSM.¹²

Normalmente cuando la estación móvil es autenticada, el *VLR* asigna un número de identificación temporal adecuado a la estación y se denomina *TMSI* (*Temporal Mobile Subscribe Identification*). La transmisión del *TMSI* también es cifrada, y es imposible relacionarla con el *IMS* correspondiente.

¹² BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.173

Es importante tener en cuenta que mientras la estación móvil depende del mismo VLR, el TMSI se utilizará para todas las transacciones que se vayan realizando. En la figura 2.4 se muestra una representación gráfica de la encriptación de datos en GSM.

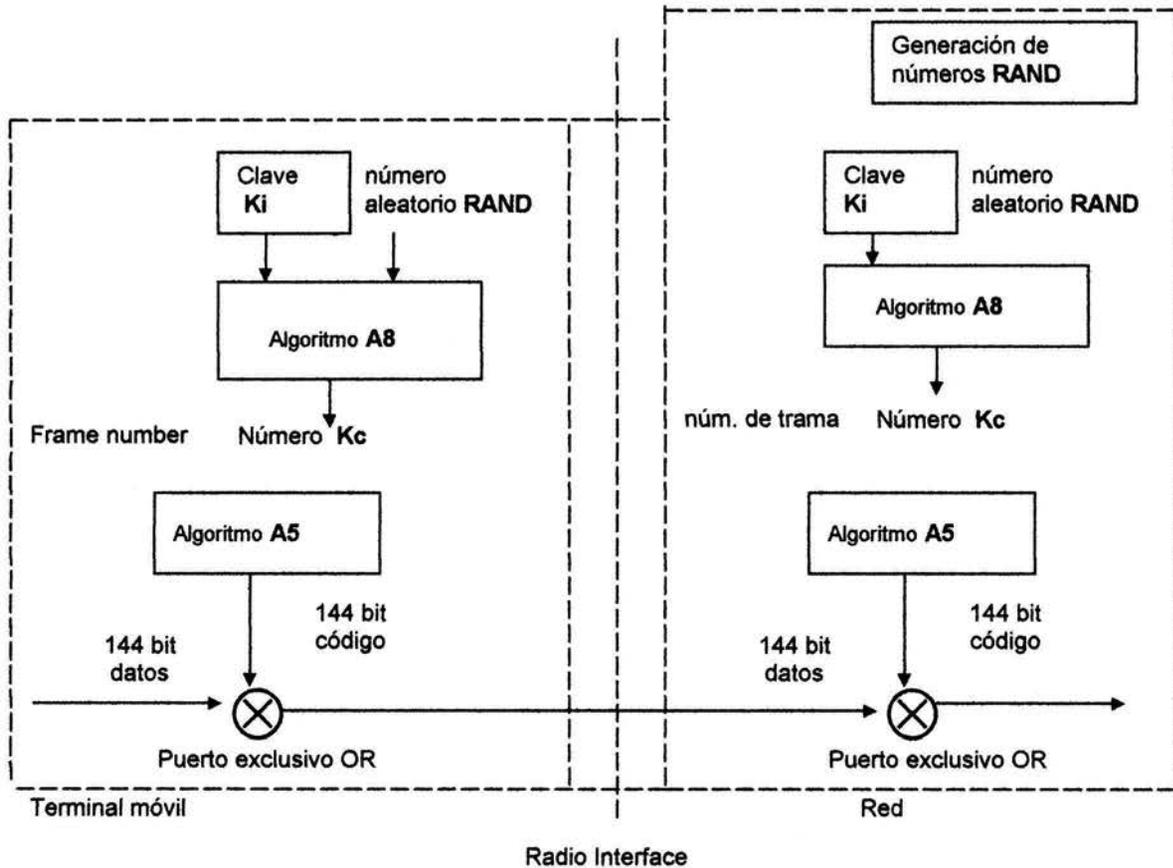


Figura 2.4 Encriptación de datos para GSM.¹³

¹³ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.174

2.4 Interconexión con otras redes

Las especificaciones GSM menciona algunas redes con las cuales la interconexión es posible. La voz en redes GSM esta en una forma digital. Para la interconexión con redes de telefonía, esta información es convertida en audio con un ancho de banda de 3.1 [KHz], la cual es utilizada en conexiones telefónicas. Este no significa, sin embargo, que los *modems* de teléfono puedan ser utilizados en ambos finales cuando un *subscriber GSM* se está comunicando con un usuario de la red de telefonía fija. Debido a la compresión, el audio en la conexión GSM es adecuado solo para voz. Equipos de *modems* son usados en la red GSM de datos para la interconexión con la red telefónica. Estos equipos son un conjunto de *modems* montados en la IWF. En la figura 2.5 se muestra una representación gráfica de este tipo de interconexión.

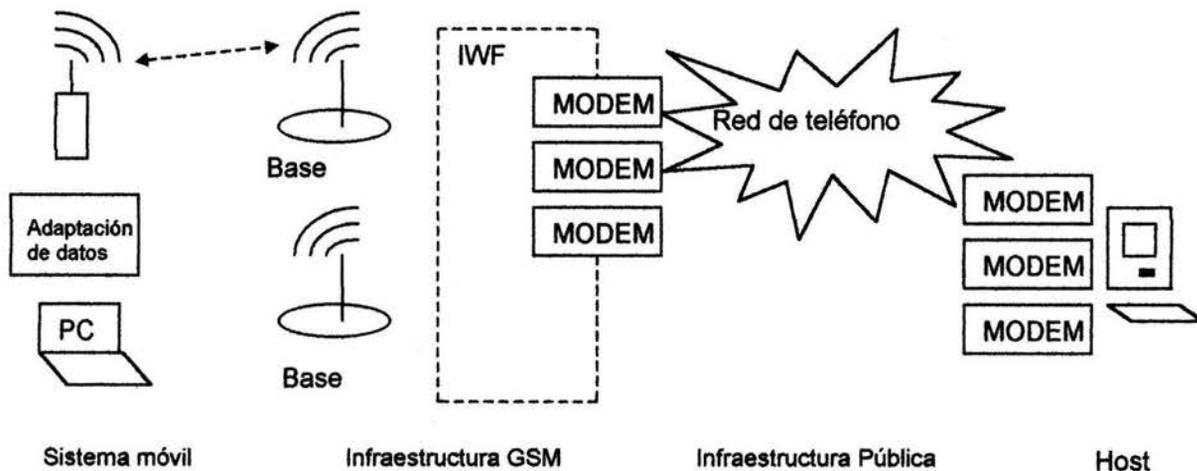


Figura 2.5 Interconexión de GSM con la red telefónica.

Cuatro diferentes servicios son posibles en el caso de la interconexión con una red de conmutación por paquetes con acceso X.25:

- Los servicios básicos *PAD* (*Packet Assembler/Disassembler*).
- Los servicios dedicados *PAD*.
- Servicios por paquete de datos básicos.
- Servicios por paquete de datos dedicados.

En el caso de los servicios *PAD* básicos, el usuario móvil llama a un *PAD*, teniendo acceso vía un equipo *modem*. El *PAD* establece una conexión con la conexión deseada de la red de conmutación por paquetes. Para poder hacer esto el usuario debe especificar dos números cuando inicializa la llamada, unos de los cuales indica la posición de la *PAD* en la red de telefonía y el otro indica el deseo del suscriptor en la red de datos de conmutación por paquetes. El usuario debe por supuesto tener suscripciones en ambas redes a menos que la conexión deseada a la red de datos admite no identificación, es decir números libres de pagos. La función del *PAD* es convertir los datos de conmutación de circuitos de los suscriptores *GSM* a la conmutación de paquetes para el suscriptor de dicha red y viceversa. Únicamente la terminal *GSM* puede iniciar una sesión básica del *PAD*. Tan pronto como se hace el contacto, las comunicaciones de las dos formas son posibles. Esta llamada es idéntica a la llamada del *PAD* de la red telefónica fija. Esto es porque la red del *GSM* no necesita facilidades especiales excepto para dicha interconexión de datos con la red de teléfono. En la figura 2.6 se muestra una representación de la línea de conexión.

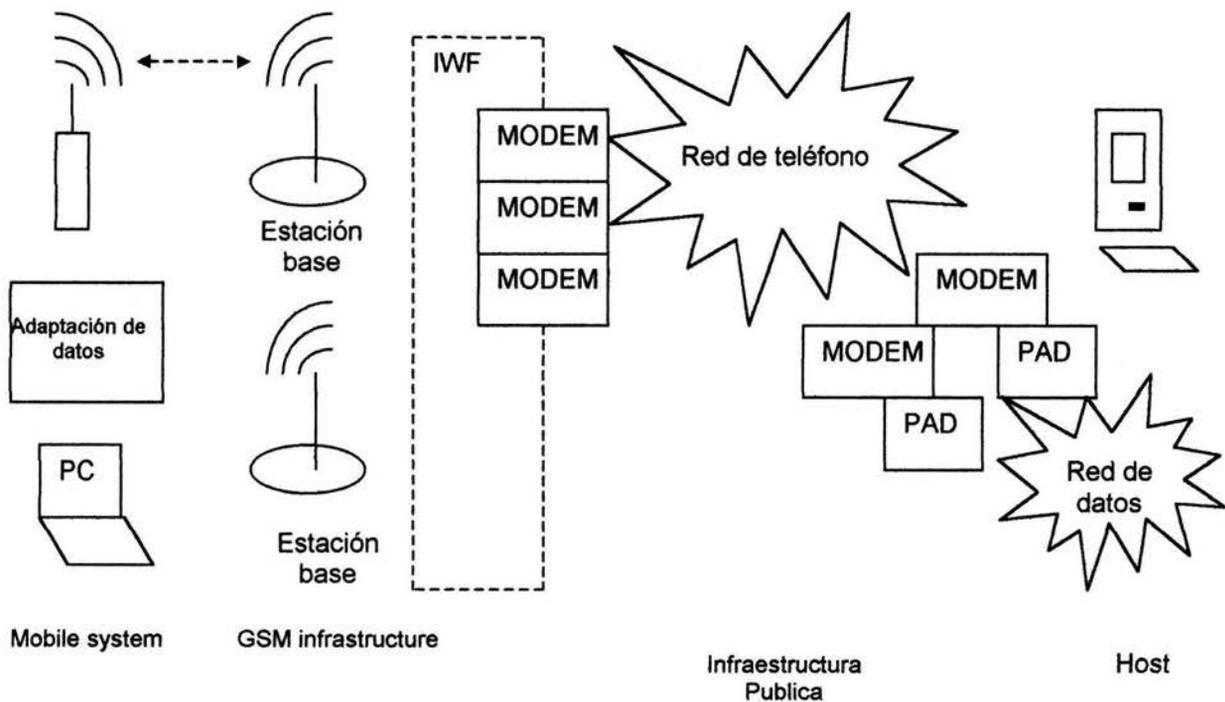


Figura 2.6 Interconexión de GSM con PSDN: PAD básico.

En el caso del servicio del *PAD* dedicado, un usuario puede comunicarse con un número limitado de conexiones que el operador de la red ha seleccionado en una red de datos de conmutación del paquete. El usuario únicamente tiene que especificar el tipo de servicio deseado y la red *GSM*

seleccionará un *PAD* adaptable que establece la conexión. El usuario puede, por ejemplo, pedir conexiones con un servicio de mensajes sin tener subscripción a la red de datos por conmutación de paquetes. El operador de la red GSM puede establecer los costos incurrido con el operador del servicio de datos cargando estos costos al subscritor. De nuevo, el teléfono GSM puede iniciar una sesión con este servicio. La figura 2.7 muestra una representación gráfica de la línea de conexión.

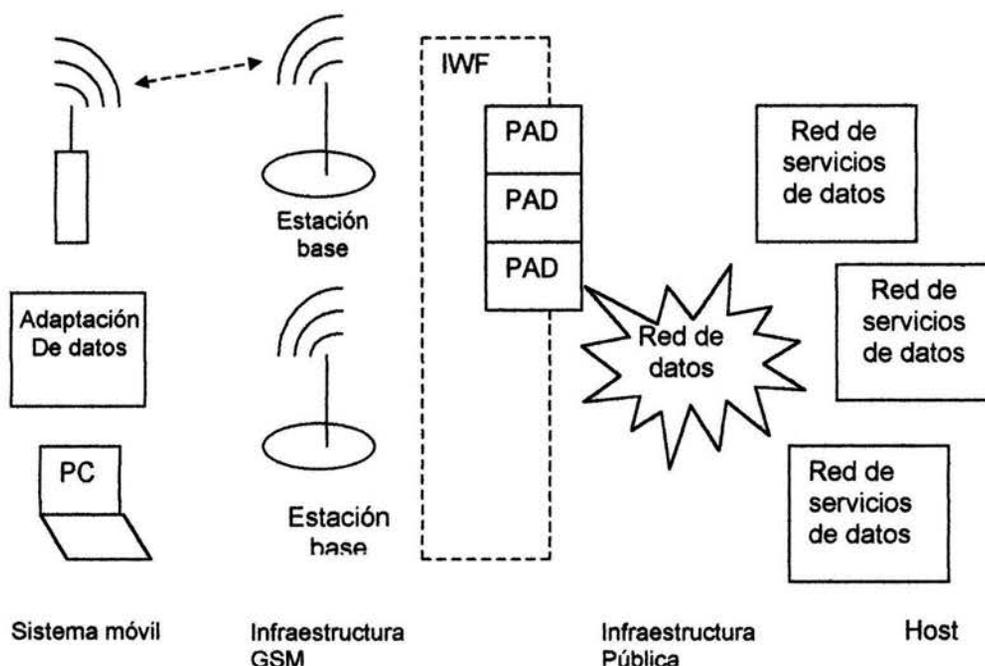


Figura 2.7 Interconexión de GSM con PSDN: PAD dedicado.

Los dos últimos modos, el servicio básico de paquetes de datos (*basic packet data service*) y el servicio dedicado de paquetes de datos (*dedicated packet data service*), son comparables a los servicios antes mencionados. La diferencia es que no se utiliza un *PAD* sino que las terminales móviles también operan mediante un "método sabio de paquete" (*packet-wise method*). Estos servicios generalmente no están llevados a cabo, además de que las actuales terminales móviles GSM son incompatibles con estos dos servicios. En estos servicios son conexiones de radio por conmutación de circuitos. La posición que solía estar ocupada por la *PAD* ahora está ocupada por un manipulador de paquete (*PH: Packet Handler*) que desbarata paquetes entre las dos redes.

Un caso especial es la interconexión con *ISDN*. Aunque una gran cantidad de GSM se basa en *ISDN*, las actuales funciones son prácticamente las mismas como la unión a una red telefónica normal. De nuevo existen cambios ligeros en el caso de las uniones a las redes de datos de

conmutación de paquetes vía una red intermediaria *ISDN*, esto porque *ISDN* también usa un tipo de solución *PAD*.

La interconexión para los servicios de mensajes cortos es un caso separado. El tipo *GSM* indica que estos mensajes pueden combinarse vía un centro de mensajes *SMS*. Un operador de red puede realizar varias uniones a otras redes fijas, por lo que se catalogan como teleservicios (concepto descrito en el próximo tema de este capítulo). La mayoría de los métodos para entregar mensajes corresponden en cierto modo a los métodos que se usan para compaginar las redes.¹⁴

¹⁴ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.175-178

2.5 Servicios

El estándar GSM ofrece:

1. Comunicaciones de voz y datos entre una terminal móvil y un usuario de una red fija.
2. Comunicaciones de voz y datos entre un usuario de una red fija y una terminal móvil.
3. Comunicaciones de voz y datos entre dos usuarios móviles.
4. Una conexión no orientada, servicio de mensajes cortos entre un usuario móvil y un SMS centro de mensajes y entre usuarios móviles.

Sin embargo las especificaciones definen muchos servicios. Estos son por lo regular combinados con una interconexión particular y existe una diferencia entre los servicios de transmisión, que son independientes del "contenido" de la información, es decir existe un intercambio de información entre usuarios, independiente de la forma del contenido por lo que se refiere solo a una capacidad de transportación y no al procesamiento de la información y los Teleservicios. Los servicios de transmisión se muestran en la tabla 2.3.

| TIPO DE SERVICIO | CARACTERÍSTICAS | INTERCONEXIÓN DE REDES | VELOCIDADES [BITS/S] | MODO DE TRANSMISIÓN |
|---------------------------|---|-------------------------|-------------------------------------|-----------------------------|
| Voz | 3.1 [KHz] en interconexión de red | <i>PSTN, ISDN</i> | No aplicable | No aplicable |
| Datos vía red de teléfono | Comunicación de datos asíncronos dúplex. | <i>PSTN, ISDN</i> | 300, 1200, 1200/75, 2400, 4800 9600 | <i>T, NT</i> |
| | Comunicación de datos síncronos dúplex | <i>PSTN, ISDN, CSDN</i> | 1200, 2400, 4800, 9600 | <i>T, NT (1200 solo NT)</i> |
| | Alternando voz y datos | <i>PSTN, ISDN</i> | Síncrono y Asíncrono servicio | |
| | Voz seguido por datos | <i>PSTN, ISDN</i> | Síncrono y Asíncrono servicio | |
| Datos vía red de datos | Acceso asíncrono <i>PAD</i> (básico y dedicado) | <i>PSDN</i> | 300, 1200, 1200/45 2400, 4800, 9600 | <i>T, NT</i> |
| | Acceso por paquetes síncrono | <i>PSDN</i> | 2400, 4800 y 9600 | <i>T, NT</i> |

Tabla 2.3 Servicios de transmisión para GSM.¹⁵

Los Teleservicios son aquellos servicios de Telecomunicación que proporcionan plena capacidad de comunicación entre usuarios o terminales, de acuerdo a protocolos preestablecidos. La tabla 2.4 muestra los más importantes teleservicios para GSM.

| TELESERVICIO | SERVICIO DE TRANSMISIÓN UTILIZADO. | COMENTARIO |
|-----------------------------|---|---|
| Telefonía | Voz 3.1 [KHz] sobre interconexión de red. | _____ |
| Llamada de emergencia | Voz 3.1 [KHz] sobre interconexión de red. | Trabaja sin la intervención de modulo de identificación del suscriptor. |
| Servicio de mensajes cortos | Transmisión vía señalización de canal. | Mensajes de y para una terminal móvil. |
| Grupo de fax 3 | Comunicación síncrona dúplex de datos. | Usa un fax – módem. |

Tabla 2.4 Teleservicios más importantes para GSM.¹⁶

Por último se encuentran los servicios suplementarios, los cuales no disponen de transmisión de datos, pero sí extra facilidades. Estas facilidades se pueden ver en la tabla 2.5.

| TIPO DE FUNCIÓN | SERVICIOS SUPLEMENTARIO |
|--------------------------|---|
| Salvo llamada | <ul style="list-style-type: none"> • Todas las llamadas salientes. • Llamadas de salida internacional. • Todas las llamadas entrantes. • Todas las llamadas entrantes cuando el suscriptor este en el extranjero. |
| Enrutación de llamada | <ul style="list-style-type: none"> • Siempre • Cuando el suscriptor este en el extranjero • Cuando el suscriptor este ocupado en una llamada • Cuando el suscriptor no responda |
| Llamada en espera | <ul style="list-style-type: none"> • Teniendo una llamada entrante esperando • Colocando el modo de segunda llamada en espera |
| Servicios de Información | <ul style="list-style-type: none"> • Mostrar el número del que llama antes de contestar (puede ser bloqueado). • Mostrar los cargos de nuestra cuenta |
| Otros | <ul style="list-style-type: none"> • Conferencia telefónica con un máximo de 5 suscriptores. • Grupo de usuarios cerrado |

Tabla 2.5 Servicios Suplementarios para GSM.¹⁷

¹⁵ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.181

¹⁶ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.182

¹⁷ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.183

2.6 Codificación de voz

Como se menciono anteriormente, el servicio más importante que ofrece GSM al usuario es la transmisión de voz y el requerimiento técnico general es simple: transmitir señales de voz con un nivel aceptable de calidad minimizando el número de bits que se necesita transmitir.

El codificador utilizado en GSM es RPE-LTP (*Regular Pulse Excited Long Term Prediction*). El RPE-LTP combina ventajas de un codificador RELP (propuesto por Francia), que tiene una buena calidad con baja complejidad, pero en donde la calidad de voz es limitada debido al ruido tonal introducido por la regeneración de altas frecuencias y por errores de bit introducidos en la transmisión, con las ventajas de un codificador MPE-LTP (*Multi-Pulse Excited Long Term Prediction*, propuesto por Alemania). El MPE-LTP por otra parte produce una alta calidad de sonido, y no le afectan mayormente los errores de bits en la transmisión, pero su complejidad es mayor.

Modificando el codificador RELP para incorporarle ciertos rasgos de MPE-LTP, la tasa de bits neta pudo ser reducida de 14.77 [Kbps] a solo 13 [Kbps] sin pérdida en la calidad de voz. La modificación más importante, fue la adición de un ciclo LTP (*Long Term Prediction*).

En la figura 2.8 se puede observar un diagrama general de un codificador RPE-LTP, que se divide en cuatro bloques fundamentales de procesamiento. La secuencia de voz es preenfatzada y ordenada en segmentos de 20 [ms] de duración cada una. Posteriormente a las tramas se les aplica una ventana *Hamming*, posteriormente entran a un filtro de predicción de tiempo corto (*STP: Short Term Prediction*) donde se extrae 8 parámetros LAR. Estos son codificados con distinta resolución de bits cada uno, ya que presentan diferentes rangos dinámicos y distribuciones probabilísticas.

Los parámetros LAR obtenidos del STP, son transmitidos en el canal, y a la vez alimentados hacia un filtro LPC inverso, que reconstruye la señal con el error residual (igual que en RELP), que luego es minimizado por el ciclo RPE-LTP. Durante el análisis LTP además son extraídos los datos sobre la tonalidad (*Pitch*) y el factor de ganancia. Estos factores son transmitidos a una tasa de 3.6 [kbps].

La señal residual del LTP es ponderada y descompuesta en tres secuencias de excitación que son candidatos a ser transmitidas. La secuencia con mayor energía es seleccionada para representar

el residuo del *LTP*. Los pulsos en la secuencia de excitación son normalizados a amplitud máxima, cuantizados y luego transmitidos a una tasa de 9.6 [kbps].

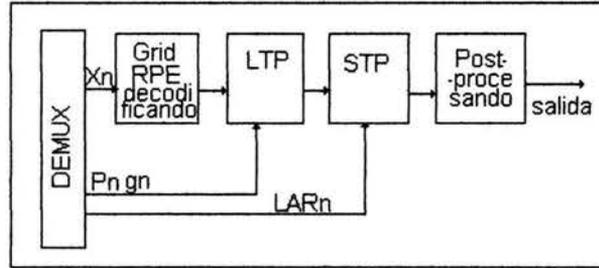


Figura 2.8 Transmisor RPE-LTP

La figura 2.9 se muestra el diagrama de bloques de un receptor *RPE-LTP*. Los parámetros de *RPE* son decodificados por el primer bloque, y transferidos al sintetizador *LTP* que los combina con el *pitch* y la ganancia para generar la señal de largo tiempo. Esta señal es procesada por el *STP* que incluye la información del vector *LAR* para reconstruir una señal de corto tiempo, que es luego acondicionada para la salida de voz.

Este codificador se verá a detalle en el siguiente capítulo.

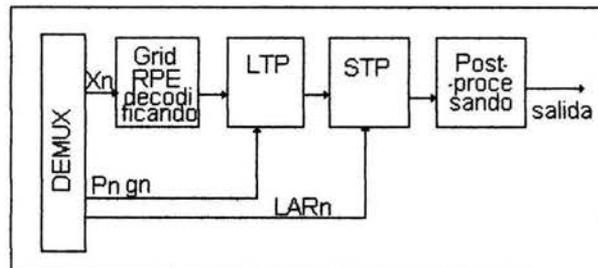


Figura 2.9. Receptor RPE-LTP

2.7 Nuevos desarrollos en GSM.

Una de las graves confrontaciones que están teniendo las redes GSM, es la escasez de frecuencias. En muchos países europeos, más de la mitad de la banda GSM esta ocupada por una red interina (como NMT-900 o TACS). Mientras que el espectro restante debe ser usualmente compartido por dos operadores de red GSM. Sin embargo la Oficina de Radiocomunicaciones Europea publicó una recomendación en 1996 para reguardar la disponibilidad de esta banda extra en diferentes países. Esto está anticipando que la infraestructura será considerablemente fácil adaptarla a esta banda, sin embargo antes la mayoría de las terminales están siendo hechas para soportar estos cambios.

Otro importante desarrollo consiste en la implementación de una nueva codificación de voz, sobre todo considerando un crecimiento anticipado de la capacidad computacional en el Procesamiento Digital de Señales y por consecuencia otro modo fue provisto en el cual se obtenía la misma calidad de voz con *slots* de tiempo a la mitad, lo que implica que la capacidad de una red puede incrementarse considerablemente sin un alto costo, sin embargo solo unos cuantos fabricantes han implementado esto en su infraestructura.

Sin embargo parecía que los proveedores estaban interesados en otra codificación de voz: EFR: *Enhanced Full Rate*. Este protocolo fue desarrollado inicialmente para PCS-1900 para competir con el estándar americano. Como el convencional protocolo *full-rate*, EFR utiliza 13 [Kbit/s] pero se obtiene una considerable calidad de voz y por otra parte es menos sensible a los errores de transmisión, así se necesitaran *teléfonos full-rate/half-rate/ERF* que deberán tener la capacidad de conmutar a *half-rate*. En la tabla 2.6 varios codificadores de voz son listados.

En lo que respecta a la comunicación de datos por GSM, dos nuevos desarrollos son relevantes: un programa propuesto para modificar el modo de conmutación por circuitos por un modo de conmutación por paquetes, GPRS: *General packet radio services*. La conmutación por paquetes tiene algunas ventajas sobre la conmutación por circuitos, haciendo este más adecuado para el entorno del tráfico tanto en redes LAN e Internet. Sin embargo las modificaciones a larga escala tanto en la infraestructura y la terminal son necesarias para introducir GPRS.

El desarrollo de GPRS ha sido retrasado y solo parte de las especificaciones fueron estandarizadas en 1997, así las investigaciones para las modificaciones involucran grandes costos y sin una visión clara del tipo y tamaño del mercado y del costo, los operadores no se vieron interesados.

En cambio de *GPRS*, algunos proveedores (*Nokia* en particular) han puesto empeño en introducir un modo de datos más simple, que no requiere grandes modificaciones. Esta propuesta, también llamada *HSCCD: High Speed Circuit- Switched Data*, utiliza algunos *slots* de tiempo para una sesión de comunicación simple. Esto significa que los usuarios pueden tener un múltiplo de 9.6 [Kbit/s] a su disposición. Se han mostrado acceso a Internet o transmisiones de vídeo muy cercanas a transmisiones de tiempo real en cuatro *slots* de tiempo (38.4 [Kbit/s]). Sin embargo esta asociado a algunos problemas para los operadores de red: un usuario utilizando esos *slots* de tiempo constituye una gran carga dentro de una célula y si muchos subscriptores utilizan *HSCCD*, la distribución de tráfico actual cambiará considerablemente.

| CODIFICACIÓN DE VOZ | VELOCIDAD DE DATOS | DISPONIBLE | COMENTARIOS |
|---------------------------------|--------------------|---|--|
| <i>Full rate</i> | 13 [Kbit/s] | Desde de los inicios de <i>GSM</i> . | Protocolo en un uso para todas las terminales producidas para datos. |
| <i>Half rate</i> | 6.5 [Kbit/s] | Estándar desde la mitad de los 90's, implementación poco clara. | Permite mayor capacidad de la red. Calidad de voz idéntica a <i>full-rate</i> , pero ligeramente más sensible a problemas con conexiones de pobre calidad. Las terminales deben soportar <i>half-rate</i> y <i>full-rate</i> . |
| <i>Enhanced Full Rate (EFR)</i> | 13 [Kbit/s] | Desde los finales de 1997 | Mejor calidad de voz que <i>full-rate</i> y hasta menos sensibilidad a la transmisión de errores. Desarrolla para <i>PCS-1900</i> , ahora para <i>GSM</i> y <i>DCS-1800</i> . |

Tabla 2.6 Protocolos de la codificación de voz para la familia *GSM*.¹⁸

Un último desarrollo que se citará es *GSM-R*, una variante del estándar europeo que las compañías de trenes quieren implementar. En el presente todo tipo de diferentes sistemas son utilizados, por ejemplo, por conductores de trenes y personal de mantenimiento. En el futuro todo ellos deben de ser capaces de utilizar *GSMR*. *CEPS* ha reservado 2 x 4 [MHz] para *GSMR* en la banda extendida de *GSM*. Esta banda esta localizada inmediatamente debajo de la actual banda comercial *GSM*. El estándar ha sido modificado para operar en trenes de alta velocidad.

La idea de *GSMR* se origino en 1989 y recibió soporte financiero de la Unión Europea en la forma del proyecto llamado *MORANE*. Cada compañía de trenes europea a contribuido a la

¹⁸ BEKKERS, Rudi. *Mobile telecommunications: Standards, Regulations and Applications*. Boston: Artech House Publishers, 1999. P.188

implementación de GSM-R. Por lo tanto la compañía de trenes alemana *Deutsche Bahn* y la Francesa *SNCF* estas preparadas para este desarrollo.¹⁹

¹⁹ BEKKERS, Rudi. Mobile telecommunications: Standards, Regulations and Applications. Boston: Artech House Publishers, 1999. P.186-190

Capítulo 3. Codificación de Voz

3.1 Historia

La codificación de voz, que en un sistema de telefonía se puede traducir como la codificación de fuente, consiste en: convertir, procesar y transmitir la señal de voz analógica, en una señal digital con las menores pérdidas en la calidad, con el fin de optimizar el canal de voz reduciendo el número de bits a transmitir, es decir, la velocidad en bps (*bit rate*) y permitiendo en algunos casos transmitir varias comunicaciones por un solo canal e incorporar algoritmos de cifrado para establecer comunicaciones seguras.

Con lo que respecta al desarrollo de estos sistemas podríamos empezar en el año de 1605 cuando Bacon desarrolla un alfabeto de dos palabras para representar 24 letras usando 5 dígitos, a estas letras, a y b, se les llama palabras codificadas y al conjunto se le llama código (*code*) y en 1641 estas ideas se extendieron a los sistemas M-arios.

En 1703 Leibniz describió el código binario usando solo 0 y 1 para representar enteros en longitud variable.

En 1791 Wolfgang Von Kempelen construyó un rudimentario sintetizador de voz capaz de pronunciar frases cortas mediante un ingenio mecánico.

El primer sistema de comunicaciones digitales es el telégrafo, inventado por Morse en 1837, el cual usaba pulsos cortos y largos. En 1875 Banfot desarrolló el sistema de telégrafo actual con 5 dígitos por palabra. En 1879 Marconi desarrolló el radioteléfono.

En 1924 Nyquist propuso el teorema del muestreo, desarrollando ya las bases matemáticas para el análisis de la señal de voz. En 1928 estableció la máxima tasa de bits (*bps*) en un canal de cierto ancho de banda.

En 1937 Reeves desarrolló uno de los sistemas más importantes, el *PCM* (Modulación por Pulsos Codificados), planteando el problema de la codificación eficiente de la señal de voz para reducir al máximo la velocidad de transmisión y el número de bits necesarios para la codificación binaria de la señal de voz. A partir de *PCM* se desarrollaron *DPCM* y *ADPCM*, algunos propuestos como estándares por la *CCIFF* (*International Consultative Committee for Telephone and Telegraph*).

Ya en el año 1939 se presentó una versión electrónica del sintetizador de Kempelen. El sistema denominado *VODER*, el cual consistía en un teclado muy elaborado que permitía controlar la articulación y generación de sonidos vocálicos y consonánticos. Este sistema pionero sembró las bases de los modernos sintetizadores de voz, basados actualmente en tecnología digital, pero compartiendo la misma teoría sobre el modelo de producción del habla.

En el mismo año Homer Dudley de los laboratorios *Bell* demostró en la *New York World's Fair* el primer vocoder, en el cual la idea central era analizar la voz para extraer una serie de características que serían enviadas por el emisor y una vez que el receptor tuviera estas características reconstruirían la señal de voz original.

En 1959 se desarrolló la teoría de tasa de distorsión que establece límites en capacidad de un canal. Por estas fechas Fant ya trabaja en los modelos de producción de voz.

Los avances en la tecnología de integración de circuitos integrados (*VLSI*), durante los 60 y 70 permitió nuevas soluciones al problema de codificación de voz, proponiendo soluciones basadas en la transformada de Fourier.

Desde los 80's la investigación ha sido encaminada hacia conseguir codificadores que utilicen un ancho de banda cada vez menor surgiendo por ejemplo: vocoders por predicción lineal, *LPC-10* (*FS-1015*), codificación *RELTP*, codificación multiplulso (*MPC*), codificación *CELP*, codificación *VSELP*, *HFR GSM* y codificación *RPE-LTP*, los cuales serán descritos en este capítulo.

3.2 Tipos de codificación

Los codificadores de voz se clasifican generalmente en tres grupos:

1. Codificadores de forma de onda (*Waveform coders*).
2. VOCODERS.
3. Codificadores híbridos.

El objetivo de los codificadores de onda es reproducir en el decodificador la señal de voz original muestra a muestra, partiendo de una muestra inicial, es decir sin tomar en cuenta la naturaleza de la misma y pueden operar tanto en el dominio del tiempo, como en el dominio de la frecuencia. Este tipo de codificadores se caracteriza por proporcionar una alta calidad de voz pero con velocidades promedias de 32 [Kbit/s], lo cual los hace poco utilizados cuando se necesitan *bit rates* bajos.

La diferencia entre un vocoder y un codificador de forma de onda, es que los vocoder dependen del tipo de señal que se desea codificar. El vocoder lo que hace es extraer información acerca de la modulación de la voz y discriminar entre diferentes tipos de sonidos (sordo o sonoro) y su correspondiente frecuencia. En este tipo de codificación la información a transmitir puede ser reducida considerablemente en comparación con los codificadores de forma de onda teniendo una calidad aceptable, pero sonando demasiada sintética.

Por último los codificadores llamados híbridos combinan las técnicas de los dos tipos de codificaciones mencionadas, tomando las ventajas de cada una de ellas y consiguiendo una alta calidad de voz a bajos *bit rates*.

3.2.1 Codificadores de forma de onda

El funcionamiento de este tipo de codificadores es intentar reproducir la forma de onda de la señal de entrada sin importar que señal de entrada sea, por lo que se utilizan para codificar una amplia variedad de señales. Aunque presentan una degradación aceptable en presencia del ruido y errores de transmisión sólo se consiguen *bit rates* medios y como se había mencionado se pueden desarrollar tanto en el dominio del tiempo como en el dominio de la frecuencia siendo los más importantes:

En el dominio del tiempo:

- PCM.
- DPCM.
- ADPCM.

En el dominio de la frecuencia:

- Codificación por subbandas.
- Codificación por transformada.

PCM

La conversión de una señal a PCM envuelve un muestreo, cuantización y la codificación de los niveles de cuantización, que son los datos que se envían al receptor. Un convertidor analógico – digital es el que hace los tres procesos.

El muestreo de una señal de banda base y frecuencia máxima $f_{\text{máx}}$ está determinada en forma única por el conjunto de sus valores equiespaciados y separados en el tiempo por $T_s \leq \frac{1}{2f_{\text{máx}}}$, es decir con una frecuencia de muestreo $f_s \geq 2f_{\text{máx}}$, lo que constituye el teorema de muestreo de Nyquist.

Con lo que respecta a la cuantización, consiste en asignar un código binario a cada nivel de cuantización, dado que su compromiso es contener tan pocos niveles como sea posible para mantener un nivel aceptable de ruido, en telefonía se utiliza un cuantizador escalar de niveles fijos no uniforme con 128 niveles para asegurar una calidad aceptable o 256 niveles para telefonía de calidad.

Con base en lo anterior, para telefonía si calculamos la velocidad en bps con la ecuación 3.1

$$\text{velocidad en bps (bit rate)} = \left(\frac{n_o \text{ de bits}}{\text{muestra}} \right) \times f_s \quad (3.1)$$

Con 256 niveles, que se traduce en un $n_o \text{ de bits} / \text{muestra}$ de 8 [bits] y una frecuencia de muestreo de 8000 [Hz] resulta un *bit rate* de 64 [Kbps], que por sí solo, este *bit rate* no es

problemático, pero tengamos en cuenta que si se trata de transmitir miles de conversaciones simultáneas por un mismo canal, este tipo de codificación no es tan conveniente.

DPCM

Una ventaja que existe en las señales de voz es su alta correlación entre las muestras adyacentes, particularmente en regiones de voz sonora, en consecuencia la señal de diferencia formada por la substracción adyacente de muestras tiene un rango dinámico mucho menor que la señal original, por lo tanto esta señal puede ser codificada con unos pocos bits. Esta característica es explotada en DPCM, donde la muestra actual $x[n]$ es estimada utilizando la muestra anterior, substrayendo esta de la actual muestra $x[n]$ y cuantizando la diferencia de la señal o residuo de predicción, $d[n] = x[n] - \hat{x}[n]$, como se muestra en la figura 3.1.

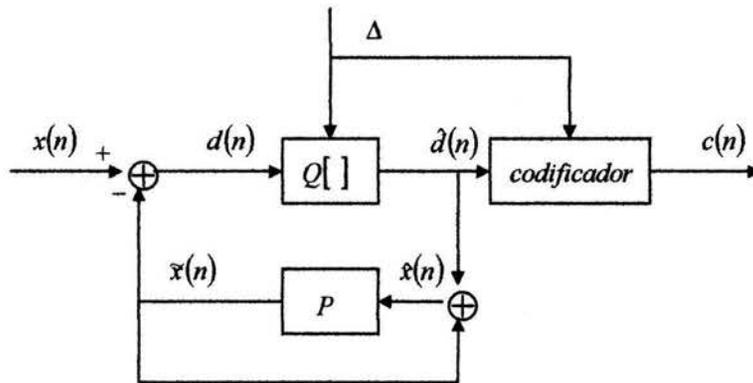


Figura 3.1 Codificador DPCM.

La obtención de la señal reconstruida de la señal $\hat{x}[n]$, se obtiene sumando la señal cuantizada $d[n]$ a la señal estimada $\hat{x}[n]$. \hat{x} es la salida del predictor que estima la siguiente muestra $\hat{x}[n]$, la cual es normalmente estimada como una combinación lineal de pocas muestras pasadas \hat{x} , como se muestra en la siguiente expresión.

$$\hat{x}[n] = \sum a_k \hat{x}[n - k] \tag{3.2}$$

Para el caso más simple (N=1), la señal estimada es simplemente $\hat{x}[n] = a_1 \hat{x}[n - 1]$ y los coeficientes de predicción a_k se elige adecuadamente para minimizar el error cuadrático medio. El diagrama del decodificador se muestra en la figura 3.2.

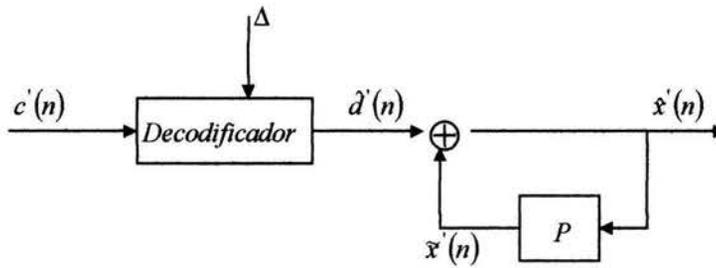


Figura 3.2 Decodificador DPCM

ADPCM

En la técnica *DPCM* descrita anteriormente tanto como el cuantificador como el predictor son fijos. Una substancial eficiencia se podría conseguir si el cuantificador o un predictor fuera adaptable a los cambios del residuo de prediccion. Con esto se aseguraría que la raíz cuadrada del error de prediccion se minimice continuamente.

Existen dos métodos para adaptar los cuantificadores y los predictores, uno de ellos es llamado *feedforward*, donde los niveles de reconstrucción y los coeficientes de prediccion se calculan en el emisor, usando un bloque de voz. Después son cuantificados y transmitidos al receptor como información lateral. Tanto el emisor como el receptor usan estos valores para hacer las predicciones y cuantificar el residuo. En la figura 3.3 y figura 3.4 se muestran un codificador y decodificador *feedforward* respectivamente.

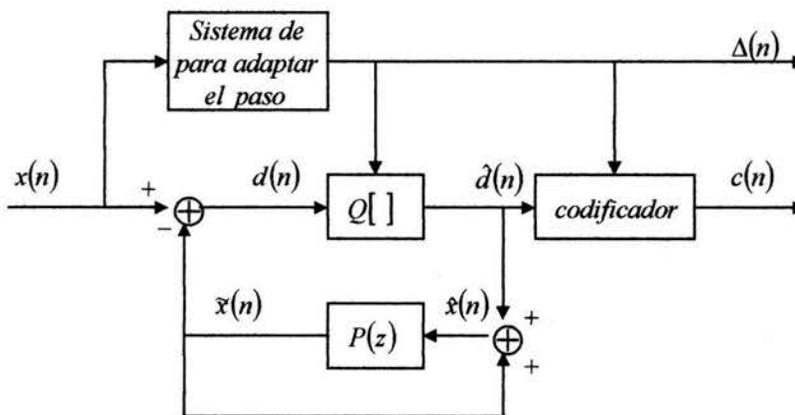


Figura 3.3 Codificador ADPCM feedforward

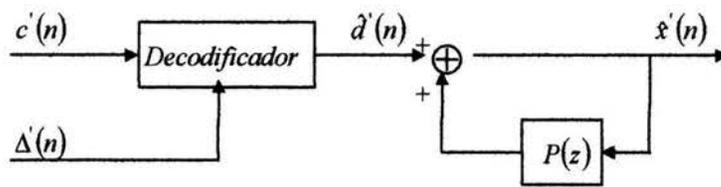


Figura 3.4 Decodificador ADPCM feedforward.

En el segundo, llamado *backward* los niveles de reconstrucción y los coeficientes de predicción se calculan a partir de la señal codificada. Dado que la señal es conocida tanto para emisor como para el receptor, no hay necesidad de transmitir información lateral, así el predictor y el cuantizador pueden actualizarse para cada muestra. La adaptación *backward* puede dar menores *bit rates*, pero es más sensible a los errores de transmisión que la adaptación *feedforward*. En la figura 3.5 se muestra un esquema de codificador y decodificador ADPCM *backward*.

La CCITT propone un estándar de codificación de voz telefónica (estándar G.721) a una velocidad de 32 [Kbit/s], utilizando un esquema de adaptación *backward* tanto para el cuantizador como para el predictor. El predictor en este caso tiene dos polos y seis ceros.¹

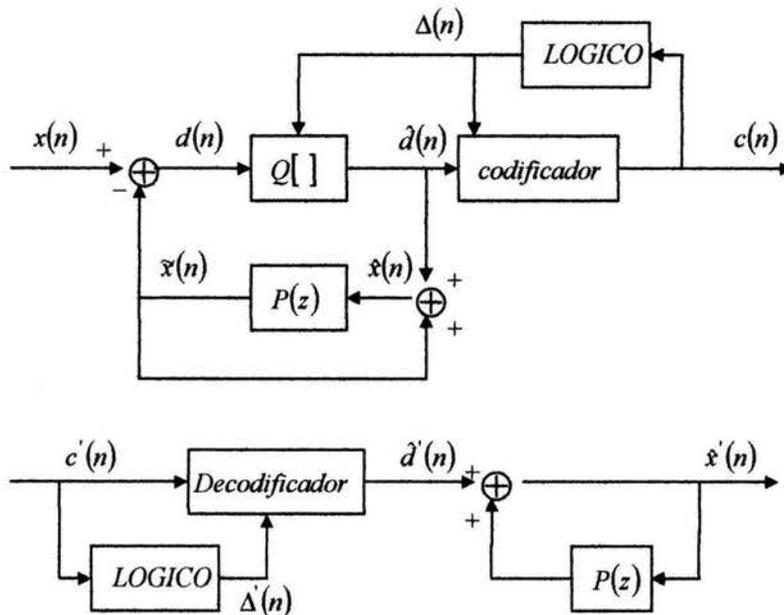


Figura 3.5 Codificador y decodificador ADPCM backward.

¹ OWENS, Frank , Signal Processing of Speech. McGraw Hill, 1987. P.30-31

Codificación por sub – bandas

La codificación por sub – bandas es un tipo de cuantización que utiliza previamente una codificación por transformación. Esta última efectúa una división en frecuencias o división espectral de la señal, como se muestra en la siguiente figura.

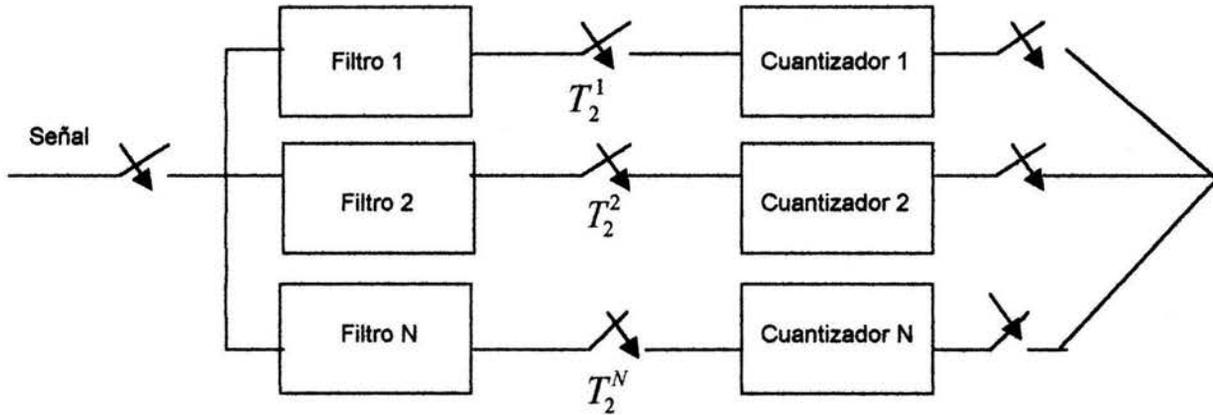


Figura 3.6 Representación de una codificación por sub – bandas por una descomposición espectral del canal.

En algunas aplicaciones los filtros tienen la siguiente respuesta en frecuencia.

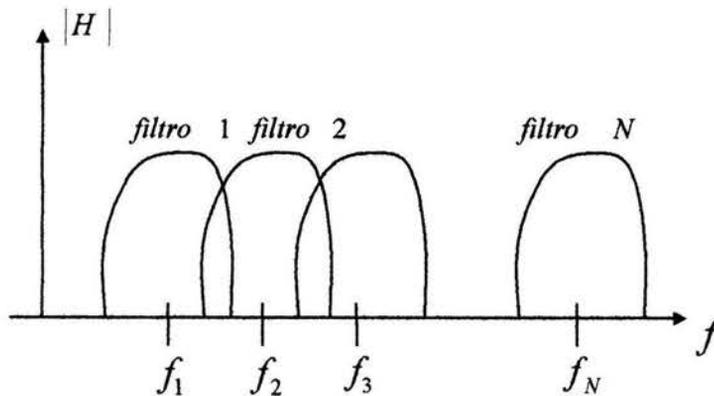


Figura 3.7 Representación de la respuesta en frecuencia de los filtros.

Es el más sencillo de los métodos del dominio de la frecuencia. La señal atraviesa un conjunto de filtros paso – banda (BPF), donde cada sub – banda se pasa a banda baja y se realiza un proceso de decimación. Las sub – bandas se codifican usando algún método basado en el dominio del tiempo. El número de bits asignados a cada banda puede variar en función de la importancia de

dicha banda. En el receptor, se añaden muestras y se vuelve a modular las bandas a sus posiciones originales. Al final, se suman para obtener la señal de voz de salida.

La ventaja que se puede mencionar es que se tiene un mayor control del ruido de cuantización, ya que el ruido generado por cada banda queda confinado en la misma. En la figura 3.8 se muestra un diagrama de codificación por sub – bandas.

La codificación por sub – bandas se utilizan para señales con gran ancho de banda y la CCITT propone el estándar G.722 con una frecuencia de muestreo de 16 [KHz], una codificación ADPCM a 64 [Kbps] y con tres modos de funcionamiento a 48, 56 o 64 [Kbps].²

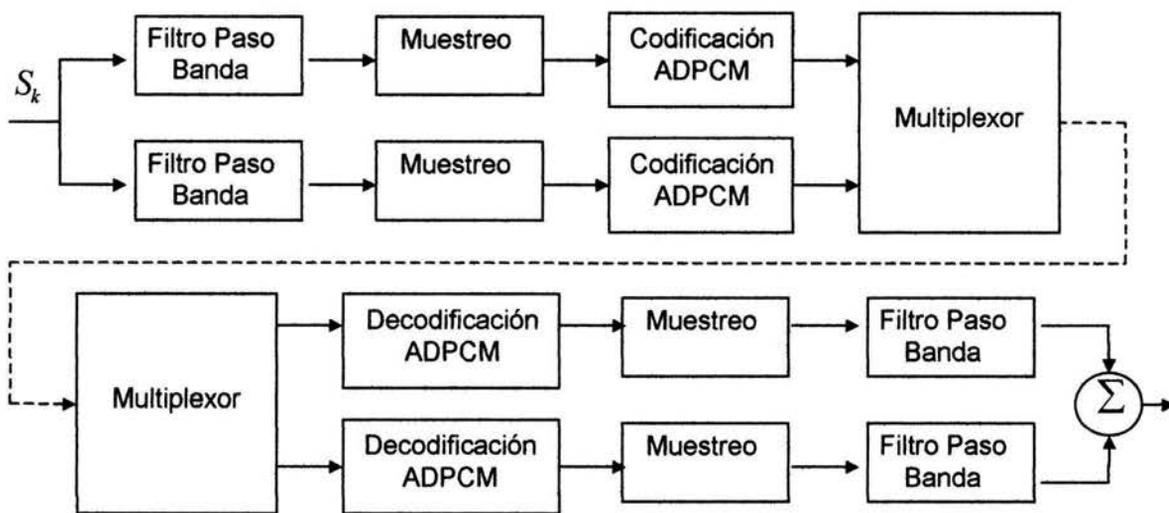


Figura 3.8 Codificador por sub – bandas.

Codificación por transformada

Consiste en una codificación por bloques, donde la señal de entrada se transforma en un dominio diferente, comúnmente Transformada Discreta Coseno y se codifican los coeficientes de transformación. En el receptor, el decodificador calcula la transformada inversa para obtener la señal original reconstruida.

Este tipo de codificación se utiliza en señales de banda ancha de imagen y sonido y muy poco en codificación de voz debido a su complejidad.

² HERRERA, Abel. Apuntes de Comunicaciones Digitales. FI, UNAM, 2000.

3.2.2 Vocoders

Los vocoder aprovechan las características de la señal de voz para realizar una codificación más eficiente basándose en el modelo de producción de voz, el cual se muestra en la figura 3.9

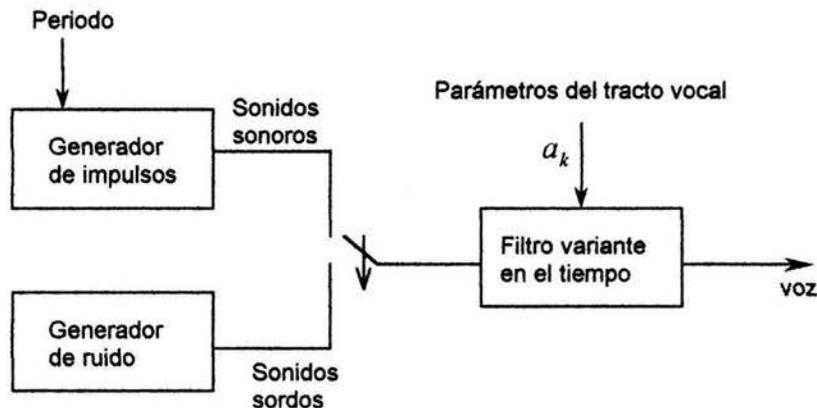


Figura 3.9 Modelo de producción de voz.

Los vocoders independientemente de que si la forma de onda se parece a la original, intentan producir una señal que suene como la original extrayendo los parámetros del tracto vocal y la frecuencia de excitación en el transmisor. Esta información se envía al receptor donde se sintetiza la voz, produciendo una señal de voz con bajo *bit rate*, pero como se menciono anteriormente sonando poco natural.

Vocoder por predicción lineal

Debido a que este vocoder se basa en el modelo de producción de voz antes mencionada es necesario clasificar de forma general los sonidos, es decir determinar si se trata de sonidos sonoros o sordos. En los primeros se abren y cierran las cuerdas vocales, modificando el área de la tráquea y produciendo un tren de impulsos casi periódicos. El período o frecuencia fundamental de este tren es el concepto antes mencionado como *pitch*. Los sonidos no sonoros o sordos los podemos representar como el aire que fluye libremente hasta alcanzar el tracto vocal cuando permanecen abiertas las cuerdas vocales y presenta una estructura ruidosa tanto en el dominio del tiempo como en la frecuencia, además de que la energía de la señal es considerablemente menor que la de los sonidos sonoros. Es por eso que para las señales sonoras, la excitación será un tren de impulsos con la frecuencia fundamental obtenida, mientras que para las señales no sonoras, la

excitación será un ruido aleatorio. Existen diferentes métodos para determinar la frecuencia fundamental, entre los más importantes están: Gold-Rabiner, el algoritmo de autocorrelación, *AMDF* (*Average Magnitude Difference Function*), *SIFT* (*Simplified Inverse Filtering Technique*), *CEP* (*Cepstrum method*), etc. Sin embargo no se puede dejar de lado, que este sigue siendo un problema importante y se continúa investigando para lograr un algoritmo mucho más certero en la obtención del *pitch*.

La parte central de un vocoder de predicción lineal es el modelado del filtro del tracto vocal y secundariamente la obtención de la ganancia.

En *LPC* el tracto vocal es modelado como un filtro digital todo polos que se muestra en la ecuación (3.3), donde p es el orden del modelo o del filtro.

$$H(z) = \frac{G}{1 + a_1 z^{-1} + \dots + a_p z^{-p}} = \frac{S(z)}{E(z)} \quad (3.3)$$

Si definimos a $s(n)$ como la salida del modelo y $e(n)$ como la excitación de entrada, la ecuación 3.3 puede ser escrita en el dominio del tiempo como

$$s(n) = Ge(n) - a_1 s(n-1) - \dots - a_p s(n-p) \quad (3.4)$$

Así, cada muestra de voz es obtenida como una combinación lineal de las muestras previas con una contribución de la excitación.

Sin embargo, se tiene que tomar en cuenta que el método *LPC* solo es aplicable a señales estacionarias, es decir que su comportamiento no cambie con el tiempo y la señal de voz no es una de ellas.

Sin embargo, cuando una señal de voz se analiza en periodos de tiempo muy cortos (de 5 a 100 [ms]), la variación de sus características estadísticas es muy pequeña y se puede considerar una señal cuasi estacionaria.³ Por lo tanto, lo que se hace es segmentar la señal de voz en pequeños segmentos llamadas tramas (*frames*). Esta segmentación se realiza al multiplicar la señal de voz $s(n)$, con una señal ventana $w(n)$, la cual tiene la característica de que es cero fuera del intervalo que nosotros queremos extraer. Algunos ejemplos de ventanas son: rectangular, Bartlett

³ RABINER, L., Fundamentals of Speech Recognition. New Jersey: Prentice Hall, 1993.

(triangular), Hamming, Hanning y Blackman; cada una de ellas con características particulares. La más popular es la ventana Hamming, la cual esta dada por la ecuación 3.5.

$$w(n) = \begin{cases} 0.54 - 0.46 \cos \frac{2\pi}{N} n, & 0 \leq n \leq N-1 \\ 0 & \text{otro valor} \end{cases} \quad (3.5)$$

Donde N es la longitud de la ventana en muestras y es generalmente escogida en un rango de entre 20-40 [ms], con 30 [ms] como valor típico. Así considerando una frecuencia de muestreo de 8000 [Hz], se obtiene una longitud de $N=240$ muestras.

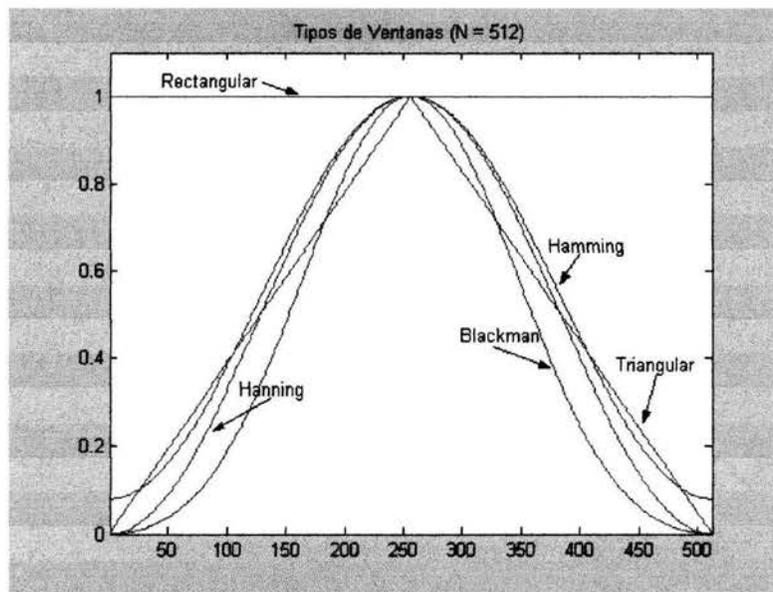


Figura 3.10 Ventana rectangular, Bartlett, Hamming, Hanning y Blackman con $N=512$.

Usualmente se utiliza un traslape entre ventanas, donde la distancia entre ventanas es llamada periodo de trama (*frame period*) con un valor de entre 10-30 [ms]. El tamaño de esta dependerá de la calidad de voz, entre más pequeño sea el periodo de trama, mejor será la calidad, sin embargo esto se reflejará en el *bit rate* debido a que entre más seguido se transmitan las tramas, más alto será el *bit rate*. Este concepto será aplicado posteriormente en el análisis *LPC* por correlaciones.

Inicialmente es necesario aplicar un filtro de preénfasis ya que el modelo *LPC* funcionará adecuadamente con las frecuencias bajas, pero hará un pobre trabajo con las frecuencias altas debido a que en el espectro de la voz existe una caída de -6 [dB/octava], conforme la frecuencia aumenta. Esto se debe a la combinación de una caída de -12 [dB/octava] ocasionada por la fuente

de excitación de la voz y un incremento de + 6 [dB/octava] ocasionado por la radiación de la boca. Esto significa que, cada vez que la frecuencia aumenta al doble, la amplitud de la señal se reduce en un factor de 16. Por lo que se desea compensar esta caída de - 6 [dB/octava] con una manipulación de la señal de voz que de un incremento de + 6 [dB/octava] en el rango apropiado, de manera que la medición del espectro tenga un rango dinámico similar a lo largo de todo su ancho de banda.⁴

La función de transferencia de dicho filtro es $1 - az^{-1}$, el cual enfatiza las altas frecuencias antes de ser procesada. El valor del coeficiente a esta alrededor de 0.9, con un valor típico de $a = 15/16 = 0.9375$.

Por último, para completar la representación del modelo *LPC*, necesitamos obtener los coeficientes del filtro a_i que minimicen el error de predicción cuadrático medio y obtener la ganancia G . Para lo cual, empezamos con definir el segmento de voz y el segmento de error en un tiempo n como

$$\begin{aligned} s_n(m) &= s(n+m) \\ e_n(m) &= e(n+m) \end{aligned} \quad (3.6)$$

Se busca minimizar la señal del error cuadrático medio en el tiempo n

$$E_n = \sum_m e^2(m) \quad (3.7)$$

Al usar la definición de $e_n(m)$ en términos de $s_n(m)$, se puede escribir como

$$E_n = \sum_m \left[s_n(m) - \sum_{k=1}^p a_k s_n(m-k) \right]^2 \quad (3.8)$$

Para resolver esta ecuación, se deriva parcialmente E_n con respecto a cada a_k y el resultado es igualado a cero.

⁴ DELLER, PROAKIS & HANSEN. Discrete Time Processing of Speech Signals. New Jersey: Prentice Hall, 1987

$$\frac{\partial E_n}{\partial a_k} = 0, \quad k = 1, 2, \dots, p \quad (3.9)$$

Resultando:

$$\sum_m s_n(m-i)s_n(m) = \sum_{k=1}^p \hat{a}_k \sum_m s_n(m-i)s_n(m-k) \quad (3.10)$$

Observando que los términos de la forma $\sum_m s_n(m-i)s_n(m-k)$ representan las *covariancias* de los segmentos $s_n(m)$, se pueden escribir como:

$$\phi(i, k) = \sum_m s_n(m-i)s_n(m-k) \quad (3.11)$$

Por lo que la ecuación (3.11) se puede escribir en forma más compacta como:

$$\phi_n(i, 0) = \sum_{k=1}^p \hat{a}_k \phi_n(i, k) \quad (3.12)$$

Esta última ecuación describe un conjunto de p ecuaciones con p incógnitas. Y para obtener los coeficientes de predicción óptimos, se tiene que calcular $\phi_n(i, k)$ para $1 \leq i \leq p$ y $0 \leq k \leq p$ y resolver el conjunto de p ecuaciones simultáneas.

Existen varios métodos para calcular los coeficientes de predicción, algunos de estos son: *covariancia*, *autocorrelación*, *enrejado*, *filtro inverso*, *estimación espectral*, *máxima probabilidad* y *el de producto interno*. El más utilizado es el de autocorrelación, debido a su eficiencia computacional y a su estabilidad inherente produciendo un filtro de predicción cuyos ceros se encuentran adentro del círculo unitario en el plano Z .⁵

Método de autocorrelación

Al aplicar la ventana a la señal de voz se logra determinar los límites de las sumas de las ecuaciones anteriores, por lo que se consigue expresarlas de la siguiente forma

⁵ RABINER, L., *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall, 1993. P. 102-103

$$s_n(m) = s(m+n)w(m) \quad (3.13)$$

$$E_n = \sum_{m=0}^{N+p-1} e_n^2(m) \quad (3.14)$$

y $\phi_n(i, k)$ es definida como

$$\phi_n(i, k) = \sum_{m=0}^{N+p-1} s_n(m-i)s_n(m-k), \quad \begin{matrix} 1 \leq i \leq p \\ 0 \leq k \leq p \end{matrix} \quad (3.15)$$

o por un cambio de variable

$$\phi_n(i, k) = r_n(i-k) = \sum_{m=0}^{N-1-(i-k)} s_n(m)s_n(m+i-k), \quad \begin{matrix} 1 \leq i \leq p \\ 0 \leq k \leq p \end{matrix} \quad (3.16)$$

Además, como la función de autocorrelación es simétrica, $r_n(-k) = r_n(k)$, las ecuaciones LPC pueden expresarse como:

$$\sum_{k=1}^p r_n(|i-k|) \hat{a}_k = r_n(i) \quad 1 \leq i \leq p \quad (3.17)$$

o en forma matricial:

$$\begin{bmatrix} r_n(0) & r_n(1) & r_n(2) & \cdots & r_n(p-1) \\ r_n(1) & r_n(0) & r_n(1) & \cdots & r_n(p-2) \\ r_n(2) & r_n(1) & r_n(0) & \cdots & r_n(p-3) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ r_n(p-1) & r_n(p-2) & r_n(p-3) & \cdots & r_n(0) \end{bmatrix} \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \vdots \\ \hat{a}_p \end{bmatrix} = \begin{bmatrix} r_n(1) \\ r_n(2) \\ r_n(3) \\ \vdots \\ r_n(p) \end{bmatrix} \quad (3.18)$$

Esta matriz de orden $p \times p$ con los valores de autocorrelación, es una matriz *Toeplitz* (simétrica con los elementos de la diagonal principal iguales), que puede resolverse eficientemente con el uso de varios procedimientos numéricos. Uno de ellos es el algoritmo de *Levinson Durbin*.⁶

Algoritmo de Levinson-Durbin

$$E^{(0)} = r(0) \tag{3.19}$$

$$k_i = \frac{\left[r(i) - \sum_{j=1}^{i-1} \alpha_j^{(i-1)} r(i-j) \right]}{E^{(i-1)}} \quad \text{para } 1 \leq i \leq p \tag{3.20}$$

$$\alpha_i^{(i)} = k_i \tag{3.21}$$

$$\alpha_j^{(i)} = \alpha_j^{(i-1)} - k_i \alpha_{i-j}^{(i-1)} \quad \text{para } 1 \leq j \leq i-1 \tag{3.22}$$

$$E^{(i)} = (1 - k_i^2) E^{(i-1)} \tag{3.23}$$

La solución final esta dada por:

$$a_m = \text{coeficientes LPC} = \alpha_m^{(p)} \tag{3.24}$$

$$k_m = \text{coeficientes PARCOR} \tag{3.25}$$

$$g_m = \text{coeficientes log area ratio} \tag{3.26}$$

Es importante resaltar que entre los coeficientes de reflexión o *PARCOR* y los coeficientes *LPC* existe una equivalencia

$$a_i^{(i)} = k_i \tag{3.27}$$

$$a_j^{(i)} = a_j^{(i-1)} + k_i a_{i-j}^{(i-1)} \quad \begin{matrix} i = 1, \dots, p \\ j = 1, \dots, i-1 \end{matrix} \tag{3.28}$$

Por último, para determinar la ganancia G que necesitamos en el modelo de síntesis o de producción de voz, la obtenemos de los parámetros *PARCOR* con la siguiente expresión

$$G^2 = E(p) = (1 - k_1^2)(1 - k_2^2) \dots (1 - k_p^2) r(0) \tag{3.27}$$

Por lo que se podrían transmitir únicamente los parámetros *PARCOR*, $r(0)$ y el *pitch*, para posteriormente en el receptor calcular los coeficientes *LPC* y la ganancia y aplicar el modelo de producción de voz.⁷

LPC 10

Este algoritmo de codificación se le conoce como el estándar americano *FS 1015* fue desarrollado por el Departamento de Defensa Estadounidense a finales de los años 70's. Fue estandarizado primero por dicha institución y posteriormente por la Organización del Tratado del Atlántico Norte (*OTAN*), para convertirse en estándar federal en 1984.

La base de este vocoder, como todos, es el modelo de producción de voz, pero con determinadas características que se enunciarán a continuación. En la figura 3.11 se muestra un esquema del sistema codificador decodificador *LPC10*.

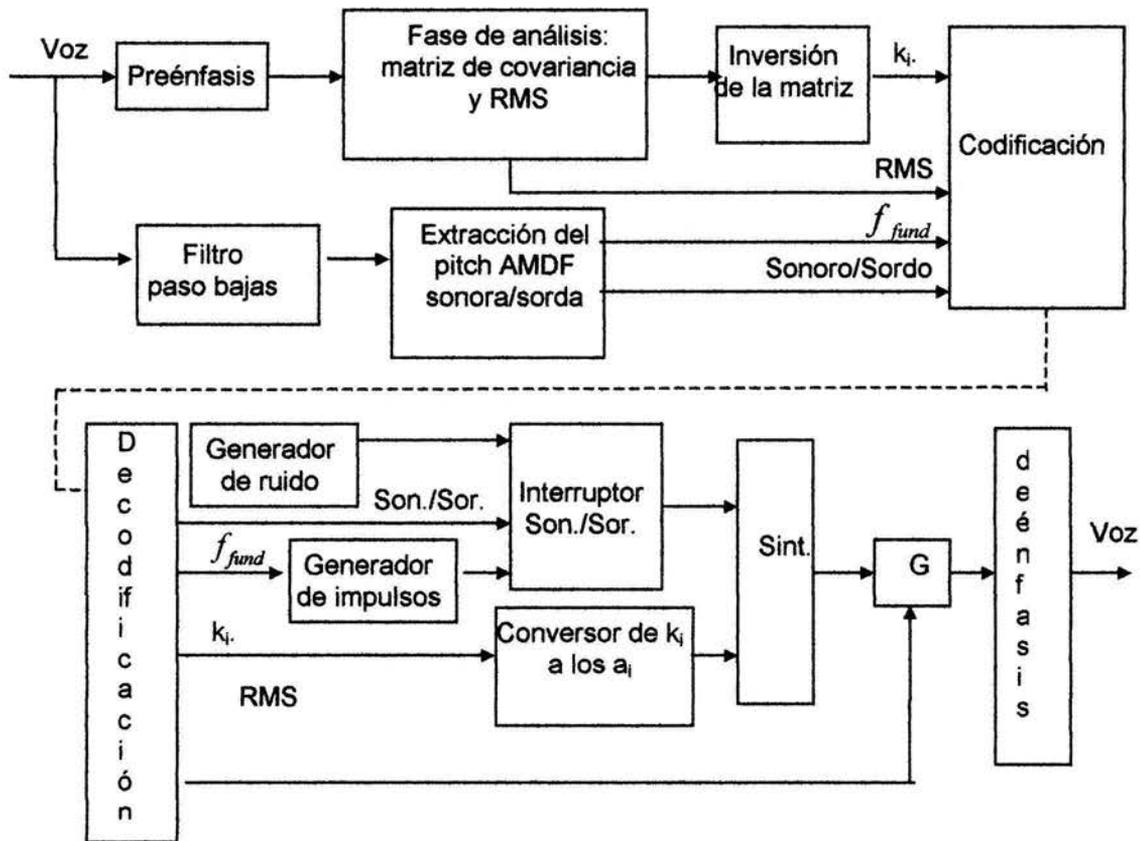


Figura 3.11 Codificador y decodificador *LPC10*

⁶ RABINER, L., *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall, 1993. P. 102-103

A la señal de voz digital se le aplica un filtro de preénfasis, igual al antes mencionado. Para la fase de análisis se utiliza una trama de longitud de 22.5 [ms], se calculan los coeficientes *LPC* mediante el algoritmo de covariancia y con el orden del filtro igual a 10.

Método de covariancia.

Así partiendo de $\phi_n(i,0) = \sum_{k=1}^p \hat{a}_k \phi_n(i,k)$ ecuación (3.12) o lo que es lo mismo en forma matricial

$$\begin{bmatrix} \phi_{1,0} \\ \phi_{2,0} \\ \phi_{3,0} \\ \dots \\ \phi_{p,0} \end{bmatrix} = \begin{bmatrix} \phi_{1,1} & \phi_{1,2} & \phi_{1,3} & \dots & \phi_{1,p} \\ \phi_{2,1} & \phi_{2,2} & \phi_{2,3} & \dots & \phi_{2,p} \\ \phi_{3,1} & \phi_{3,2} & \phi_{3,3} & \dots & \phi_{3,p} \\ \dots & \dots & \dots & \dots & \dots \\ \phi_{p,1} & \phi_{p,2} & \phi_{p,3} & \dots & \phi_{p,p} \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \dots \\ a_p \end{bmatrix} \quad (3.28)$$

o simplemente

$$\Phi_0 = \Phi a \quad (3.29)$$

Así la solución del método de covariancia es obtener la matriz inversa Φ^{-1} y premultiplicarla por el vector Φ_0 como se expresa en la siguiente ecuación

$$a = \Phi^{-1} \Phi_0 \quad (3.30)$$

Hay que notar que la matriz es simétrica, es decir, $\phi(i,k) = \phi(k,i)$ y se pueden utilizar numerosos métodos para obtener la matriz inversa, como por ejemplo el método de descomposición Cholesky.⁸ Como se había mencionado con anterioridad se hace una conversión de los coeficientes *LPC* a a los coeficientes de reflexión o *PARCOR* por medio de la ecuación (3.27)

Posteriormente la señal de voz es transferida a la fase de obtención de la frecuencia fundamental con el algoritmo *AMDF* (*Average Magnitud Diference Function*), el cual se describe a continuación.

⁷ PAPANICHALIS, E. Panos. *Practical Approaches to Speech Coding*. New Jersey: Prentice – Hall, INC, 1987

⁸ RABINER, L., *Fundamentals of Speech Recognition*. New Jersey: Prentice Hall, 1993. P. 102-103

AMDF (Average Magnitud Difference Function).

Este algoritmo está basado en el de autocorrelación, en cual tiene su sustento en la propiedad de que si se tiene una señal con un cierto periodo y a esa señal se le aplica la función de autocorrelación, la señal resultante tendrá el mismo periodo y se podrá obtener midiendo las distancias entre máximos. Sin embargo el aplicar este tipo de algoritmo requiere una carga computacional grande de multiplicaciones y sumas. Por esta razón el *AMDF* fue desarrollado como "una función de autocorrelación de hombres pobres".

El algoritmo *AMDF* está definido para una señal ventaneada $s(n)$, la cual no es cero entre $n = 0$ y $N - 1$ como

$$D(k) = \sum_{n=0}^{N-1} |s(n) - s(n+k)| \quad k = 0, 1, 2, \dots \quad (3.31)$$

De esta ecuación se puede observar que el periodo del *pitch* se logra cuando $s(n)$ y $s(n+k)$ tienen la misma amplitud y por lo tanto $D(k)$ mostrará un valle en vez de un pico si se utiliza la función de autocorrelación, por lo que existe una analogía entre estos dos algoritmos.

La clasificación de los sonidos se lleva a cabo con el análisis de un umbral de energía y la propia frecuencia fundamental con base en las propiedades de sonidos sordos o sonoros.

La información de la ganancia se transmite en forma del valor cuadrático medio de cada trama.

Por último con lo que respecta al emisor, los 10 coeficientes de reflexión se codifican utilizando 41 [bits] para tramas sonoras y 20 [bits] para tramas sordas, para la ganancia son necesarios 5 [bits] y para el periodo fundamental y la sonoridad se utilizan 7 [bits], los que en un total se obtienen 54 [bits/trama] que completan los 2400 [bps].

Ya en el receptor, empieza la reconstrucción de la señal, primero decodificando los parámetros y siguiendo el modelo de producción de voz, simulando la excitación sonora por un tren de impulsos con la frecuencia correspondiente y la sorda como ruido. Se realiza la conversión de los coeficientes *PARCOR* a coeficientes *LPC* y se aplica el filtro del tracto vocal a la excitación, para

que por último la señal obtenida sea escalada en función de la energía *RMS* y obtener la señal de voz.⁹

3.2.3 Codificadores híbridos

La principal ventaja de estos codificadores radica en combinar las técnicas de los codificadores de la forma de onda con las de los vocoders, obteniendo la alta calidad de voz de los primeros y los bajos *bite rates* de los segundos, por lo que en su mayoría contienen o están basados en el esquema de predicción lineal.

También son conocidos como codificadores de análisis por síntesis ya que en el emisor se lleva a cabo un análisis en el cual se obtienen los parámetros de la señal para luego sintetizarla y conseguir el mayor parecido a la original.

Codificación Multipulso

La diferencia central entre los vocoders y este tipo de codificación es que en los primeros como ya se había mencionado, utilizan un modelo muy simple de la excitación, separando los sonidos en sordos o sonoros dando como resultado un sonido muy mecánico de la voz. En este tipo de codificación las tramas no son separadas por sonoras o sordas, sino los que se hace es tratar de determinar una serie de impulsos para cada trama (por lo regular de 8 a 10), determinando la posición y la amplitud de cada impulso sucesivamente para obtener una excitación que hará que al final de la síntesis la señal sea lo más parecida a la original. Cuando un impulso es calculado, su efecto es abstraído de la señal de voz y el procedimiento es repetido.

Si $s(n)$ es la señal original y $s'(n)$ es la señal sintética, entonces para minimizar la energía de la señal error tenemos

$$e(n) = [s(n) - s'(n)] * w(n) \quad (3.32)$$

donde $w(n)$ es una función ponderación y el asterisco denota una convolución. Así en el dominio de la frecuencia,

$$E(z) = [S(z) - S'(z)]W(z) \quad (3.33)$$

⁹ PAPAMICHALIS E., Panos. Practical Approaches to Speech Coding. New Jersey: Prentice-Hall. Inc. 1987. P. 167-168

Una manera usual de escoger la función ponderación es si $H(z)$ es el filtro de tracto vocal LPC, como

$$W(z) = \frac{H(bz)}{H(z)} \quad (3.34)$$

donde

$$H(bz) = \frac{G}{1 + a_1 b^{-1} z^{-1} + \dots + a_p b^{-p} z^{-p}} \quad (3.35)$$

y b es un número entre 0 y 1, donde se ha obtenido experimentalmente que $b = 0.8$ da una buena calidad de voz.

Observando que $S(z)/H(z)$ es la señal residual y $S'(z)/H(z)$ es la excitación de la voz sintética, la cual en este caso es el pulso con características de amplitud A_m , con una posición m a ser determinados. Por otro lado si se define a $h(n)$ como la respuesta al impulso de $H(bz)$ y sustituimos estos valores en la ecuación (3.33), la energía del error minimizada esta dada por

$$\sum_n e^2(n) = \sum_n [d(n) - A_m h(n-m)]^2 \quad (3.36)$$

donde $d(n)$ es el resultado de filtrar la señal residual $r[n]$ a través de $H(bz)$. La minimización es realizada cuando

$$A_m = \frac{\sum_n d(n)h(n-m)}{\sum_n h^2(n-m)} = \frac{C_m}{\phi} \quad (3.37)$$

donde $C_m = \sum d(n)h(n-m)$ es la correlación cruzada entre $h(n)$ y $d(n)$, y $\phi = \sum h^2(n-m)$ es la energía de $h(n)$. La correspondiente energía mínima es

$$E = \left(\sum d^2(n) \right) - \left(\frac{C_m^2}{\phi} \right) \tag{3.38}$$

En la ecuación anterior, C_m es la única cantidad que depende de m . Por lo tanto, para la minimización, la posición m del pulso es seleccionada para que el valor absoluto de la correlación cruzada C_m sea maximizada. Luego, la amplitud del pulso es determinada por la ecuación (3.37). Una vez que este pulso está determinado, es restado de la señal residual. Para determinar el siguiente pulso, se realiza el mismo procedimiento utilizando la última diferencia calculada como la señal residual. En la figura 3.12 se muestra el diagrama de bloques de la obtención de la amplitud y posición del pulso.

Los impulsos calculados son utilizados como la excitación del sintetizador y se logran *bit rates* alrededor de 9,600 [bits/s].¹⁰

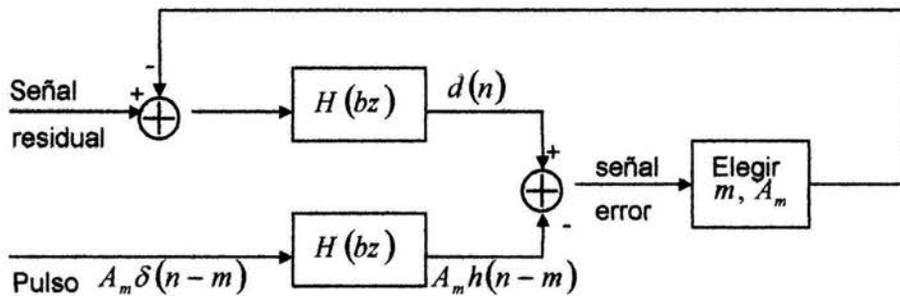


Figura 3.12 Diagrama de la obtención de la amplitud y posición del pulso en un codificador Multipulso

Codificación RELP (Residual Excited Linear Predictive)

Este tipo de codificación nace de la idea de que si se pudiera codificar y transmitir la señal residual, obtenida de hacer pasar la señal de voz por el filtro inverso LPC, se podría tener una muy buena calidad de voz reconstruida, por lo que este concepto es tomado y realizado por el codificador RELP.

La señal residual contiene toda la información de excitación y cualquier otra información que el filtro LPC no considera y esta información es más o menos uniforme sobre las diferentes

¹⁰ PAPANICHALIS E., Panos. Practical Approaches to Speech Coding. New Jersey: Prentice-Hall, Inc. 1987 P.134-136

frecuencias. Como resultado, la necesaria codificación y transmisión eficiente puede ser realizada solo codificando y transmitiendo una pequeña porción del espectro de la señal residual, por ejemplo entre 0-1000 [Hz]. Para poder realizar esto la codificación de la señal residual es llevada a cabo en el dominio de la frecuencia. Así la transformada rápida de Fourier (*FFT: Fast Fourier Transform*) de la señal residual es calculada y se extraen las frecuencias de 0-1000 [Hz]. Estas muestras son codificadas en magnitud y fase para su transmisión.

Después de transmitir dichos datos, en el receptor, una transformada inversa de la *FFT* es calculada para recobrar la banda de la señal residual de la información transmitida, esta banda base es repetida para generar las frecuencias altas en un procedimiento de copiado. Luego la señal residual restaurada es utilizada como la excitación del filtro *LPC*.

Un problema en el procedimiento de copiado es que las armónicas alrededor de 1000 [Hz], 2000 [Hz] y sucesivamente, no estarán apropiadamente espaciadas, debido a que es poco probable que la última armónica de la banda base este exactamente en los 1000 [Hz]. Sin embargo el problema puede ser aligerado si se extrae el *pitch*. Luego en vez de seleccionar la banda base exactamente en 1000 [Hz], esta selección se hace variable, para que las armónicas estén espaciadas correctamente. En este caso obviamente, es necesario transmitir en la información el *pitch*, para que el receptor sepa la longitud de la banda base.

Este incremento del *bit rate* por la transmisión de *pitch*, puede llegar a ser una ventaja si se utiliza para crear un sistema *multirate*, es decir, si la transmisión del canal llega a estar muy cargada, se podría cambiar a un *bit rate* significativamente más bajo y no transmitir la señal residual, sino utilizar exclusivamente el resto de la información (coeficientes *lpc* y *pitch*).¹¹

Por último la codificación *RELTP* utiliza un *bit rate* típico de 9600 [bits/s].

Codificación CELP (Code Excited Linear Prediction)

Este es una de las más eficientes aproximaciones para la codificación de voz de alta calidad a *bit rates* bajos. Esta codificación esta basada en procedimientos de búsqueda de análisis por síntesis, cuantización de vectores con pesos (*VQ*) y predicción lineal. Es utilizado un filtro *LPC* de décimo orden para modelar los formantes de retardo corto de la señal de voz: la periodicidad de retardo largo de la señal se modela con un diccionario adaptivo *VQ*. El error de la predicción lineal de retardo corto y el *pitch* *VQ* se cuantifican usando un libro de secuencias estocásticas, es decir

¹¹ PAPANICHALIS E., Panos. Practical Approaches to Speech Coding. New Jersey: Prentice-Hall, Inc. 1987 P.136-139

muestras de ruido aleatorio blanco gaussiano con media cero y varianza uno. Por lo que al receptor solo se envía como excitación el índice del diccionario y la ganancia.

En la figura 3.13 se puede observar la etapa de análisis de un transmisor CELP. A la señal de entrada, se le resta la salida del sintetizador CELP, esta diferencia se pasa por un filtro perceptual. El error perceptual se utiliza después para realizar la búsqueda en el diccionario. El procedimiento de búsqueda encuentra las ganancias de los diccionarios estocásticos y adaptivos que minimizan el error cuadrático medio. El filtro de predicción lineal puede determinarse por técnicas de análisis de retardo corto de la señal de entrada.

El objetivo del filtro perceptual es atenuar aquellas frecuencias en las que el oído humano detecta menos los posibles errores y atenuarlas para aquellas en que el oído es más perceptivo. Es decir, se lleva a cabo una comparación dando más importancia a aquellas zonas donde el oído aprecia más el ruido.

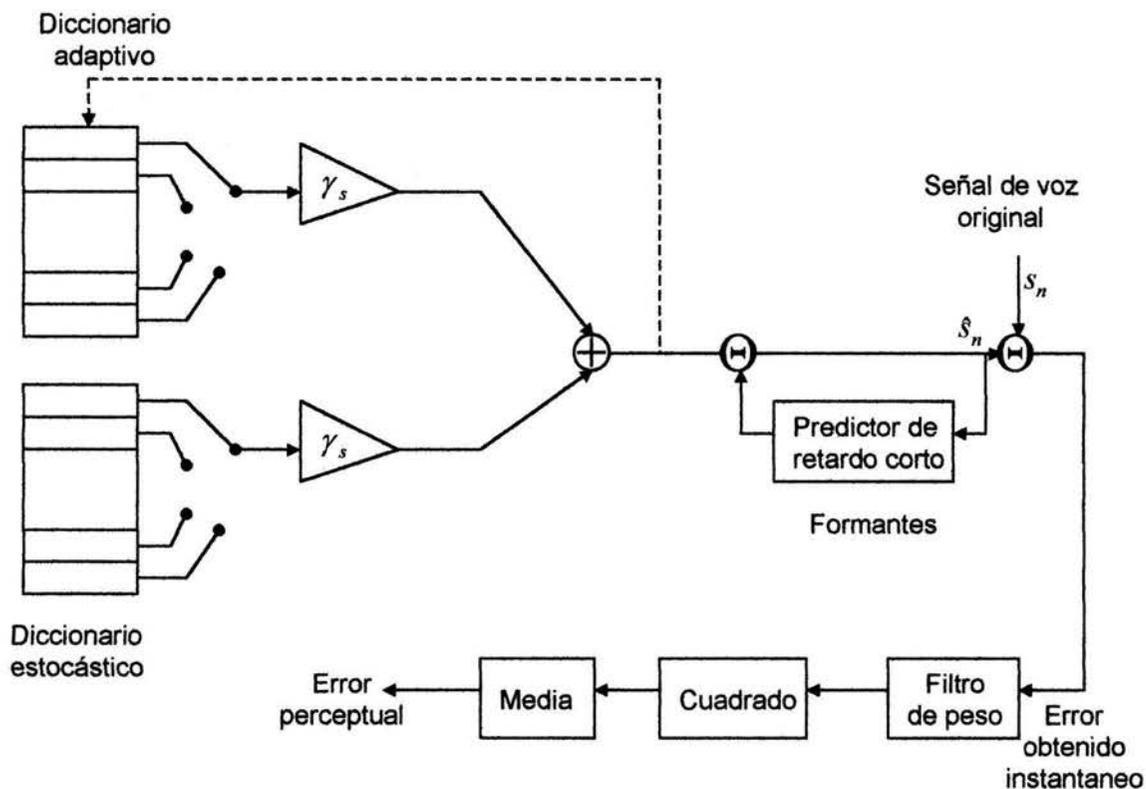


Figura 3.13 Etapa de análisis de un transmisor CELP

El sintetizador *CELP* se basa en el principio del uso de un filtro de predicción lineal cuya ganancia se busca en una tabla. La excitación se genera usando, en paralelo, un diccionario estocástico fijo y otro diccionario adaptivo. Posteriormente, se actualiza el diccionario adaptivo con esta excitación.

Así resumiendo el procedimiento esencial de una señal en un típico codificador *CELP*. Por ejemplo, el estándar *FS-1016 CELP* utiliza muestras a 8 [KHz], tiene una trama de 30 [ms], es decir 240 [muestras/trama] y reserva 10 bits para 1024 vectores código y su análisis consiste en:

1. Una predicción por trama de análisis corto, utilizando el análisis *LPC* de décimo orden.
2. Búsqueda del *pitch* cuatro veces por trama (cada 7.5 [ms]).
3. Búsqueda en 1024 vectores código (en un lazo cerrado), buscando cuatro veces por trama para seleccionar el óptimo index del diccionario para minimizar el peso del error medio cuadrático.

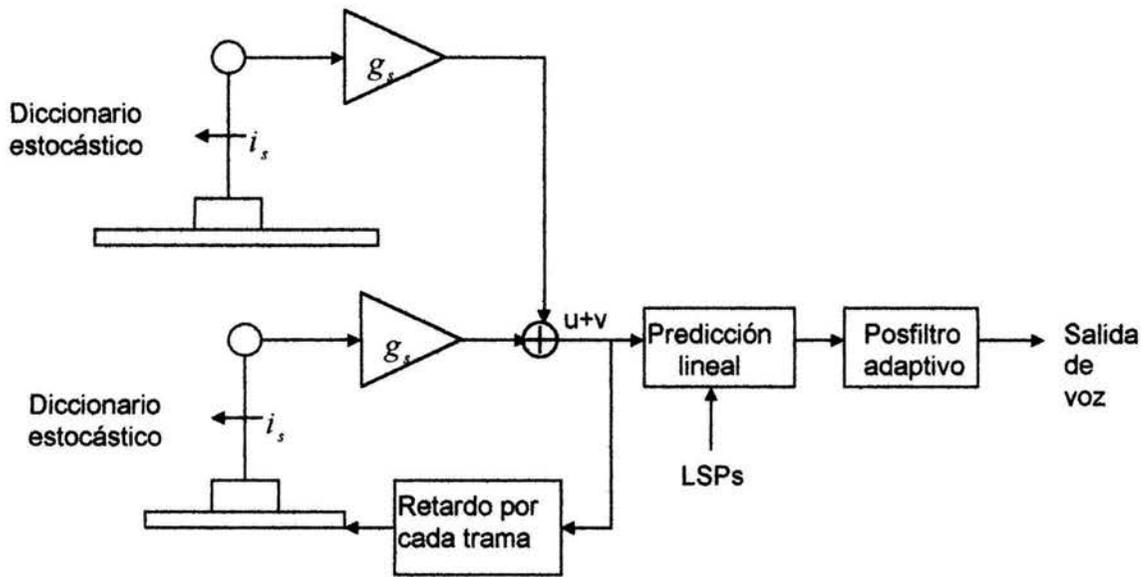


Figura 3.14 Decodificador CELP

Codificación VSELP (Vector Sum Excited Linear Predictive)

El esquema de *VSELP* conocido como el estándar de *Half-Rate GSM* se muestra en la figura 3.15, donde existen 4 diferentes modos de operación y dos diferentes diagramas de bloques para dichos modos. Los coeficientes del filtro de síntesis son determinados cada 20 [ms] con orden 10 ($p=10$), este intervalo es dividido en cuatro subsegmentos de 5 [ms] para la optimización de la excitación.

A cada uno de los cuatro modos de síntesis antes mencionados, les corresponde diferentes modos de excitación, con lo cual implica una determinación de diferentes grados de sonoridad en la señal de voz, esta determinación esta basada en la ganancia LTP, la cual es típicamente alta para segmentos sonoros altamente correlacionados y baja para segmentos sordos no correlacionados como el ruido.

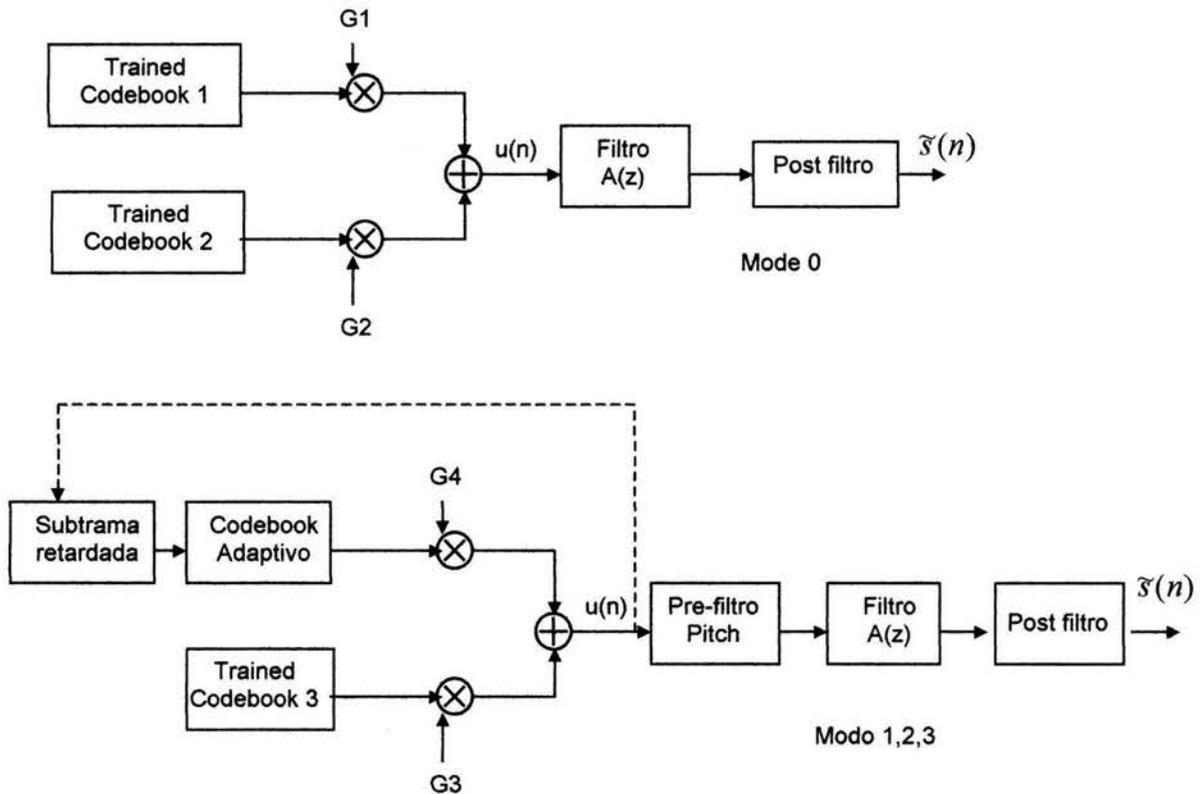


Figura 3.15 Esquema del codificador VSELP half-rate GSM

En el modo 0 o sordo, la señal de voz es sintetizada por la superposición de las dos salidas escaladas por G_1 y G_2 de los *trained codebooks* de 128 entradas para generar la señal de excitación, la cual es filtrada por el filtro de síntesis $A(z)$ y un post - filtro espectral.

En los modos 1, 2 y 3 donde la entrada de voz tiene algún grado de sonoridad, la excitación es ahora generada por la superposición de un *trained codebook* de 512 entradas escalado por G_3 y un *codebook* adaptivo escalado por G_4 .

En GSM VSELP, los 10 coeficientes de reflexión que son codificados para cada trama son agrupados en tres vectores $v_1 = [r_1 r_2 r_3]$, $v_2 = [r_4 r_5 r_6]$ y $v_3 = [r_7 r_8 r_9 r_{10}]$. Los vectores son cuantizados secuencialmente, de v_1 a v_3 , usando un VQ codebook (C_i) b_i -bit para v_i , donde b_i , $i = 1,2,3$ son 11,9 y 8 b respectivamente. El vector v_i es cuantizado para minimizar la energía del error de predicción del j th estado del filtro lattice donde r_j es el orden más alto del coeficiente en el vector v_i . La complejidad computacional asociada con cuantizar v_i es reducida por buscar solo un pequeño subconjunto del vector código de C_i . El subconjunto es determinado primero al buscar en un codebook precuantizador de tamaño c_i bits, donde $i = 1,2,3$ son 6, 5 y 4 b respectivamente. Cada vector en el codebook precuantizador esta asociado con $2^{b_i-c_i}$ vectores en el target codebook. El subconjunto es obtenido de combinar todos lo vectores código en C_i que están asociados con la mejor concordancia de los vectores precuantizados. De esta manera, un factor de reducción en complejidad computacional cercana a $2^{b_i-c_i}$ es obtenida para la cuantifican de v_i .

| Parámetros | Bits/trama |
|--------------------------------|--------------|
| Coeficientes LPC | 28 |
| Bandera de interpolación LPC | 1 |
| Modo de excitación | 2 |
| Modo 0: | |
| Índice de Codebook1 | 4x7=28 |
| Índice de Codebook2 | 4x7=28 |
| Modos 1,2,3 | |
| LTPD (subtrama 1) | 8 |
| Δ LTPD (subtrama 2,3,4) | 3x4=12 |
| Índice de Codebook 3 | 4x9=36 |
| Energía de la Trama | 5 |
| Ganancia de excitación | 4x5=20 |
| Total no. De bits | 112bits/20ms |
| Bitrate | 5.6 [Kbps] |

Tabla 3.1 Bitrate de VSELP Half-Rate GSM¹²

¹² HANZO, Lajos. Voice Compression and Communications. NY: John Wiley & Sons, Inc. 2001. P 234.

Con este tipo de codificación y con base en la tabla 3.1 obtenemos un *bitrate* de 5.6 [Kbps], sin embargo la carga computacional debido a la complejidad del algoritmo es de 13.5 [MIPS] (millones de instrucciones por segundo). Lo cual implica que como ya habíamos visto en el capítulo 2, que el retardo sea mayor y por lo tanto provee un servicio regular, por lo que existen muy pocas implementaciones de este codificador.

Enhanced Full-Rate GSM.

Este codificador está basado en el modelo de excitación *ACELP* y su funcionamiento empieza con el preénfasis de la señal de voz utilizando un filtro paso altas para aumentar la baja energía de los componentes de las frecuencias altas antes que la señal de voz sea procesada. De nueva cuenta la cuantización espectral es realizada trama por trama mientras que la optimización de la excitación es calculada por subsegmentos.

Para la cuantización espectral se utiliza un análisis *LPC* de orden 10 el cual es realizado dos veces por trama de 20 [ms], utilizando dos diferentes ventanas asimétricas de 30 [ms] de duración. Los coeficientes de reflexión son calculados de la señal ventaneada con el algoritmo de Levinson Durbin, posteriormente los coeficientes *LPC* son convertidos a *LSFs* y cuantizados usando *SMQ* (*split matrix quantizer*). Este método consiste en que el término largo de *LSF* es quitado, definiendo los vectores p_1^n y p_2^n para la trama n , correspondiendo a las dos ventanas antes mencionadas. Posteriormente el conjunto de tramas n *LSF* es predecido de el conjunto *LSF* previamente cuantizado \tilde{p}_2^{n-1} , tomando en cuenta la correlación de 0.65 de sus términos largos. Ambos *LSFs* vectores diferencia son luego entrada a la *SMQ*. Una submatrix de 2×2 de los primeros dos *LSF* de ambos vectores *LSF* y cuantizados buscando a través de un codebook de 128 entradas.

Similarmente, el tercer y cuarto vector *LSF* de ambos *LSF* son cuantizados utilizando un codebook de 256 entradas. Finalmente, después de encontrar la mejor entrada del codebook para todas las submatrices de 2×2 , los valores previos predecidos son sumados para producir el vector *LSF* cuantizados, \tilde{p}_1^n y \tilde{p}_2^n respectivamente.

Con lo que respecta a la búsqueda en el codebook adaptivo, un análisis del pitch de lazo cerrado y abierto son usados como sigue: Con base en la voz ponderada una búsqueda del pitch en lazo abierto se lleva a cabo dos veces por cada trama de 20 [ms] o una vez cada dos subtramas favoreciendo valores de bajas frecuencias de muestreo con el objeto de evitar que se repita dicha frecuencia de muestreo.

Posteriormente una búsqueda de lazo cerrado para valores de frecuencia de muestreo entera es conducida sobre una base de subtramas, esto se restringe al rango $[T_0 \pm 3]$ en la primera y tercer subtrama, con el objeto de mantener una baja complejidad de búsqueda. Como las subtramas segunda y cuarta el lazo cerrado de búsqueda se concentra alrededor de los valores de la frecuencia de muestreo de la subtrama previa en el rango de $[-5 \dots +4]$.

Por último los retardos de la frecuencia de muestreo fraccional se prueban también alrededor del valor de lazo cerrado en las subtramas segunda y cuarta, aun que únicamente para los retardos de la frecuencia de muestreo debajo de 95 en las subtramas primera y tercera, correspondiendo a las frecuencias de la frecuencia de muestreo en exceso de 84 [HZ].

Ya determinado el lazo óptimo, la entrada del *codebook* adaptivo es únicamente identificada, mientras que su ganancia es restringida al recorrido de $[0 \dots 1.2]$ y cuantizada.

Por último tocaremos el tema de la optimización para búsqueda del *codebook* fijo, se utiliza el mismo principio que en ACELP, sin embargo los codificadores conocidos como de la familia ACELP no son tocados en este trabajo.

En este tipo de codificación se logra un bitrate de 12.2 [kbps] con base en la siguiente tabla.

| Parámetros | 1er. y 3er. subtrama | 2da y 4ta subtrama | No de bits | Total (kbps) |
|-------------------------|-------------------------|--------------------|--------------------------------|--------------|
| Dos conjuntos de LSF | | | 38 | 1.9 |
| Ganancia de CB fijo | 5 | 5 | $4 \times 5 = 20$ | 1 |
| ACELP | 35 | 35 | $4 \times 35 = 20$ | 7 |
| Índice de CB adaptivo | 9 | 6 | $2 \times 9 + 2 \times 6 = 30$ | 1.5 |
| Ganancia de CB adaptivo | 4 | 4 | 16 | 0.8 |
| Total | | | 240/20ms | 12.2 |

Tabla 3.2 Bitrate de la codificación Enhanced full – rate GSM

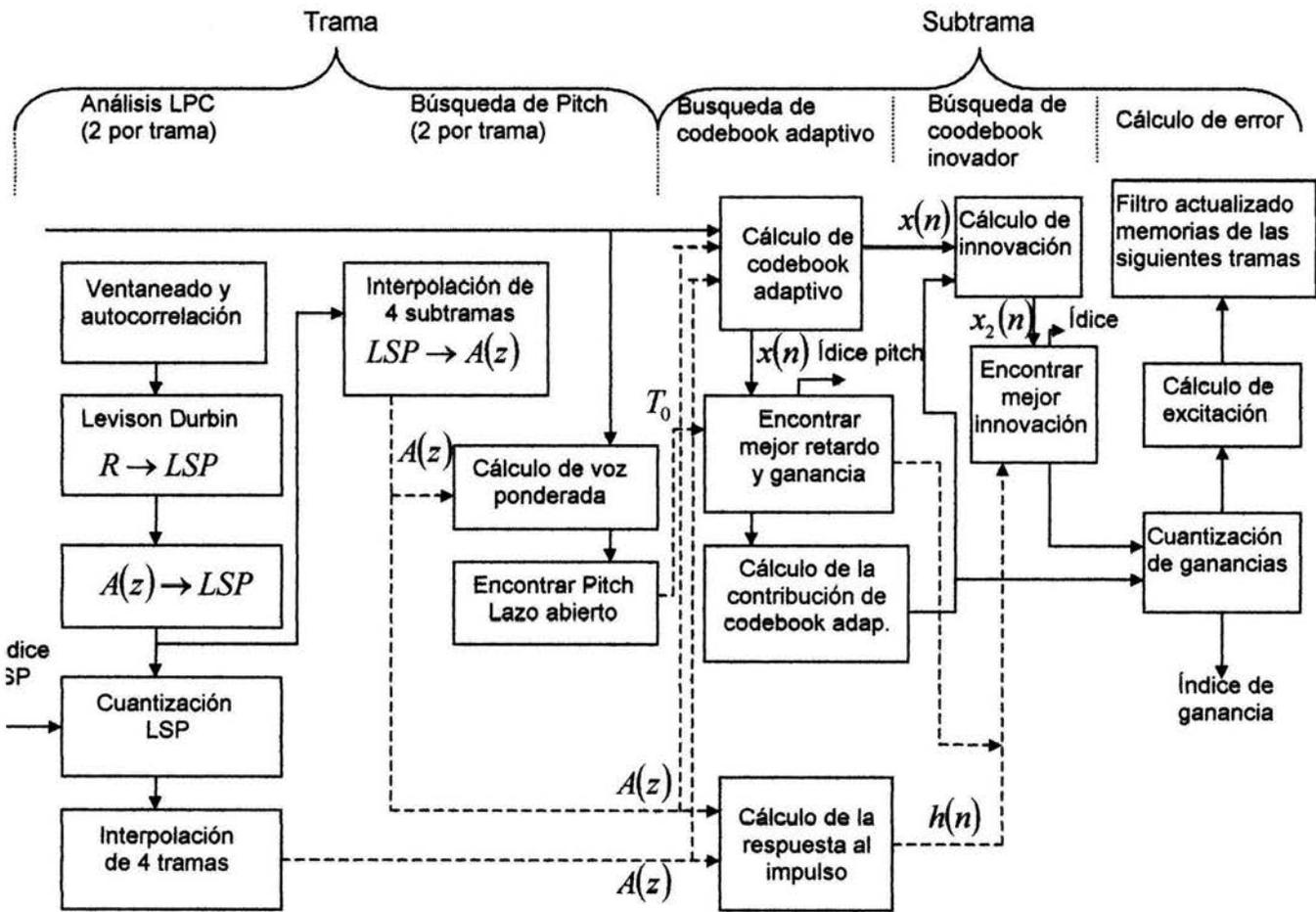


Figura 3.16 Codificador Enhanced full – rate GSM

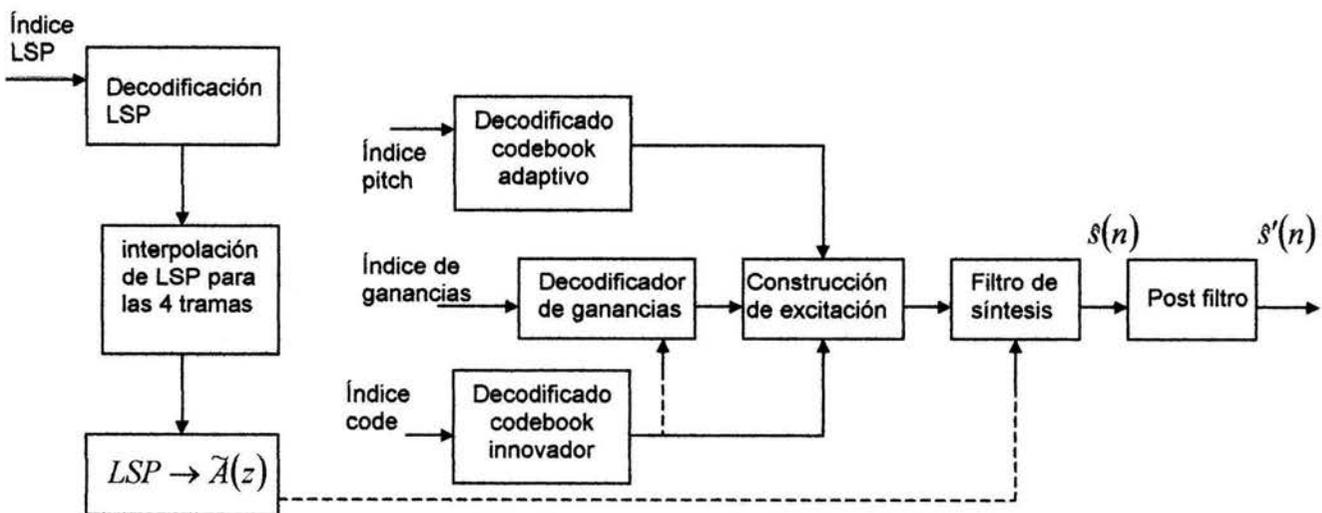


Figura 3.17 Decodificador Enhanced full – rate GSM

Y se habla que tiene una calidad igual e inclusive superior al codificado RPE – LTP, sin embargo esta codificación no es usada como estandar quizá debido a la complejidad de su algoritmo que se traduce en más retraso temporal y más utilización de recursos y por supuesto a una comunidad conservativa con respecto a los cambios.¹³

Codificación RPE – LTP (Regular Pulse Excitation – Long Term Prediction)

La codificación RPE – LTP es una recomendación ETSI GSM 6.10 “GSM Full Rate Speech Transcoding”. RPE – LTP combina las ventajas de un codificador RELP con las de un codificador MPE – LTP (Multi – Pulse Excited Long Term Prediction). Las ventajas de combinación, son que RELP entrega una buena calidad de voz con poca complejidad, pero hay que tomar en cuenta que la calidad de voz sigue siendo limitada debido al ruido tonal por la regeneración de altas frecuencias. Por otro lado la codificación MPE – LTP consigue una alta calidad de voz, pero tiene una complejidad alta. Así modificando RELP para incorporarle algunas características de MPE – LTP, se pudo obtener una reducción de 14.77 [Kbps] a 13 [Kbps] sin pérdida en la calidad de voz. Siendo la modificación más importante la introducción del ciclo LTP.

Así este codificador utiliza un predictor lineal para eliminar la correlación entre tramas, posteriormente mejora la calidad de la señal con un predictor de retardo largo y una secuencia de excitación de patrones de pulsos regularmente espaciados como se muestra en la figura 3.18.

Pre – procesamiento

Como principio, a la señal de voz se aplica un filtro de pre – énfasis de primer orden necesario para el análisis LPC. De la señal resultante se toman tramas de voz de 20 [ms], que a una frecuencia de muestreo de 8000 [Hz], obtenemos 160 muestras por trama a las cuales se les aplica una ventana de Hamming, para disminuir el efecto producido en el dominio de la frecuencia por la oscilación de Gibbs, causada por el truncamiento de la señal de voz fuera de la trama analizada.

Análisis de predicción lineal de tiempo corto

A las tramas resultantes del pre – procesamiento se les aplica un análisis LPC de orden $p = 8$. por lo que como primer paso se obtienen $p + 1$ coeficientes de autocorrelación para cada trama, a

¹³ MOULY, Michel. Current Evolution of the GSM Systems. IEEE Personal Communications. October 1995.

partir de los coeficientes de correlación, son calculados los coeficientes de reflexión k_i utilizando la recursión de Schur.

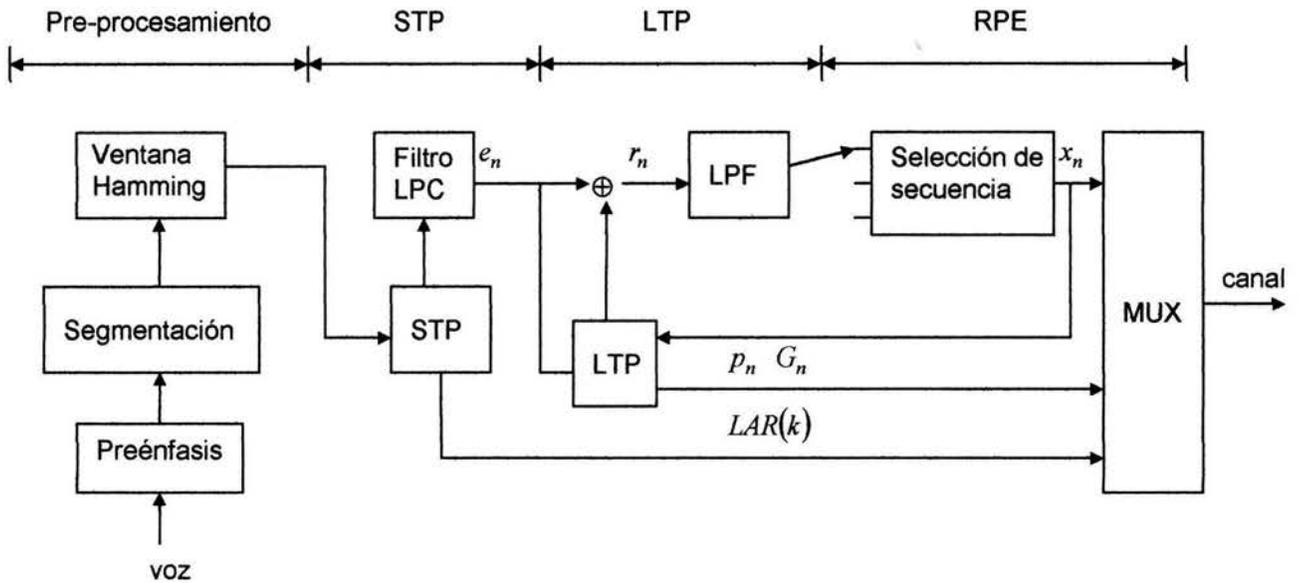


Figura 3.18 Codificador RPE – LTP

Posteriormente los coeficientes de reflexión son transformados en los LAR para ser transmitidos por el canal y a la vez alimentados a un filtro inverso LPC , para reconstruir la señal con el error residual como en $RELTP$.

Filtro de predicción de retardo largo LTP

El objetivo principal de este modulo es minimizar el error de predicción de retardo largo y se consigue por aquel retardo d , que maximiza la correlación cruzada entre el residuo actual $r_{STP}(n)$ y los residuos previamente recibidos $r_{STP}(n-d)$. Para lograr esto las 160 muestras del residuo de retardo largo se dividen en cuatro subsegmentos de 40 muestras y para cada uno de estos se calcula una correlación cruzada entre el segmento actual y las 128 muestras de residuos STP precedentes. El máximo de la correlación se encuentra para un retardo d , que con mucha probabilidad corresponde al periodo del pitch o a un múltiplo de éste. Por lo que restando el segmento altamente correlacionado multiplicado por un factor de ganancia G , que es la correlación cruzada normalizada para el retardo d , se puede eliminar buena parte de la redundancia, una vez obtenidos G y d se cuantizan con 2 y 7 bits respectivamente, llamándolos G' y d' .

G' y d' son descodificados localmente para obtener G'' y d'' y poder producir el residuo localmente descodificado STP $r'_{STP}(n)$. El residuo $r_{LTP}(n)$ se calcula como la diferencia entre el residuo $r_{STP}(n)$ y la estimación $r''_{STP}(n)$, que se obtiene con los parámetros G'' y d'' , con las siguientes expresiones

$$r_{LTP}(n) = r_{STP}(n) - r''_{STP}(n)^{n=1..40} \quad (3.39)$$

$$r''_{STP}(n) = G'' r'_{STP}(n-d)^{n=1..40} \quad (3.40)$$

donde $r'_{STP}(n-d)$ representa un segmento pasado. Finalmente esa representación se actualiza usando el residuo $r'_{LTP}(n)$ y la estimación del residuo $r''_{STP}(n)$ para formar $r'_{STP}(n)$

$$r'_{STP}(n) = r'_{LTP}(n) + r''_{STP}(n)^{n=1..40} \quad (3.41)$$

Módulo RPE

El residuo $r_{LTP}(n)$ se pesa con un filtro fijo que es esencialmente un filtro paso bajas con frecuencia de corte de $4 \text{ [KHz]}/3 = 1.33 \text{ [KHz]}$ y poder hacer una decimación por un factor de 3. La señal residual filtrada $r_{FLTP}(n)$ se descompone en tres excitaciones candidatas, descartando la muestra número 40. Se calcula las energías de cada secuencia decimada y la que tiene mayor energía se elige como representante del residuo LTP . Los pulsos de excitación son posteriormente normalizados a la máxima amplitud $v_{max}(k)$ de las trece muestras y son cuantizadas uniformemente con tres bits, mientras que el máximo de la amplitud se cuantiza logarítmicamente con seis bits. Como se tienen tres posibles posiciones iniciales de la secuencia decimada elegida, se necesitan dos bits para codificar la posición inicial de cada subsegmento. Esto se conoce como posición de *grid*. Las amplitudes de los pulsos $\beta(k, i)$, la posición de *grid* y los máximos de la amplitud son localmente decodificados para obtener el residuo $r'_{LTP}(n)$, donde la secuencia decimada se rellena con ceros.

Con lo que respecta al decodificador, en la figura 3.19 se muestra el diagrama de bloques de dicho decodificador, el cual cuenta con una decodificación *RPE*, síntesis del filtro *LTP*, síntesis del filtro *STP* y postprocesado.

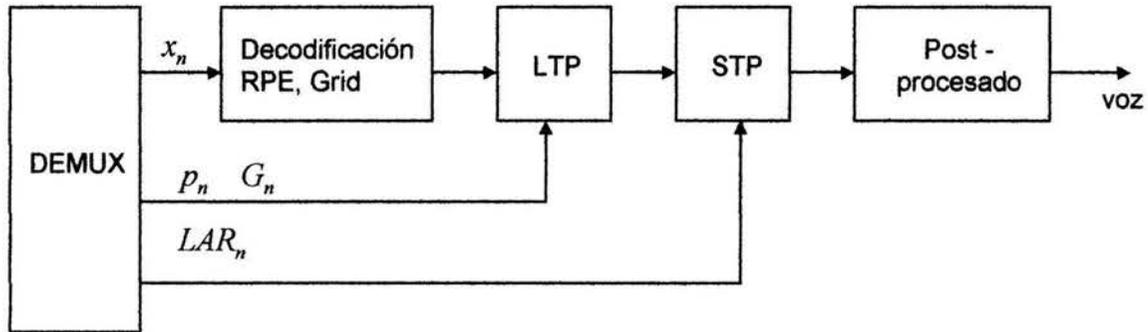


Figura 3.19 Decodificador RPE – LTP

Decodificación RPE

En el decodificador, las amplitudes de los pulsos $\beta(k, i)$, la posición del *grid* y los máximos de la amplitud son localmente decodificados y se calculan las amplitudes de los pulsos actuales multiplicando su amplitud actual por la amplitud máxima del bloque. El modelo del residuo $r'_{LTP}(n)$ se recupera colocando apropiadamente las amplitudes de los pulsos de acuerdo con el *grid* inicial.

Síntesis del filtro LTP

Inicialmente, los parámetros del filtro LTP (G', d') son cuantizadas inversamente para derivar los parámetros del filtro LTP. A continuación, el modelo de excitación recuperado $r'_{LTP}(n)$ se utiliza para excitar el filtro y recuperar un nuevo subsegmento de longitud 40 del residuo estimado $r'_{STP}(n)$. Para realizar esto se usa la historia del residuo recuperado $r'_{STP}(n)$, retardado por d' muestras y multiplicado por G' , y así se obtiene la estimación del residuo $r''_{STP}(n)$ de acuerdo con

$$r''_{STP}(n) = G' r'_{STP}(n - d') \tag{3.42}$$

y entonces $r''_{STP}(n)$ se usa para calcular el residuo del subsegmento recuperado más reciente como

$$r'_{STP}(n) = r''_{STP}(n) + r'_{LTP}(n) \quad (3.43)$$

Para la síntesis del filtro STP los parámetros $LAR'(i)$ son decodificados utilizando el cuantizador inverso y convertidos a los coeficientes de reflexión por medio de la siguiente expresión

$$k(i) = \frac{10^{LAR'(i)} - 1}{10^{LAR'(i)} + 1} \quad (3.44)$$

Por último el post-procesamiento es desarrollado por el deénfasis, utilizando el filtro inverso de énfasis.¹⁴

¹⁴ HANZO, Lajos. Voice Compression and Communications. New York: John Wiley & Sons, Inc. 2001. P 133-137

3.3 Simulación de codificaciones

La simulación de los codificadores se llevó a cabo con base en los algoritmos descritos en la sección pasada, con lo que respecta a los codificadores utilizados en GSM, con base en los puntos obtenidos en el capítulo 2 y capítulo 3, se determino implementar el algoritmo más utilizado y por lo tanto el más importante de los tres codificadores antes mencionados y sus antecesores: DPCM, ADPCM, Vocoder LPC, LPC10, RELP, Multipulso y por RPE – LTP . Estos algoritmos se desarrollaron en el Laboratorio de Procesamiento de Voz del edificio de Posgrado de la Facultad de Ingeniería. Se utilizó la versión 6 de Matlab con una computadora pentium IV.

A continuación se mencionan las características de dichas implementaciones y se muestran las gráficas de la palabra original y la palabra codificada en tiempo y en frecuencia. Para todas las codificaciones se utiliza la palabra nueve con mi voz en formato "wav" y una frecuencia de muestreo de 8000 [Hz] a 16 bits mono.

3.3.1 DPCM

Para este codificador se utilizan 128 niveles y un predictor de primer orden $\hat{x}[n] = a_1 \hat{x}[n-1]$,

donde $a_1 = \frac{2 * x_{\text{máx}}}{M}$, donde M es el número de niveles. Se desarrolló en el programa dpcm128.m

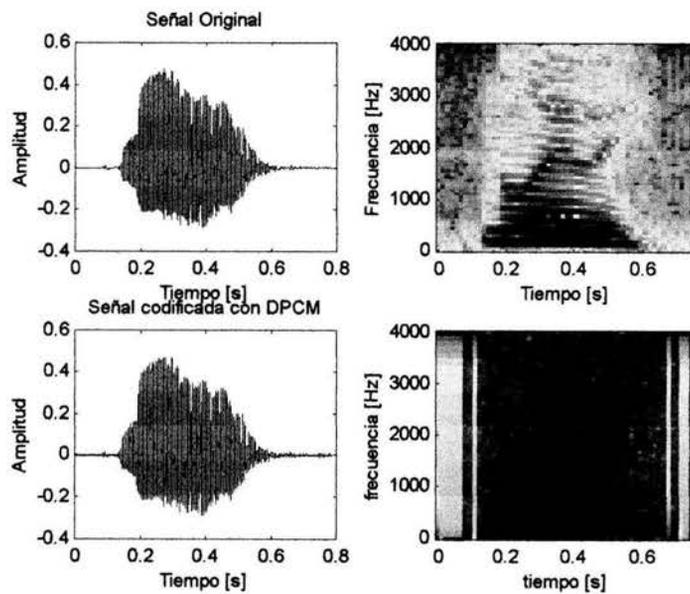


Figura 3.20 Señal original y señal codificada con DPCM con espectrograma

3.3.2 ADPCM

Para esta codificación se utilizó ADPCM *backward*, adaptando el paso con el predictor de primer orden utilizado en DPCM y con 128 niveles. El desarrollo está en el programa `adpcm128.m`

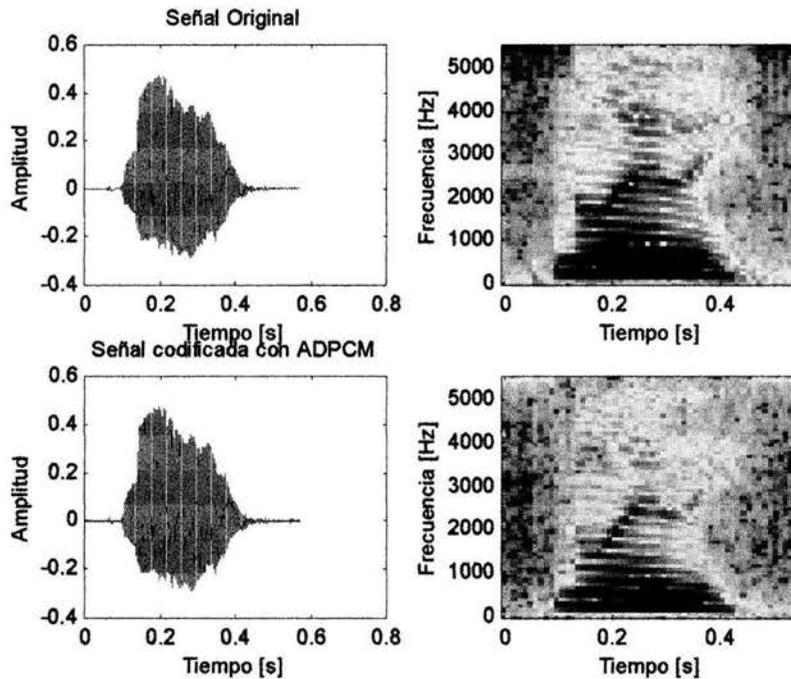


Figura 3.21 Señal original y señal codificada con ADPCM con espectrograma

3.3.3 Vocoder LPC

Para la implementación de este codificador primero se implementó un filtro paso bajas con frecuencia de corte de 3400 [Hz]. Posteriormente un filtro de preénfasis con coeficiente $a=0.95$.

Las tramas fueron de 30 [ms] con un traslape de 10 [ms] entre tramas utilizando una ventana Hanning.

Para el análisis de los coeficientes del filtro LPC se utilizó el método de autocorrelación con el método de Levinson – Durbin de orden 8.

También se utilizó el método de autocorrelación para la obtención de la frecuencia fundamental y la discriminación entre tramas sordas o sonoras. El desarrollo está en el programa `vocoderlpc.m`

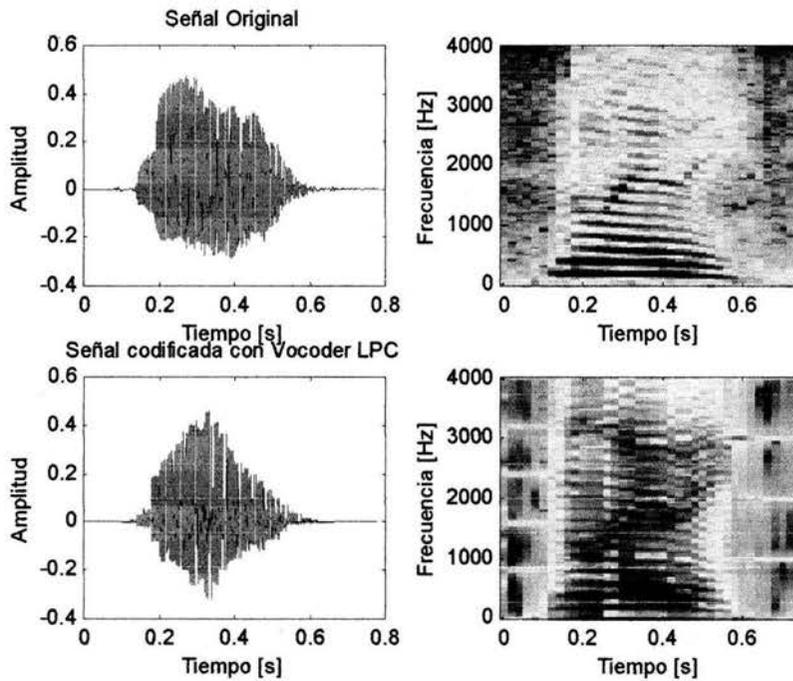


Figura 3.22 Señal original y señal codificada con Vocoder LPC y espectrograma

3.3.4 Vocoder LPC10

En este caso, siguiendo con el estándar, las tramas fueron de 22.5 [ms] sin traslape utilizando de nuevo una ventana Hanning.

Para el análisis LPC se utiliza el método de covariancias con un orden de 10. Así como para la obtención de la frecuencia fundamental se utilizó el método AMDF antes descrito.

En la figura 3.23 se muestra el resultado gráfico de esta codificación. Se prueba con el programa lcp10.m

3.3.5 Codificación RELP

En esta codificación se utiliza un análisis LPC de orden 8, con tramas de 30 [ms] sin traslape, utilizando una ventana de Hanning para el análisis y usando la FFT para obtener las muestras de la señal de 0-1000 [Hz].

En la figura 3.24 se muestra el resultado gráfico de esta codificación. El programa desarrollado es vorelp.m

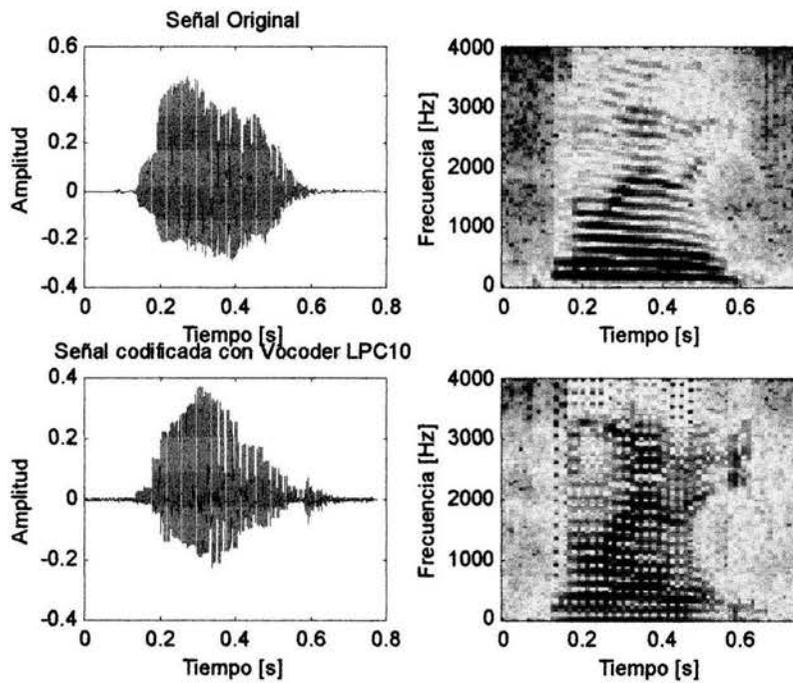


Figura 3.23 Señal original y señal codificada con Vocoder LPC10 y espectrograma

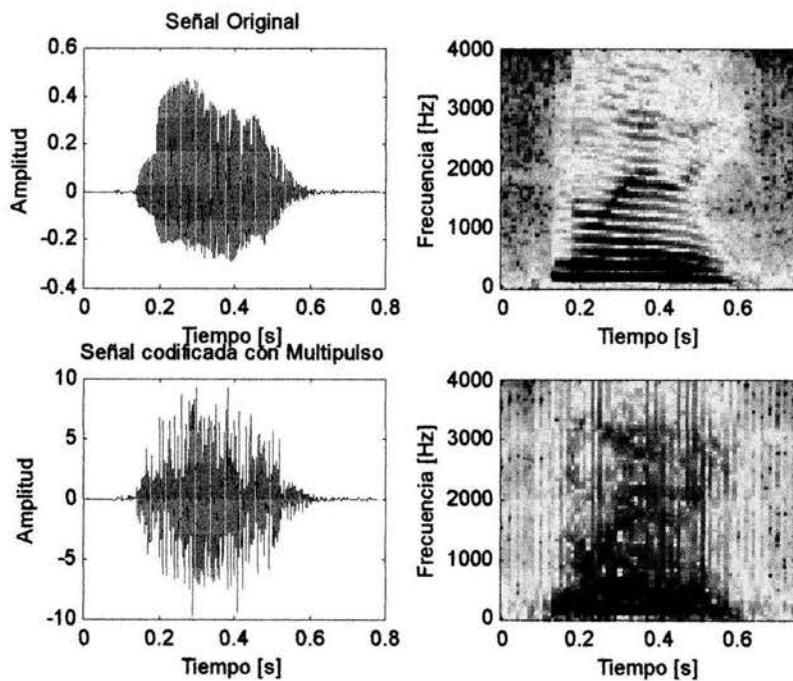


Figura 3.24 Señal original y señal codificada con RELP y espectrograma

3.3.6 Codificación Multipulsos.

En multipulso, para el análisis LPC se utilizó el método de autocorrelación con se respectivo preénfasis y ventaneado, utilizando un coeficiente del filtro de $a=0.95$ y una ventana de Hanning, con tramas de 30 [ms] sin traslape. Con los coeficientes se obtuvo el filtro $H(bz)$ con $b=0.8$. El programa desarrollado es mp.m

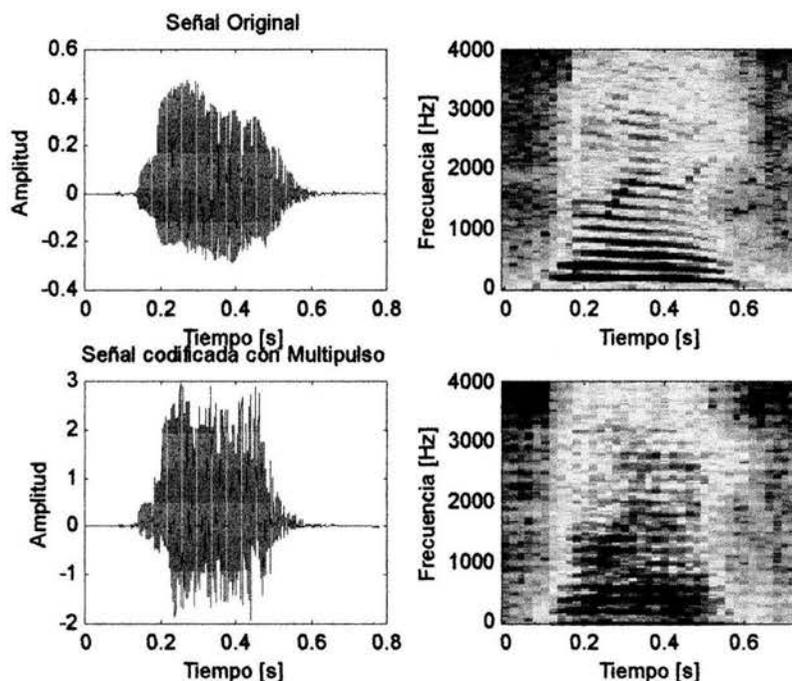


Figura 3.25 Señal original y señal codificada con codificación Multipulso y espectograma.

3.3.7 RPE-LTP

Los parámetros utilizados para esta codificación fueron, para la etapa de pre-procesado primero se utilizo un filtro paso bajas con frecuencia de corte de 3600 [Hz] y posteriormente se aplico un filtro de preénfasis con coeficiente de 0.95, se utilizo una ventana Hamming, la frecuencia de muestreo es de 8000 [hz] y la segmentación de la señal se realizó cada 20 [ms]. Para el análisis STP, el orden del filtro LPC fue de 8 y por último para el análisis LTP se utilizaron subsegmentos de 5[ms].

Por último para esta codificación se desarrolló el algoritmo en el programa rpeltp.m

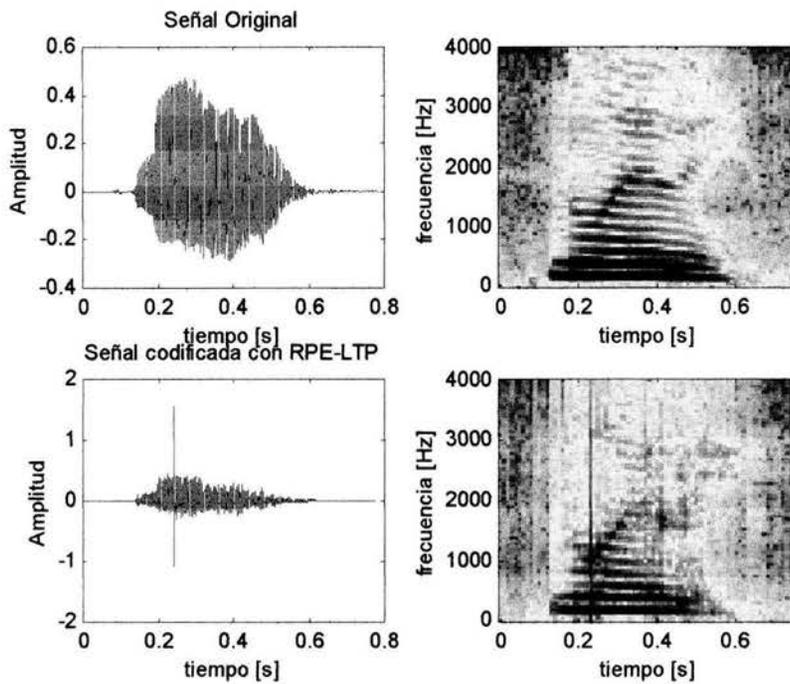


Figura 3.26 Señal original y señal codificada con codificación RPE – LTP y espectrograma.

Las gráficas anteriores muestran la señal original en la parte superior izquierda y en la parte superior derecha su espectrograma, por otro lado en la parte inferior se muestra la señal codificada con su debido espectrograma.

Como era de esperarse, comparando la señal original con la codificada, para DPCM y ADPCM tanto la señal en tiempo como espectralmente son muy parecidas debido al algoritmos que realiza el análisis muestra por muestra.

Con lo que respecta a las otras codificaciones se puede observar, que las señales en el tiempo no son tan parecidas, sin embargo espectralmente no son tan diferentes, esto debido a la naturaleza de sus algoritmos en donde no importa si la señal codificada es igual a la original, sino importa la naturaleza de la señal, en este caso que suene lo más parecido posible.

Capítulo 4

Evaluación de los codificadores seleccionados

4.1 Sistemas de Prueba

La tecnología del habla ha alcanzado un elevado grado de inteligibilidad, que hacen pensar que pronto se podrá obtener niveles propios del habla natural, sin embargo no se ha alcanzado dicha calidad con los *bit rates* necesarios en telefonía celular, por lo que uno de los principales puntos en los codificadores de voz es la calidad de dicha señal sintetizada y poder clasificar la señal en aceptable o no aceptable y al mismo tiempo poder hacer una comparación entre los diferentes tipos de codificadores con lo que respecta a calidad de la señal de voz.

Los procesos que involucran señales audiovisuales se pueden medir con dos tipos de parámetros:

1. Objetivamente
2. Subjetivamente

4.1.1 Medidas Objetivas

Este tipo de medición involucra un análisis matemático y en la práctica se utiliza la medida matemática denominada relación señal a ruido SNR (*Signal to Noise Ratio*), la cual se define como

$$SNR = 10 \log_{10} \left(\frac{\sum_{\text{muestras}} x^2}{\sum_{\text{muestras}} (x - x')^2} \right) \quad (4.1)$$

Otra manera de comparar objetivamente los archivos de voz que resultan de las simulaciones en este trabajo es utilizar la relación señal ruido segmental (SNR_{seg}), que promedia la relación señal ruido sobre segmentos cortos de la forma de onda de la voz y matemáticamente tiene la siguiente expresión

$$SNR_{seg} = \frac{1}{N} \sum_{i=1}^N SNR_i \quad (4.2)$$

donde N es el número de tramas y SNR_i es la relación señal ruido de la trama i .

Con esta relación es posible comparar la semejanza entre el archivo original y el decodificado, de forma que si son muy parecidos, el término del denominador es muy pequeño, por lo que la será muy alta. El valor de esta relación depende también de la energía del archivo de voz origen, por lo que será más alta cuanto mayor sea su energía.

La medición objetiva no siempre es relevante, pero puede ser un buen parámetro a considerar dado que las pruebas subjetivas son complejas y llevan tiempo.

Para realizar las pruebas y obtener una cantidad SNR de los archivos, se desarrolló una función en Matlab de la ec. 4.1 llamada `snr.m` y se muestran los resultados en la siguiente tabla.

| CODIFICACIÓN | SNR |
|--------------|---------------|
| DPCM | 71.9177 [dB] |
| ADPCM | 73.7996 [dB] |
| LPC | -3.4529 [dB] |
| LPC10 | -2.2492 [dB] |
| RELTP | -4.2956 [dB] |
| MULTIPULSO | -32.4196 [dB] |
| RPE – LTP | 7.9621 [dB] |

Tabla 4.1 SNR de todas las codificaciones implementadas.

4.1.2 Medidas subjetivas

En realidad la calidad de un sistema celular se mide con base en un criterio subjetivo de la señal y los dos más importantes pruebas son:

- Calificación promedio de opinión (*MOS: Mean Opinion Score*)
- Tests de rimas (*RT: Rhyme Test*)

Calificación promedio de opinión (*MOS: Mean Opinion Score*)

El concepto principal de este test, es que la calidad de la salida de un codificador – decodificador es juzgada por un considerable número de escuchas, los cuales asignan un valor a la señal de voz

después de comparar la señal original con la salida del sistema codificador – decodificado, este valor va de 1 a 5 con las siguientes características.

| Valor MOS | Característica |
|-----------|----------------|
| 1 | Mala |
| 2 | Pobre |
| 3 | Considerable |
| 4 | Buena |
| 5 | Excelente |

Tabla 4.2 Características de los valores MOS.

Por último, para obtener el valor MOS, se hace un promedio de los valores dados por los escuchas, con este valor se mide la naturalidad o aceptación. Sin embargo este tipo de pruebas tiene algunos inconvenientes, el primero de ellos, es que un procedimiento costoso y consume tiempo.

La señal original debe tener un resultado de 5. Por lo que la mayoría de los sistemas tienen un resultado entre 3 y 4. Un sistema codificador con un resultado arriba de 4 se puede catalogar como uno de muy alta calidad.¹

Test de Rimas

El test fue diseñado por Fairbanks en 1958 y su versión más actual se conoce como Test de Rimas Modificado (*MRT, Modified Rhyme Test*). Se trata de un test formado por estímulos consistentes en palabras monosilábicas con la estructura consonante-vocal-consonante, en el que los escuchas deben elegir una palabra entre seis alternativas. Las palabras difieren en un único segmento, que se encuentra o en posición inicial o en posición final.

La adaptación al castellano ha sido llevada a cabo por Aguilar en 1991, manteniendo las características del test original inglés: estímulos monosilábicos y estructura CVC (aceptando CV o VC en algunos casos). El requisito de la monosilabidad plantea problemas importantes, dado que en ciertos casos ha debido recurrirse a palabras poco familiares o reducirse el número de alternativas ante la imposibilidad de encontrar seis palabras que sólo difieran en la consonante inicial o en la final. Por este motivo, el número de alternativas en la respuesta se ha reducido a 4.

¹ GOLDBERG, Randy. *A Practical Handbook of Speech Coders*. CRC Press LLC. 2000 p. 117-118

En el momento de seleccionar los monosílabos, se ha tenido en cuenta la mayor o menor frecuencia de aparición de la consonante, tendiendo a una aparición proporcional a la que se encuentra en la lengua; aún así, el equilibrio fonético no es un requisito del Test de Rimas.

En la tabla 4.3 y 4.4 se presentan algunos ejemplos de los estímulos del test, para realizarlo, el oyente debe señalar únicamente cuál es la palabra que oye entre todas las de la serie.

Existen además otras herramientas, entre las que citaremos el *Diagnostic Rhyme Test (DRT)* de Voicers (1984), adaptado al castellano por Nadeu (1987) y actualmente en curso de revisión. También se dispone de una versión castellana del *Fast Diagnostic Test (FDT)*, originalmente concebido por Loman y Van Beezoujen (1988), en el que se contempla la inteligibilidad de todas las combinaciones posibles de consonante vocal en palabras de estructura CVC y VCV.²

| | A | B | C | D |
|---|-----|-----|-----|-----|
| 1 | Van | Vas | Bah | Bar |
| 2 | Ved | Ven | Ves | Ver |
| 3 | Dad | Dan | Dar | Das |
| 4 | Sol | Son | Sor | Sos |

Tabla 4.3 Algunos ejemplos de estímulos correspondientes a la adaptación castellana del test de rimas modificado, donde se evalúa la inteligibilidad de la consonante final.

| | A | B | C | D |
|---|-----|-----|------|-----|
| 1 | Sal | Cal | Chal | Tal |
| 2 | Gol | Col | Sol | Rol |
| 3 | Can | Tan | Pan | Dan |
| 4 | Hez | Pez | Tez | Vez |

Tabla 4.4 Algunos ejemplos de estímulos correspondientes a la adaptación castellana del test de rimas modificado, donde se evalúa la inteligibilidad de la consonante inicial.

Los resultados de esta encuesta son reportados como el porcentaje de respuestas correctas con un ajuste de incógnitas. El rango de posibles valores es de 0 a 100 % y es calculado como:

$$DRT = \frac{\text{correctas} - \text{incorrectas}}{\text{total}} \times 100 \quad (4.1)$$

² LLISTERRI, J y POCH, D. Caracterización fonética del bilingüismo. análisis acústico del habla espontánea y evaluación de sistemas de síntesis de habla. Simposio de la Lengua Española. Ciencia y Tecnología. Barcelona. Octubre 1991

4.2 Resultados

Las pruebas fueron aplicadas a 10 personas en el laboratorio de procesamiento digital de voz del edificio Bernardo Quintana de posgrado. Las pruebas fueron aplicada con el ruido de las máquinas propio de un laboratorio.

El formato de las pruebas fue el siguiente:

- **Test de Rimas**

Análisis de codificaciones de voz utilizadas en telefonía celular GSM

Test de Rimas

Test de Rimas

Codificación (una para cada codificación)

Instrucciones:

1. Se escuchará un archivo de voz en formato "wav". Este archivo contiene 8 palabra codificadas, la primera palabra es una de las cuatro posibles opciones (A, B, C o D) del renglón 1 y así sucesivamente.
2. Se marcará con una X la palabra que corresponda con el archivo escuchado.

| | A | B | C | D |
|----------|----------|----------|----------|----------|
| 1 | Van | Vas | Bah | Bar |
| 2 | Ved | Ven | Ves | Ver |
| 3 | Dad | Dan | Dar | Das |
| 4 | Sol | Son | Sor | Sos |
| 5 | Sal | Cal | Chal | Tal |
| 6 | Gol | Col | Sol | Rol |
| 7 | Can | Tan | Pan | Dan |
| 8 | Hez | Pez | Tez | Vez |

• **Test MOS**

Análisis de codificaciones de voz utilizadas en telefonía celular GSM

Prueba MOS

Calificación promedio de opinión (MOS: Mean Opinion Score)

Codificación (una para cada codificación)

Instrucciones:

1. Se escuchará un archivo de voz el cual contiene dos secuencias de voz con la palabra "nueve". La primera de ellas es la señal original y la segunda en la señal codificada.
2. Se comparará dichas señales y se le asignará un valor de acuerdo a que tan similar suena la señal codificada con respecto a la original con base en la siguiente tabla, marcando una X.

| Valor MOS | Característica | MARCA X |
|-----------|----------------|---------|
| 1 | Mala | |
| 2 | Pobre | |
| 3 | Considerable | |
| 4 | Buena | |
| 5 | Excelente | |

Cada formato fue aplicado a todas las codificaciones, obteniendo y presentando para cada codificación los siguientes resultados.

Prueba MOS

| ESCUCHAS | VALOR MOS DPCM | VALOR MOS ADPCM | VALOR MOS LPC | VALOR MOS LPC10 | VALOR MOS RELP | VALOR MOS MULTIPULSO | VALOR MOS RPE – LTP |
|-----------------|----------------|-----------------|---------------|-----------------|----------------|----------------------|---------------------|
| 1 | 4 | 5 | 1 | 1 | 3 | 3 | 4 |
| 2 | 3 | 4 | 3 | 2 | 3 | 2 | 4 |
| 3 | 5 | 4 | 2 | 2 | 3 | 3 | 4 |
| 4 | 4 | 4 | 1 | 2 | 3 | 3 | 3 |
| 5 | 4 | 5 | 2 | 3 | 4 | 4 | 4 |
| 6 | 4 | 4 | 2 | 2 | 3 | 2 | 4 |
| 7 | 4 | 4 | 1 | 1 | 3 | 3 | 4 |
| 8 | 4 | 4 | 2 | 2 | 2 | 3 | 3 |
| 9 | 5 | 5 | 3 | 3 | 4 | 4 | 4 |
| 10 | 4 | 4 | 1 | 2 | 2 | 2 | 3 |
| Promedio | 4.1 | 4.3 | 1.8 | 2 | 3 | 2.9 | 3.7 |

Tabla 4.5 Resultados de la prueba MOS aplicada a los codificadores

En resumen mostrando los resultados MOS de cada codificación

| CODIFICACIÓN | MOS |
|--------------|-----|
| DPCM | 4.1 |
| ADPCM | 4.3 |
| LPC | 1.8 |
| LPC10 | 2 |
| RELP | 3 |
| MULTIPULSO | 2.9 |
| RPE – LTP | 3.7 |

Tabla 4.6 Tabla final de los valores MOS.

Test de Rimas

| ESCUCHAS | DPCM | ADPCM | LPC | LPC10 | REL P | MULTIPULSO | RPE – LTP |
|----------|------------------|-------|-------|--------|-------|------------|-----------|
| | ACIERTOS/ERRORES | A / E | A / E | A / E | A / E | A / E | A / E |
| 1 | 6 / 2 | 7 / 1 | 4 / 4 | 6 / 24 | 6 / 2 | 2 / 6 | 5 / 3 |
| 2 | 8 / 0 | 8 / 0 | 5 / 3 | 4 / 4 | 4 / 4 | 4 / 4 | 6 / 2 |
| 3 | 6 / 2 | 8 / 0 | 3 / 5 | 6 / 2 | 6 / 2 | 3 / 5 | 5 / 3 |
| 4 | 6 / 2 | 7 / 1 | 5 / 3 | 7 / 1 | 6 / 2 | 4 / 4 | 6 / 2 |
| 5 | 8 / 0 | 8 / 0 | 7 / 1 | 5 / 3 | 3 / 5 | 1 / 7 | 7 / 1 |
| 6 | 7 / 1 | 7 / 1 | 5 / 3 | 5 / 3 | 6 / 2 | 3 / 5 | 6 / 2 |
| 7 | 8 / 0 | 8 / 0 | 6 / 2 | 7 / 1 | 4 / 4 | 2 / 6 | 7 / 1 |
| 8 | 6 / 2 | 7 / 1 | 5 / 3 | 4 / 4 | 5 / 3 | 3 / 5 | 6 / 2 |
| 9 | 6 / 2 | 8 / 0 | 5 / 3 | 5 / 3 | 6 / 2 | 2 / 6 | 7 / 1 |
| 10 | 8 / 0 | 8 / 0 | 6 / 2 | 6 / 2 | 6 / 2 | 4 / 4 | 7 / 1 |

Tabla 4.7 Resultados de la prueba MOS aplicada a los codificadores

Por último aplicando la ecuación 4.1 para obtener el porcentaje de inteligibilidad, este porcentaje es mostrado en la siguiente tabla.

| CODIFICACIÓN | % DE INTELIGIBILIDAD |
|--------------|----------------------|
| DPCM | 72.5 |
| ADPCM | 90 |
| LPC | 27.5 |
| LPC10 | 30 |
| REL P | 30 |
| MULTIPULSO | -30 |
| RPE – LTP | 67.5 |

Tabla 4.8 Tabla con el porcentaje de inteligibilidad de cada codificador.

Por último se mostrara una tabla general con los parámetros determinados de la codificaciones.

| CODIFICACIÓN | SNR | MOS | % DE INTELIGIBILIDAD | VELOCIDAD EN bps. (BIT RATE) |
|--------------|---------------|-----|----------------------|------------------------------|
| DPCM | 71.9177 [dB] | 4.1 | 72.5 | 64 [kbps] |
| ADPCM | 73.7996 [dB] | 4.3 | 90 | 64 [kbps] |
| LPC | -3.4529 [dB] | 1.8 | 27.5 | 2 [kbps] |
| LPC10 | -2.2492 [dB] | 2 | 30 | 2 [kbps] |
| RELTP | -4.2956 [dB] | 3 | 30 | 16 [kbps] |
| MULTIPULSO | -32.4196 [dB] | 2.9 | -30 | 3.2 [kbps] |
| RPE – LTP | 7.9621 [dB] | 3.7 | 67.5 | 13 [kbps] |

Tabla 4.9. Parámetros determinados de los codificadores.

Con base en los resultados mostrados en la tabla anterior, si comparamos las codificaciones con la mayoría de los parámetros más altas, las cuales son ADPCM, DPCM y RPE – LTP, podemos observar que los resultados en las pruebas tanto MOS como en el test de rimas son similares o por lo menos están en un rango de calificación de buena calidad, sin embargo el valor en el que difieren mucho es la velocidad en bps (*bit rate*), ya que la utilización de la codificación RPE – LTP reduce en un 80% la velocidad en bps, por lo cual este codificador es utilizado en GSM.

Conclusiones

De acuerdo con el trabajo realizado en esta tesis acerca de la codificación de voz en telefonía celular GSM se puede concluir que con base en lo observado, los sistemas analógicos de comunicaciones móviles presentan muchas limitaciones debido a la creciente demanda de usuarios, que se traduce en una escasez del espectro de frecuencia, servicios limitados y una incompatibilidad entre usuarios móviles, por lo que la solución fue crear un sistema basado en las ventajas de la transmisión digital, con un mejor aprovechamiento en el espectro de frecuencias y una verdadera libertad de movimiento para el usuario, a lo que esto último se traduce en un entendimiento con las redes existentes en las diferentes ciudades y países del mundo, este sistema es llamado *GSM (Global System for Mobile communications)* y hoy por hoy se ha convertido en uno de los sistemas mundiales más importantes y que actualmente se está introduciendo en nuestro país, ganando un importante número de abonados, además de ser la base de nuevas tecnologías tanto para la transmisión de voz como la transmisión de datos.

Por otro lado, debido a que en la actualidad el servicio más importante en una red *GSM* es la comunicación de voz y con la problemática que *GSM* empieza a sufrir con la saturación del espectro de frecuencias, la codificación de voz juega un papel importante y es una de las partes en donde se esté buscando hacer más desarrollo para obtener codificaciones con un menor número de bits a transmitir y obviamente con una buena calidad de voz.

Dentro de los diferentes tipos de codificaciones que se estudiaron se pudo determinar que aun que existen parámetros establecidos que pudieran clasificar la calidad de las codificaciones de voz tal como la relación señal a ruido, los más importantes son las pruebas subjetivas ya que el usuario final es el que determinará la calidad de la señal que escucha, por otro lado la determinación que se realiza al ver si los codificadores de voz son implementados como nuevas tecnologías tiene muchas vertientes, no solo la calidad de voz, obviamente la velocidad en bps (*bit rate*) juega un papel importante debido a que entre más comunicaciones puedas tener en un canal más abonados podrá tener un sistema y eso se ve traducido en dinero, y por otro lado como vimos con *GSM*, aunque existen tecnologías nuevas que ofrecen una calidad mejor al sistema utilizado, este cambio de tecnología no se produce tan fácilmente debido en parte a lo conservador del mercado y a que un cambio también se traduce en inversión, por lo que en conclusión para poder desarrollar un codificador de voz para telefonía celular *GSM* se debe de tener en cuenta: un mejoramiento en la calidad de voz, una velocidad en bps (*bit rate*) menor y una compatibilidad con el sistema de codificación de voz utilizado actualmente.

Con este trabajo fue posible determinar gracias a la simulación la velocidad en bps (*bit rate*) y la calidad de voz utilizando diferentes métodos como la obtención de la relación señal a ruido, su calificación MOS y un test para medir el % de inteligibilidad.

Con estos resultados puede concluir que la tendencia al desarrollo de codificadores de voz seguirá hacia los codificadores híbridos ya que como se puede observar en este trabajo, este tipo de codificadores obtiene una calidad de voz muy similar a los codificadores de onda (en este caso ADPCM con 64 [Kbps], una calidad MOS= 4.3 y inteligibilidad=90%), pero con una reducción del bit rate considerable (13 [Kbps] de RPE – LTP con MOS= 3.7 e inteligibilidad=67.5%). En especial la modificación de parámetros de los codificadores ya establecidos pueden generar cambios importantes en la calidad de voz, tal es el caso de los codificadores utilizados en GSM, como RPE-LTP que fue una modificación de RELP y mejora la calidad de voz.

Por último en este trabajo se plantearon las bases teóricas para el trabajo futuro de la modificación de parámetros en el codificador mas utilizado en GSM, es decir RPE-LTP, para diseñar y desarrollar codificaciones más eficientes tanto en la calidad de voz como el número de parámetros a enviar.

Bibliografía.

1. BEKKERS, Rudi. Mobile Telecommunications: Standards, Regulations and Applications. Artech House Publishers, 1999.
2. MEDINA N., Guadalupe, Telefonía Digital, Ingenierías, Abril – Junio 2001 Vol. IV No.11 p.53
3. HANZO, Lajos. Voice Compression and Communications: Principles and Applications for Fixed and Wireless Channels. USA: Wiley-Interscience, 2001.
4. HERRERA, Abel. Apuntes de Comunicaciones Digitales. FI, UNAM, 2000.
5. OWENS, Frank , Signal Processing of Speech. McGraw Hill, 1987
6. RABINER, L., Fundamentals of Speech Recognition. New Jersey: Prentice Hall, 1993
7. DELLER, PROAKIS & HANSEN. Discrete Time Processing of Speech Signals. Prentice Hall, 1987
8. PANOS, E. Papamichalis. Practical Approaches to Speech Coding. New Jersey: Prentice – Hall, INC, 1987
9. GOLDBERG, Randy. A Practical Handbook of Speech Coders. CRC Press LLC. 2000
10. HERRERA, Abel. Apuntes de procesamiento digital de voz. FI, UNAM, 2000.
11. LLISTERRI, J y POCH, D. Caracterización fonética del bilingüismo. análisis acústico del habla espontánea y evaluación de sistemas de síntesis de habla. Simposio de la Lengua Española. Ciencia y Tecnología. Barcelona. Octubre 1991
12. DARWIN, C.J. Analysis and Synthesis of mixed excitation LPC Coded Speech. IPO Annual Progress Report, 1982 No.17 p. 51-56
13. KONDOZ, A.M. Coding for Low Bit Rate Communications Systems. John Wiley & Sons.
14. KLEIJN, W.B. Speech Coding and Synthesis. Elsevier. 1998
15. ATAL S., Bishnu. Speech and Audio Coding for Wireless and Network Applications. Kluwer Academic Publishers. 1993
16. BARNWELL P., Thomas. Speech Coding: A Computer Laboratory Textbook. John Wiley & Sons, Inc. 1996

-
17. MOULY, Michel. Technical aspects of the GSM system and its derivatives Current Evolution of the GSM Systems. IEEE Personal Communications. October 1995. p 9-19
 18. GERSON, Ira y JASIUK, Mark. Vector Sum Excited Linear Prediction (VSELP) Speech Coding at 8 kbps. IEEE: Acoustics, Speech and Signal Processing. P 461-464
- <http://www.acm.org/crossroads/espanol/xrds7-2/cellular.html>
 - http://ceres.ugr.es/~alumnos/c_avila/gsm0.htm
 - <http://neutron.ing.ucv.ve/revista-e/No2/Gsmseuly.htm>
 - <http://www.yucatan.com.mx/especiales/celular/3g.asp>
 - <http://www.revista.unam.mx/vol.2/num2/art3/>
 - http://agamenon.uniandes.edu.co/~revista/articulos/redes_moviles/rm.html