

UNIVERSIDAD NACIONAL AUTÓNOMA  
DE MÉXICO



FACULTAD DE INGENIERÍA

“TRANSMISIÓN DE AUDIO VÍA INTERNET  
PARA LA IMPLEMENTACIÓN DE ASESORÍAS  
A DISTANCIA A BAJO COSTO”

TESIS

QUE PARA OBTENER EL TÍTULO DE:  
INGENIERO ELÉCTRICO ELECTRÓNICO

PRESENTA:

OSCAR ADRIÁN MENDOZA PABLO  
*acompañado de un libro de computación*  
DIRECTOR DE TESIS:

ING. MARIO A. IBARRA PEREYRA



MÉXICO, D.F. 2002

TESIS CON  
FALLA DE ORIGEN



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.

# PAGINACIÓN DISCONTINUA

ESTA TESIS NO SALE  
DE LA BIBLIOTECA

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

NOMBRE: Oscar Adrián Mendoza Pablo

FECHA: 18-10-02

FIRMA: Mendoza Pablo Oscar A.

ESTE LIBRO  
NO SE SALE

# INDICE

Agradecimientos .....	I
Objetivos .....	VI

## CAPÍTULO 1

### INTRODUCCIÓN

1.1 Antecedentes de la transmisión de voz en la historia del hombre .....	1
1.2 El comienzo de una nueva era: la transmisión de voz vía Internet .....	3

## CAPÍTULO 2

### LA SEÑAL DE VOZ

2.1 Características de la señal de voz .....	8
----------------------------------------------	---

## CAPÍTULO 3

### OPCIONES DE TRANSMISIÓN DE VOZ

3.1 Satelital .....	16
3.2 Videoconferencia .....	22
3.3 Teleconferencia .....	27
3.4 Técnicas de compresión / descompresión por software y hardware vía Internet.	29

## CAPÍTULO 4

### TÉCNICAS DE COMPRESIÓN POR SOFTWARE O POR ALGORITMOS A TRAVÉS DE DSP's

4.1 Compresión de señales de voz .....	35
----------------------------------------	----

## **CAPÍTULO 5**

### **ESTUDIO DEL HARDWARE**

5.1 Requerimientos del equipo de cómputo para implantar el sistema de ..... transmisión de voz.	48
5.2 Tipos de módem .....	49
5.3 Proveedores de Internet .....	51

## **CAPÍTULO 6**

### **SOFTWARE Y HARDWARE PARA LA TRANSMISIÓN DE VOZ**

6.1 Evaluación del software y hardware que hay en el mercado para ..... transmisión de voz.	53
------------------------------------------------------------------------------------------------	----

## **CAPÍTULO 7**

### **DESARROLLO**

7.1 Desarrollo del sistema para impartir asesorías a distancia a bajo costo .....	62
7.2 Resultados Experimentales .....	112

## **CAPÍTULO 8**

<b>CONCLUSIONES</b> .....	<b>120</b>
---------------------------	------------

<b>APÉNDICE A</b> .....	<b>124</b>
-------------------------	------------

<b>BIBLIOGRAFÍA</b> .....	<b>128</b>
---------------------------	------------

## *Agradecimientos*

### *Gracias Señor*

*Porque de alguna u otra forma te has hecho presente en este largo caminar de mi vida como estudiante.*

*Porque a través de todas aquellas pruebas a las que fui sometido, he podido madurar un poco más y de ellas he tomado las lecciones que has querido que yo aprendiera.*

### *Gracias*

*Porque has permitido que tuviera una guía como lo han sido mis padres, hermanos y todos aquellos que me han apoyado.*

### *Gracias*

*Porque has estado en los momentos de gloria en mi vida universitaria y de igual forma en los que he sido doblegado por la adversidad de las situaciones.*

### *Gracias*

*Porque después de varios años, aun sigo vivo para agradecértelo.*

### *Gracias*

*Por escucharme.*

*Y perdona si en algún momento me olvide de ti, si en algún momento perdí la fe en ti.*

*Pero hoy Señor, en este espacio en donde puede culminar esta gran obra que hemos construido todos a lo largo de varios años, sé muy bien que me darás la serenidad y fortaleza necesaria para defender las ideas que he plasmado en este trabajo; y sé de antemano que el lugar que he reservado para ti en esta sala hoy, estará ocupando por ti, para que de esta forma, pase lo que pase, cierre mi promesa ante ti, de no encerrarme sobre cuatro paredes, porque detrás de ellas hay mundo en el cual yo puedo colaborar sin tiempos ni medidas.*

*Oscar Adrián Mendoza Pablo.*

## ***A mis padres***

***Dos personas que me han guiado con sus consejos por las múltiples situaciones a las que me ha sometido la vida.***

***Dos personas que en su momento quizás no supe comprender el valor de sus palabras, pero hoy en mi realidad sé que nunca me mintieron.***

***Dos personas que se preocuparon hasta el último momento en el que fui estudiante.***

***A ellas mi respeto y admiración, a ellas mi agradecimiento...***

***Pero de alguna u otra forma, puedo decirles que han dejado una estela sobre mí que habrán de reflejarse en los valores que me inculcaron desde niño; que habrán de reflejarse en el trato con los demás al conservar la humildad y sencillez, aun cuando el título profesional pudiese ser imponente.***

***Y hoy en este día tan decisivo sólo recordaré, que la sencillez y humildad es una estructura fundamental para el hombre y si se conjuga la fortaleza y dedicación, quizás pueda parecerme un poco a mis padres.***



***A mi director de tesis Ing. Mario A. Ibarra Pereyra***

***Por la paciencia y dedicación que tuvo en el desarrollo de este trabajo.  
Por todos aquellos consejos que me ha dado en este tiempo de trabajo, que gracias a Dios  
a culminado satisfactoriamente, el cuál comenzó como una idea de dos y hoy se observa  
tangible en este pequeño libro, que ha deservir a muchos para alguna investigación.***

### ***A Miriam Ista Solís***

***Por el apoyo que me mostró en la última etapa de mi vida universitaria, cuando las cosas no salieron como yo deseaba.***

***Por sus consejos que si bien escuchaba eran los de mis padres en sus labios, pero gracias a ellos, fueron un impulso determinante para finalizar esta obra que comencé hace algunos años.***

***Al Ing. Víctor Damián Pinilla Morán***

*Por darme su apoyo incondicional en mi estancia dentro de la Facultad de Ingeniería.  
Por haberme permitido demostrarle en alguna época de la vida, que podía realizar con mis convicciones lo que algunos dudaron y gracias a ello tenemos el día de hoy un grato recuerdo que muchos como yo llevamos en el corazón.  
Gracias, porque ahora sé que el que persevera alcanza y nadie nos quitará de la mente que hoy por hoy siempre seremos los Campeones bajo el lema CCCP.*

## OBJETIVOS

Desde la aparición del Internet, se han establecido formas de comunicación entre los usuarios que han permitido acortar distancias y disminuir costos. Esto se debe en gran medida a la estructura y funcionamiento que el Internet como red mundial posee. Las formas de comunicación por Internet se basan en la transferencia de textos en tiempo real, mejor conocidos como *chats*; el correo electrónico y la videoconferencia entre los más usuales, siendo este último uno de los más importantes dentro del ámbito empresarial y académico.

En el aspecto académico veremos que gracias a la videoconferencia, ha sido posible la implantación de cursos, conferencias, carreras y diplomados a distancia; aun cuando los costos de la videoconferencia han sido y siguen siendo elevados. Por tal motivo a través de un análisis y estudio minucioso, surge la idea de realizar un proyecto que de alguna manera resuelva las necesidades de impartir asesorías a distancia en diplomados y carreras como complemento de su formación sin el uso de la videoconferencia.

Por ello el objetivo de esta tesis es el desarrollar, mediante hardware y software, un pequeño sistema de comunicaciones, en el que la voz como parte fundamental de la comunicación, sea transmitida vía Internet con calidad telefónica, en una forma unidireccional o bidireccional y con la optimización de los recursos que se poseen, para conseguir así que la transmisión tenga un bajo costo.

Con este objetivo se conseguirá tener un sistema de comunicaciones flexible, que pueda adaptarse a cualquier entorno de trabajo y que avanzará conjuntamente con la revolución tecnológica, descartando de esta manera que el sistema quede obsoleto a corto plazo.

# CAPÍTULO 1

## INTRODUCCIÓN

---

## INTRODUCCIÓN

### 1.1.- Antecedentes de la transmisión de voz en la historia del hombre

El ser humano, eminentemente social, tiene la necesidad de comunicarse con sus semejantes para desarrollarse dentro de su comunidad; es por eso que desde principios de su existencia ha inventado formas de comunicación como han sido las señas, las pinturas rupestres, los jeroglíficos, los dialectos etc. Sin embargo, los tiempos cambian y en la actualidad, el hombre debe adaptarse a la transformación que sus semejantes le hacen a su entorno; de este modo, uno de los procesos más notables es el de crear formas de comunicación que le permitan establecer vínculos sociales, culturales y comerciales entre todos los grupos humanos y aún con animales "irracionales" y civilizaciones extraterrestres. Por el enorme anhelo (y la necesidad) del hombre por comunicarse a grandes distancias, surge uno de los grandes inventos de la era moderna, al que se le dio el nombre de **Telégrafo**, el cual, al convertir los caracteres alfanuméricos en pulsos eléctricos, abrió la gran puerta de las comunicaciones eléctricas a distancia.

Los sistemas telegráficos se hicieron universales, pero pronto ya no pudieron satisfacer las necesidades de comunicación a distancia. Para subsanar este inconveniente se desarrollaron diferentes métodos para enviar varios mensajes simultáneamente por una misma línea. En la telegrafía dúplex, el primer avance de este tipo, es posible transmitir dos mensajes simultáneos en ambos sentidos por el mismo conductor.

No fue sino hasta después de la II Guerra Mundial cuando el primer enlace comercial por microondas comenzó a operar entre Filadelfia y Nueva York (Estados Unidos) en 1947. Un año más tarde se estableció una red triangular entre Nueva York, la ciudad de Washington y Pittsburgh. El servicio se extendió rápidamente por todo el país gracias sobre todo a la gran capacidad de los sistemas desarrollados. Las microondas son capaces de transmitir, de forma casi instantánea y en grandes cantidades, mensajes de voz, textos, gráficas, fotografías y video. Las ondas de radio que se generan en la emisora se transmiten al punto de destino mediante una serie de reflectores parabólicos, dispuestos

sobre torres muy altas. A fin de prevenir el debilitamiento de la señal debido a la distancia y la curvatura de la Tierra, estas antenas de microondas se colocan a intervalos de unos 48 Km. Este servicio de transmisión por microondas lo implantó en Estados Unidos la Western Union Telegraph Company. En el caso de las comunicaciones intercontinentales se utilizan satélites artificiales geostacionarios que actúan como repetidoras de señales de voz, datos, gráficas y vídeo entre estaciones terrenas.

Otro de los grandes inventos de comunicación que hasta la fecha se sigue modernizando es el **Teléfono**, cuyo desarrollo tecnológico tiene sus orígenes en 1854 cuando el inventor francés Charles Bourseul planteó la posibilidad de utilizar las vibraciones causadas por la voz sobre un disco flexible o diafragma, con el fin de activar y desactivar un circuito eléctrico y producir unas vibraciones similares en un diafragma situado en un lugar remoto, que reproduciría el sonido original. Algunos años más tarde, el físico alemán Johann Philip Reis inventó un instrumento que transmitía notas musicales, pero no era capaz de reproducir la voz humana. En 1877, tras haber descubierto que para transmitir la voz se podía utilizar corriente continua, el inventor estadounidense de origen escocés Alexander Graham Bell construyó el primer teléfono capaz de transmitir y recibir voz humana con aceptable calidad de tono y timbre. La evolución del teléfono ha sido rápida y continua, de modo que hoy en día, se dispone de una enorme cantidad de accesorios y servicios adicionales, pero todos ellos alrededor de su función básica: la comunicación oral entre seres humanos.

La **Radio** ha jugado un papel fundamental en la vida del hombre porque ha sido un medio con el cual, se ha podido establecer comunicación con las masas, logrando ser uno de los más grandes avances tecnológicos de nuestra era, por la razón de que las ondas de radio se utilizan no sólo en la radiodifusión, sino también en la transmisión de mensajes telefónicos, la televisión, el radar, los sistemas de navegación y la comunicación espacial.

En la actualidad, las necesidades de comunicación se pueden resumir en tres palabras: voz, vídeo y datos; pero ahora, se dispone de una gran cantidad de "servicios de valor agregado" que diariamente aumentan y se mejoran, en lo que parece una carrera de obstáculos sin fin, en la que el beneficiado es el usuario, ya que con la competencia y la producción masiva, los costos se reducen o por lo menos no aumentan grandemente.

## 1.2.- El comienzo de una nueva era: la transmisión de voz vía Internet

El desarrollo tecnológico que se ha presentado durante las últimas décadas en el campo de las telecomunicaciones ha sido sorprendente, como para asegurar que la modernización tecnológica ha logrado avanzar vertiginosamente en comparación con años anteriores, siendo uno de los más importantes desarrollos la **digitalización**; es decir, la codificación de las señales en forma binaria mediante la presencia o ausencia de pulsos eléctricos. Siendo este principio el fundamento también de la tecnología de los microprocesadores para dar origen a una estrecha relación de convergencia tecnológica de los sistemas de cómputo y los de telecomunicaciones. Esta convergencia ha permitido que los sistemas de telecomunicaciones posean cierta capacidad de inteligencia, que les permite operar con características superiores de confiabilidad, eficiencia, costos y, sobre todo, con altos índices de flexibilidad.

Gracias a la modernización creciente de las comunicaciones es posible hoy en día hablar del Internet, red de comunicación que entre sus múltiples funciones permite la transmisión de textos, archivos, imágenes, audio y vídeo por un canal telefónico a bajo costo, lo que proporciona la mejor solución de comunicación de audio y vídeo a distancia sin la necesidad de implantar un sistema costoso de videoconferencias.

Con esta perspectiva con la que se visualiza la tecnología, se propondrá dar una solución a la necesidad de impartir asesorías a los diplomados en el extranjero a bajo costo. Así surge esta idea por una necesidad que tienen escuelas dentro de la universidad de dar asesorías como un complemento a la formación del alumnado dentro de los diplomados que se imparten fuera del país.

Es así como el objetivo principal del presente trabajo consiste en desarrollar, mediante hardware y software, un pequeño sistema de comunicaciones, en el que la voz como parte fundamental de la comunicación, sea transmitida vía Internet con calidad telefónica, en una forma unidireccional o bidireccional y con la optimización de los recursos que se tienen para conseguir que la transmisión tenga un bajo costo. En este sentido tendremos la tarea de realizar una digitalización de la señal de voz acompañada de una codificación o decodificación de dicha señal, la cual es comprimida de la manera más adecuada, lo que nos permitirá que ocupe el menor espacio en memoria para poder lograr que viaje de



manera rápida por el Internet. Cabe aclarar que el desarrollo de un software es imprescindible. Hay que tener en cuenta que los equipos de cómputo deberán tener características muy específicas para poder comprimir y descomprimir la voz en tiempos mínimos en los que el oído humano pueda escucharla con claridad.

Para dar una idea más concreta de lo que se pretende realizar véase el esquema de la figura 1.1

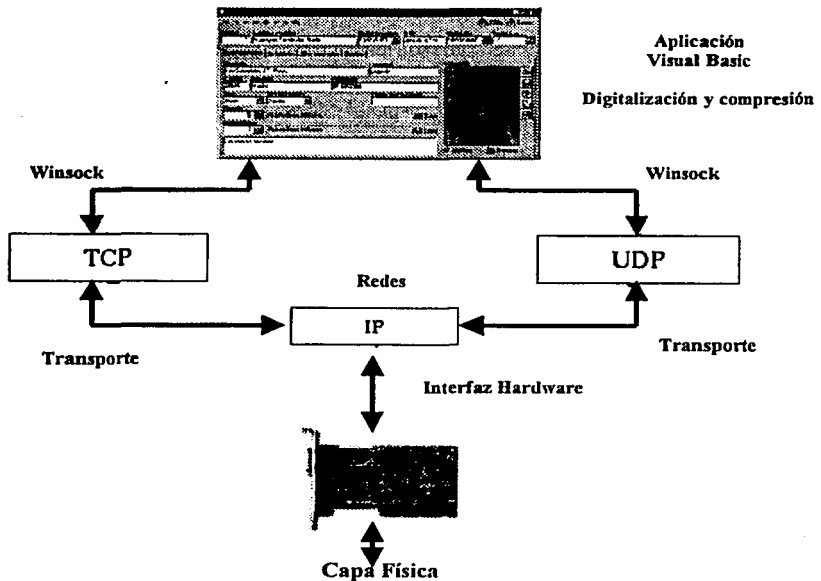


Figura 1.1 Diagrama general de la transmisión

Como podemos observar, la señal de voz es procesada por el software, que la digitaliza, codifica y comprime, para enviarla por la vía del Internet; ésta viaja hasta llegar a su destino, para encontrarse nuevamente con otra interfaz de usuario, que le permite descomprimir la secuencia de voz que viene de la red. Este software está diseñado para

trabajar con los protocolos estándar TCP y UDP y mediante la programación de Winsocks es posible comunicar procesos sobre la red mundial (Internet); es decir, se puede comunicar a dos programas sobre el Internet. Para que tengamos una calidad de voz próxima a la telefónica, se tienen que analizar de igual forma los equipos de hardware, ya que con las prestaciones de la tecnología en la actualidad, es posible tener, con una configuración adecuada de los dispositivos una calidad superior a la de un teléfono convencional, aunque normalmente lograremos una calidad bastante similar a la de un teléfono.

Habrá que considerar que un factor que influye en la calidad de la voz es el ancho de banda. Algunos proveedores de Internet tienen anchos de banda muy bajos y conseguir una buena conexión a través de ellos es difícil. En general con un ancho de banda de un módem de 14,400 bps podemos conseguir una calidad más que aceptable.

Otra forma de poder transmitir voz y audio por Internet, es con la intervención de un DSP, el cuál recurre a un algoritmo de síntesis de voz, para la digitalización de la señal, la cuál posteriormente es enviada por el protocolo de comunicación TCP/IP.

Esta opción que se ha dado al último es válida pero no práctica, ya que no cumpliría con el objetivo de ser una transmisión a bajo costo, porque el costo que tiene un DSP no es pequeño, además de que su implementación en la realidad no es posible ya que cada alumno tendría que adquirir su propio DSP para poder hacer conexión con el instructor y si hubiera algún problema no se podría dar el soporte técnico necesario, por lo que se vuelve un sistema complejo de implementar.

Hay que recordar que el presente trabajo trata de dar una solución a un problema real, de tal forma se pretende implementarlo y adaptarlo en la realidad para que trabaje de manera eficaz.

De esta manera hemos de presentar en el capítulo dos, un esquema general de las características que describen el comportamiento de la señal de voz desde el punto de vista anatómico, fundamental para realizar un análisis en la digitalización de la señal de voz. Así mismo se describen los pasos que hay que seguir para proceder a efectuar una digitalización de voz, seguido de una breve explicación de los esquemas básicos de codificación de voz. Por su parte en el capítulo tres se realiza un análisis de las opciones que se tienen hoy en día para la transmisión de voz como son la vía satelital, videoconferencia, teleconferencia y técnicas de compresión por software y hardware. Dicho análisis tiene como propósito

presentar la mejor opción de transmisión que cumpla con los objetivos establecidos en el proyecto desde sus inicios, por eso se ha investigado a profundidad cada una de las opciones antes mencionadas para llegar a una conclusión que es determinante para el desarrollo de este proyecto.

Como podremos observar, las técnicas de compresión por software y hardware son eficaces, pero siempre habrá más de una forma en las que se puedan aplicar, es decir, que una de las posibilidades en la que podemos transmitir voz es a través de software ayudados por un equipo de cómputo que cumpla con algunos requerimientos establecidos, pero también podemos hacerlo con un procesador digital de señales conocido como DSP. Por lo cual el capítulo cuatro esta dedicado a hablar del proceso de digitalización y compresión de voz por un DSP.

En el capítulo cinco, expondremos a detalle los requerimientos que son indispensables para implantar un sistema de transmisión de voz, lo que implica hablar de los requisitos que deben cumplir las computadoras, los tipos de módems que pueden ser utilizados cuando se realiza un acceso telefónico a redes y por último las consideraciones que se deben tomar al contratar el proveedor de Internet.

La transmisión de voz vía Internet no es un tema nuevo en nuestros días. En los últimos años se han desarrollado softwares que son capaces de transmitir voz, pero han tenido errores que los han dejado inutilizables. Por eso, en el capítulo seis, se evalúan los softwares más comunes en el mercado que transmiten voz, ya que se toma este parámetro de evaluación como referencia para tratar de hacer un programa que elimine en lo más posible las deficiencias que han tenido los demás a través del tiempo. Por otro lado se hacen recomendaciones indispensables a la hora de adquirir hardware, que ayudan a la transmisión de voz, como son los altavoces, tarjetas de sonido y micrófonos.

Finalmente, en el capítulo siete, se expone la solución que hemos tomado para la transmisión de la voz, fundamentada dicha solución en las investigaciones realizadas en los capítulos anteriores. Se han tenido que sortear algunos escollos, tales como entender el funcionamiento del streaming audio y de la programación de sockets, que son punto vital en el enlace de aplicaciones a través del Internet.

De la misma forma se dan a conocer los resultados obtenidos en el proyecto después de su desarrollo, ya que a partir de algunas pruebas realizadas, se ha podido verificar que sí se establece la comunicación y sí tiene calidad telefónica.

# CAPÍTULO 2

## LA SEÑAL DE VOZ

## CAPÍTULO 2

### La señal de voz

El propósito que tiene el presente capítulo, es el de presentar de forma general los aspectos que deben tomarse en cuenta para la transmisión de voz a través de un canal de banda angosta.

#### 2.1 Características de la señal de voz

La voz humana se produce voluntariamente por medio del aparato fonatorio. Éste está formado por los pulmones como fuente de energía en la forma de un flujo de aire, la laringe, que contiene las cuerdas vocales, la faringe, las cavidades oral (o bucal) y nasal y una serie de elementos articulatorios: los labios, los dientes, el alvéolo, el paladar, el velo del paladar y la lengua (figura 2.1).

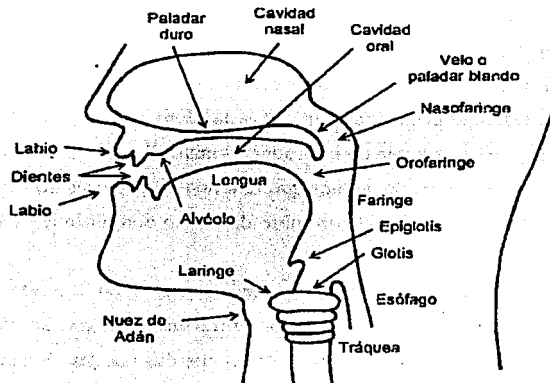


Figura 2.1 Corte esquemático del aparato fonatorio humano

Las cuerdas vocales son, en realidad, dos membranas dentro de la laringe, orientadas de adelante hacia atrás (figura 2.2).

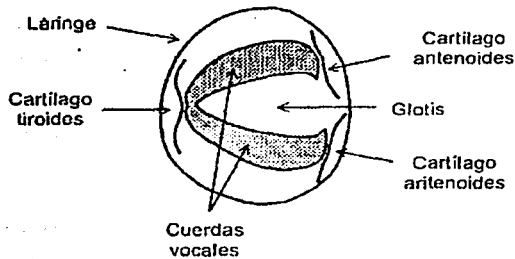


Figura 2.2 Corte esquemático de la laringe en un plano horizontal

Por adelante, las cuerdas vocales se unen en el cartilago tiroides; por detrás, cada una está sujeta a uno de los dos cartilagos aritenoides, los cuales pueden separarse voluntariamente por medio de músculos. La abertura entre ambas cuerdas se denomina glotis.

Cuando las cuerdas vocales se encuentran separadas, la glotis adopta una forma triangular.

El aire pasa libremente y prácticamente no se produce sonido. Es el caso de la respiración.

Cuando la glotis comienza a cerrarse, el aire que la atraviesa proveniente de los pulmones experimenta una turbulencia, emitiéndose un ruido de origen aerodinámico conocido como aspiración.

Al cerrarse la glotis más, las cuerdas vocales comienzan a vibrar a modo de lengüetas, produciéndose un sonido tonal, es decir periódico. La frecuencia de este sonido depende de varios factores, entre otros del tamaño y la masa de las cuerdas vocales, de la tensión que se les aplique y de la velocidad del flujo del aire proveniente de los pulmones. A mayor tamaño, menor frecuencia de vibración, lo cual explica por qué en los varones, cuya glotis es en promedio mayor que la de las mujeres, la voz es en general más grave. A mayor

tensión la frecuencia aumenta, siendo los sonidos más agudos. Así, para lograr emitir sonidos en el registro extremo de la voz es necesario un mayor esfuerzo vocal.

La frecuencia de los sonidos emitidos por el ser humano, aumenta (a igualdad de las otras condiciones) al crecer la velocidad del flujo de aire, razón por la cual al aumentar la intensidad de emisión se tiende a elevar espontáneamente el tono de voz.

La porción que incluye las cavidades faríngeas, oral y nasal junto con los elementos articulatorios se denomina genéricamente cavidad supraglótica, en tanto que los espacios por debajo de la laringe, es decir la tráquea, los bronquios y los pulmones, se denominan cavidades infraglóticas.

Varios de los elementos de la cavidad supraglótica se controlan a voluntad, permitiendo modificar dentro de márgenes muy amplios los sonidos producidos por las cuerdas vocales o agregar partes distintivas a los mismos, e inclusive producir sonidos propios. Todo esto se efectúa por dos mecanismos principales: el filtrado y la articulación.

El filtrado actúa modificando el espectro del sonido; tiene lugar en la cavidad supraglótica: la faringe, la cavidad nasal, la cavidad oral y la parte labial. Estos órganos constituyen resonadores acústicos que enfatizan determinadas bandas de frecuencia del espectro generado por las cuerdas vocales, conduciendo al concepto de formantes, es decir, una serie de picos de resonancia ubicados en frecuencias o bandas de frecuencia, que son bastante específicas para cada tipo de sonido.

La articulación es una modificación de los sonidos, principalmente a nivel temporal y está directamente relacionada con la emisión de los mismos y con los fenómenos transitorios que los acompañan. Está caracterizada por el lugar del tracto vocal en el que tiene lugar, por los elementos que intervienen y por el modo en que se produce.

Existen varias maneras para analizar y describir el habla. Los enfoques más comúnmente usados son:

1. **Articulación:** Análisis de cómo el humano produce los sonidos del habla.
2. **Acústica:** Análisis de la señal de voz como una secuencia de sonidos.
3. **Percepción Auditiva:** Análisis de cómo el humano procesa el habla.



## La señal de voz

El habla es una señal analógica, es decir un flujo continuo de ondas sonoras y silencios. Es por eso que el conocimiento de la ciencia de la acústica, se utiliza para identificar y describir los atributos del habla que son necesarios para un análisis efectivo de la señal de voz.

Las cuatro características fundamentales en el análisis acústico son:

- Frecuencia
- Amplitud
- Estructura armónica
- Resonancia

La amplitud de una onda sonora fisiológicamente provoca el movimiento del tímpano del oído. La frecuencia de la onda será el número de vibraciones por segundo; para tonos altos habrá mayor frecuencia; por el contrario, para tonos bajos hay menor frecuencia.

La mayoría de los sonidos incluyendo el habla, tienen una frecuencia dominante a la que se le denomina frecuencia fundamental. Ésta se percibe como el pitch (tono) combinado con frecuencias secundarias. En el habla, la frecuencia fundamental es la velocidad a la que vibran las cuerdas vocales al producir un fonema sonoro. Sumadas a la frecuencia fundamental hay otras frecuencias que contribuyen al timbre del sonido (son las que nos permiten distinguir una trompeta de un violín, etc. o las voces de diferentes personas). Algunas bandas de las frecuencias secundarias juegan un rol importante en la distinción de un fonema de otro; a esto se le conoce como formantes y son producidas por la resonancia. Por lo cual la resonancia se define como la habilidad que tiene un objeto de vibrar a la frecuencia de oscilación de una fuente sonora.

De acuerdo a las características antes descritas y como se había mencionado con anterioridad, la señal de voz se representa en forma continua emitiendo sonidos que consisten en variaciones en la presión del aire a través del tiempo y a frecuencias que podemos escuchar. Una de las formas en que podemos representar estos sonidos es a través de una onda, como la que se muestra en la figura 2.3.

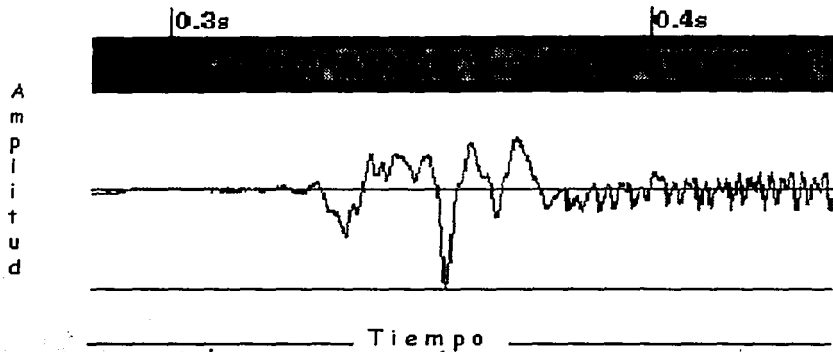


Figura 2.3 Representación gráfica de una señal de voz

Una de las grandes ventajas de éste tipo de gráficas es que no ocupa mucho espacio en memoria. Una desventaja es que no se describe explícitamente el contenido de la señal en términos de sus propiedades. Los espectrogramas contienen mayor información sobre los datos de la voz; son una transformación que muestra la distribución de los componentes de frecuencia de la señal. Un ejemplo se presente en la figura 2.4.

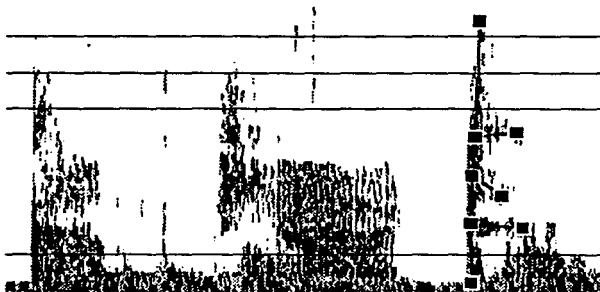


Figura 2.4 Espectro de una señal de voz

Las partes más oscuras, vistas en la figura, representan la concentración de energía y son denominadas formantes.

Por otra parte, es importante mencionar que la capacidad auditiva del ser humano se concentra en un rango de frecuencias que va de 20Hz a 20,000Hz. Los sonidos emitidos al hablar se encuentran entre 100Hz y 4,000Hz. Aunque la mayoría de la información de la voz humana se concentra en el rango de 100 a 3400 Hz.

### Codificación de voz

En la década de los 90 se ha experimentado un gran desarrollo de las comunicaciones digitales, debido a la implementación de las tecnologías de compresión de voz.

La compresión de la voz consiste en tomar una señal de voz (a través de un micrófono, por ejemplo) y una vez convertida a señal digital, realizar un procesamiento sobre ella, con objeto de reducir el número de bits a transmitir.

Toda señal analógica puede ser convertida a digital de acuerdo a los siguientes puntos:

- Hacer un muestreo de la señal a una velocidad de por lo menos dos veces el valor de la más alta frecuencia contenida en tal señal.
- Conversión de cada muestra de la señal analógica a un código binario. Este proceso recibe el nombre conversión analógico-digital.
- El "bit rate" de una señal se calcula de acuerdo a la ecuación:

$$\text{bit rate} = Q \left[ \frac{\text{muestras}}{\text{seg}} \right] \times R \left[ \frac{\text{bits}}{\text{muestra}} \right]$$

Desgraciadamente, el número resultante es a menudo más grande de lo que nos gustaría. Para voz en telefonía, la frecuencia de muestreo es 8000 muestras/seg y el número de bits por muestra es 8, produciendo un "bit rate" de 64 Kbps. Por sí solo, este "bit rate" no es problemático. Pero tengamos en cuenta que se trata de transmitir miles de conversaciones simultáneas por un mismo canal, lo cual representa un grave problema.

4 DE FEBRERO  
1992 CON

Algunos de los esquemas básicos para la codificación de la voz son:

- Modulación de Código de Pulso (PCM): ésta es la técnica normal (para la cual se ha seguido el desarrollo anterior) y no produce reducción del "bit rate", 64 Kbs.
- Modulación Delta (DM) y la Modulación Delta Adaptativa (ADM): Estos algoritmos tienen en cuenta la correlación existente entre muestras adyacentes de una señal de voz. En el caso de DM, sólo se transmite un bit por muestra, representando si la muestra actual es más grande o más pequeña en amplitud que la muestra anterior. En este sistema, un parámetro del plan básico es el tamaño del paso, que se refiere a la frecuencia de muestreo que se utiliza en la modulación. Los sistemas de DM presentan dos posibles errores: (1) "el overload" de la pendiente cuando a la entrada los cambios se producen más rápidamente que la cantidad del paso y (2) "el ruido granular", cuando la señal de entrada es constante, pero los convertidores DM producen una salida que oscila positiva y negativamente en torno al valor de 0.
- Modulación Delta Adaptativa. (ADM): En esta variante, el tamaño del paso se hace inconstante para que las dos condiciones se moderen. Se pueden conseguir "bit rates" del orden de 10 a 40 kbps.
- Códigos de Predicción Lineal (LPC): Se basan en una representación del tracto vocal humano como un filtro variante en el tiempo, excitado por energía de características particulares. En el siguiente apartado se desarrollará más detenidamente.
- Cuantización vectorial (VQ): Trata individualmente cada muestra. Los "bit rates" obtenidos están en el rango de 10 Kbps.
- Codificación de Huffman: Se basa en asignar códigos de distinta longitud de bits a cada uno de los caracteres de un archivo. En el capítulo 7 se analizará esta codificación de una manera más detallada.

Los codificadores de voz se dividen normalmente en dos grandes clases:

- "Waveform coders" ó codificadores de forma de onda: El objetivo es reproducir en el decodificador la señal de voz original muestra a muestra, partiendo de una

muestra inicial. Por ejemplo, un codificador PCM simple logra esto cuantizando cada muestra de la señal de voz a uno de un número fijo de niveles. Si tenemos en cuenta que cada muestra se codifica con 8 bits, lo cual exige que la señal analógica que tiene un número infinito de posibles valores sea representada con solo  $2^8 = 256$  valores, ya sea por redondeo o por truncamiento. Esto es lo que se llama "cuantificación". El resultado final de este proceso de conversión A/D es el ya conocido bit rate de 64 Kbps. Si se reduce el número de niveles de cuantización se puede reducir el bit rate a expensas de aumentar el ruido de cuantificación y la calidad de la voz será menor que la obtenida en el teléfono.

- **Codificadores basados en un modelo de voz:** Usan un modelo paramétrico para aproximar los segmentos cortos de voz (10-40 ms). Para cada segmento se estima un juego de parámetros y se convierte la señal analógica en una cadena de bits.

El estudio realizado de la señal de voz, es fundamental para el desarrollo del programa dentro del proyecto, ya que a partir de las características de ésta, se pueden tener diversas opciones para su digitalización y compresión. La manera en que se digitalizará la señal, depende en gran medida de las herramientas y flexibilidad de las que disponga el lenguaje de programación.

# CAPÍTULO 3

## **OPCIONES DE TRANSMISIÓN DE VOZ**

## CAPÍTULO 3

### **Análisis de las opciones de transmisión de voz**

En el presente capítulo se analizarán las diferentes opciones que existen para la transmisión de voz, siendo esencial dicho estudio para poder tener un panorama amplio que nos permita visualizar y consolidar un proyecto que cubra los objetivos que se han planteado, tomando para su implantación dos aspectos fundamentales que son:

- Una llamada por Internet con la calidad y características de la telefonía analógica actual.
- La optimización de los recursos que se tienen para conseguir que la transmisión tenga un bajo costo.

Analizando los objetivos que se han planteado y los sistemas de comunicación que pueden satisfacer las expectativas del proyecto, se han encontrado cuatro soluciones de transmisión de voz (en algunos de los casos cubren las expectativas de los objetivos planteados en forma parcial):

- Vía satélite
- Videoconferencia (conexión vía ISDN)
- Teleconferencia
- Técnicas de compresión/ descompresión de voz por software y hardware utilizando el Internet como canal de comunicación.

#### **3.1 Vía satélite**

En esta opción se utiliza la ventaja del ancho de banda que nos ofrecen las compañías de comunicaciones, además de la velocidad de transmisión, tomando en cuenta que el costo de inversión para el desarrollo del proyecto es independiente de la distancia entre los puntos a interconectar.

Al realizar la investigación correspondiente, se buscó una empresa local que tuviera soporte técnico en esta rama de transmisión en México, la cual es:

- DirecPc Internet Directo, S A de C.V ubicada en Blvd Manuel Ávila Camacho 88, 6º piso, Col. Lomas de Chapultepec (figura 3.1)

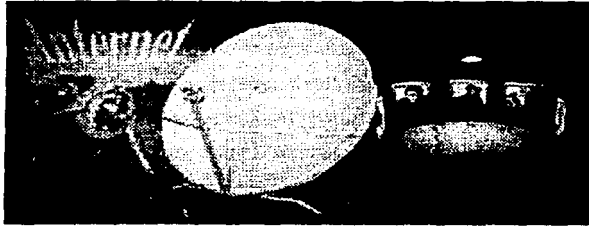


Figura 3.1 DirecPc Internet Directo

DirecPc es una solución ideal para aplicaciones educativas y de negocios que requieren gran ancho de banda, cobertura continental, precio accesible, rápida instalación y utilización. Cualquier tipo de información como: video en tiempo real, audio y paquetes de software, puede ser enviada directamente a una PC o a un servidor de red usando la tecnología satelital de DirecPc.

Dentro de las características de DirecPc se encuentran:

- Funciona utilizando la tecnología vía satélite a nivel mundial. El usuario con DirecPC solicita la información a través de su línea telefónica normal o su enlace dedicado de baja velocidad (64 kbps o 128 Kbps) pero recibe todos los servicios vía satélite a gran velocidad (400Kbps, 1.2Mbps ó 3Mbps).
- Es un sistema de transmisión vía satélite en un sólo sentido el cual cuenta con una tarjeta PCI y una antena de recepción de 66, 75, 90 o 120 centímetros de diámetro.
- Todas las solicitudes de información hechas por su línea telefónica o enlace dedicado de baja velocidad son enviadas al NOC (Centro de operaciones de la Red) y retransmitidas vía satélite.
- Utiliza un transpondedor de 12Mbps del satélite Satmex V con capacidad para distribuir información a múltiples puntos simultáneamente en casi todo el



continente Americano. La información viaja encriptada a través de algoritmo, lo que asegura que solo el usuario autorizado entienda la información que le ha sido enviada.

Cada uno de los servicios de DirecPC disponibles para los usuarios permite diseñar aplicaciones únicas. Los servicios que ofrece esta empresa son tres en esencia *Distribución satelital multipunto, Turbo Internet y Multimedia interactiva*.

#### Distribución Satelital Multipunto

Este servicio permite la distribución de datos Punto-Multipunto a gran velocidad de forma confiable, segura y simultánea desde nuestro NOC (Centro de Operaciones de la Red) hacia cualquier número de ubicaciones o localidades dentro del continente Americano. El envío de datos Punto-Multipunto es a 3 Mbps.

La distribución satelital multipunto permite a los proveedores o centros de información distribuir paquetería de software, información de mercado, documentos electrónicos, así como grandes archivos bajo tiempos y fechas predeterminadas o bajo demanda hacia las localidades seleccionadas.

#### Beneficios de Distribución Satelital Multipunto

- Ahorros significativos en costos de operación.
- Distribución de software en forma simultánea e inmediata.
- Confirmación de recepción de la información enviada.
- El emisor selecciona quien recibe la información.

#### Turbo Internet

El servicio de Turbo Internet permite la operación de todas las aplicaciones más comunes de Internet incluyendo FTP, GOPHER y HTTP.

Este servicio está disponible en versión de red, que opera entregando la información individualmente solicitada por cada uno de los usuarios de la red en forma independiente. El acceso a Internet para la solicitud de información, se logra mediante una llamada a través de la línea telefónica convencional o enlace dedicado al proveedor de servicio de Internet. En ese momento, el usuario realiza su consulta y el sistema DirecPC, envía la solicitud de información y la dirección IP a la red de Internet para obtener los datos solicitados por el usuario. La recepción de la información solicitada es entregada por Internet al Centro de Operación de la Red (NOC) de INTERDirec y éste, en forma automática e inmediata la transmite a través del satélite Satmex V hasta la PC del usuario que solicitó dicha información (véase figura 3.2). Esta es la manera en que DirecPC le brinda a los usuarios personales o de negocios en forma individual o en red, el servicio de entrega de información solicitada a Internet con gran eficiencia y a gran velocidad.

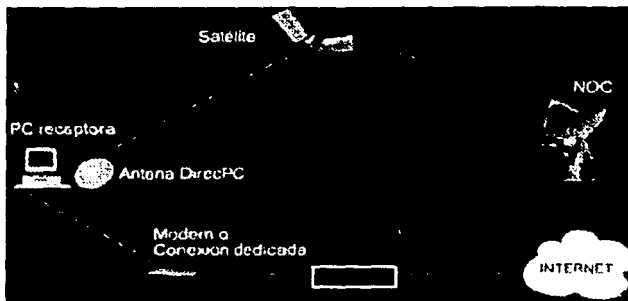


Figura 3.2 Diagrama en el que se muestra la transmisión de información por una conexión Turbo Internet de DirecPC

#### Beneficios del servicio de Turbo Internet

- Reduce drásticamente el tiempo empleado en la navegación y búsqueda de información.

- Descarga inmediata de software (Down Loads).
- Reduce el tiempo de conexión a Internet.
- La recepción de información por Internet a una mayor velocidad, incrementa la productividad del usuario particular o de los usuarios de red.
- El servicio es totalmente administrable; en cualquier momento puede cambiar el modo de recepción satelital a terrestre y viceversa.

### Multimedia Interactiva



Figura 3.3 Multimedia Interactiva

Es la solución más efectiva para el aprendizaje interactivo a distancia. En este servicio se combinan las soluciones más avanzadas, rápidas, confiables y fáciles de usar para el aprendizaje a distancia desarrolladas por One Touch Systems junto con la tecnología de DirecPC (figura 3.3).

### Ventajas de este servicio

- Cuenta con un sistema que integra herramientas muy poderosas para el instructor, el alumno y para la administración del aprendizaje.
- Optimiza los recursos de la empresa al no tener que desplazar al personal.
- Los instructores manejan la clase de una forma fácil, rápida y efectiva.

- Se cuenta con sistemas que permiten el registro, medición y control de cada sesión y estudiante de forma inmediata.
- Facilidad de tomar los cursos en cualquier lugar; por ejemplo el trabajo o en la casa.

El hardware utilizado por la empresa para cualquiera de los servicios que ofrece, consiste en una antena parabólica de 90 cm. a 120 cm., módem USB y una tarjeta PCI para el enlace bidireccional con la antena. Todo el equipo tiene un costo aproximado de \$ 1300 USD.

Las características del hardware para poder implantar este sistema se dan a conocer a continuación (figura 3.3).

#### *Hardware requerido:*

##### PC mínimo:

- \* Procesador Pentium
- \* 16 MBytes RAM
- \* 20 MBytes de espacio libre en el disco duro para el Software
- \* Módem compatible con Hayes (V.32, mínimo 9600 bps)
- \* Monitor VGA color
- \* Conexión con un proveedor de servicio de Internet

##### PC recomendado:

- \* Procesador Pentium a 90Mhz o más rápido
- \* 32 MBytes RAM ó más
- \* 20 MBytes de espacio libre en su disco duro o más
- \* Módem compatible con Hayes (V.32, 28800 bps ó más rápido)
- \* Monitor SVGA color
- \* Conexión con un proveedor de servicio de Internet



Figura 3.3  
Requerimientos de  
Hardware

**Servidor de Red:**

- Windows NT 4.0 con un mínimo de 32 Mbytes en RAM
- NetWare versión 3.12 con un mínimo de 16Mbytes en RAM.
- NetWare versión 4.10, 4.11 o IntranetWare con un mínimo de 24Mbytes en RAM
- Topologías soportadas: Ethernet y Token Ring
- Módem para el servidor (si la conexión al proveedor de servicio de Internet es vía módem).

Platicando con la empresa sobre los objetivos que se tienen estipulados en el proyecto, la cotización proporcionada por ésta asciende a \$20, 000 UDS., más \$350 UDS. por hora de transmisión. Este presupuesto incluye los aparatos necesarios de comunicaciones, equipo de cómputo especializado para dar el servicio, la renta y las horas de transmisión que se tienen programadas para las asesorías.

**3.2 Videoconferencia**

Otros de los medios de comunicación que se tienen para la transmisión de voz es la videoconferencia, capaz de lograr transmitir no sólo voz, sino también video e imágenes de gran calidad (figura 3.4).



Figura 3.4 La Videoconferencia, medio de transmisión ideal para dar asesorías a distancia

Esta tecnología, tiene cobertura más limitada, pues requiere de un mayor ancho de banda para transmitir voz e imágenes; así mismo, el equipo de comunicación en las sedes

receptoras debe cumplir con ciertos estándares para que pueda interconectarse con el de la sede emisora; sin embargo, la interactividad profesor-alumno es total ya que simultáneamente los alumnos y profesor se pueden ver y escuchar.

La videoconferencia interactiva es un recurso didáctico, por medio del cual dos o más puntos distantes establecen comunicación con capacidades de transmisión y recepción de audio y video en forma bidireccional. Esto hace que en un medio educativo el profesor y los alumnos, desde todos los sitios, se vean unos a otros y establezcan una comunicación interactiva y simultánea.

### Ventajas de la videoconferencia

Dada su flexibilidad, la videoconferencia interactiva, posee grandes ventajas sobre el resto de los medios para la educación a distancia, como son:

- La interacción es más realista, productiva y eficaz.
- Representa el modelo más cercano a la operación óptima del aula ya que permite el empleo de técnicas grupales.
- Es un medio de gran valor para capacitación y educación continua, con múltiples aplicaciones en los diferentes campos tales como telemedicina. Permite también el intercambio de datos a través de paquetes computacionales, envío de imágenes, videos, y prácticamente cualquier tipo de información audiovisual.

La información que a continuación se da a conocer fue proporcionada por DGSCA a través de sus departamentos de "conectividad y enlaces" y "clases a distancia"; en ésta se describen todo los aspectos técnicos que se requieren para llevar a cabo una videoconferencia.

Los tipos de enlace se definen en función de las necesidades de transmisión con un ancho de banda aplicable a una videoconferencia, en la que se dispone de una capacidad para enlaces desde 128 kbps hasta 384 kbps; por este motivo las conexiones de enlace que cubren las necesidades de nuestro proyecto son la conexión vía ISDN y la conexión enlace dedicado.

### Conexión ISDN (Integrated Services Digital Network)

Para este tipo de enlace se necesita tecnología tipo BRI, que consisten en dos canales de comunicación, uno de 64 kbps y otro de 16 kbps para la señalización. El problema de este tipo de videoconferencia radica en que la conexión ISDN no se encuentra disponible en todas las áreas geográficas.

En general estos sistemas ofrecen buena calidad de vídeo, debido a que el ancho de banda es mayor al de la línea telefónica. Si se requiere realizar enlaces al extranjero, es necesario que los demás países cuenten con el servicio ISDN, además de que los equipos que se utilicen en los enlaces deben contar con los mismos protocolos de comunicación o tener protocolos estándar para el caso de los equipos de distintas marcas.

### Conexión enlaces dedicados

Para este tipo de conexión, es necesario contratar un enlace dedicado entre el lugar donde está el alumno y el sitio en donde se imparte la videoconferencia, además de que los equipos deberán tener los accesorios necesarios para trabajar con este tipo de conexión que por lo general son tarjetas de comunicación que cumplen con la norma V.35.

### Elementos que integran un Sistema de Videoconferencia

#### *CODEC*

Es el elemento principal del equipo de videoconferencias y consiste básicamente en una computadora que se encarga de controlar todo el proceso de comunicaciones entre las aulas participantes, las cuales definiremos como aula local y aulas remotas. Además, tiene entradas para recibir la señal de los micrófonos, cámaras de vídeo y demás periféricos ubicados en las aulas, que mandan señales de un aula a otra y dado que cada equipo de videoconferencias tiene un CODEC, la comunicación se realiza en ambos sentidos.

#### *Tablero de Control*

Es el que nos permite operar las cámaras, tanto locales como remotas, establecer posiciones predeterminadas, controlar el volumen de audio, controlar algunos otros periféricos y ordenar la ejecución de comandos del CODEC.

### *Cámara Robótica*

Es la videocámara incluida en cualquier equipo; ésta es manejada a través del tablero de control.

### *Micrófono*

Capta el audio que se envía al otro sitio.

### *Monitores*

El sistema de videoconferencias es interactivo y cuenta con dos monitores principales; en uno de ellos se presenta la imagen del aula local y en el otro, la del aula remota. Estas imágenes pueden provenir de las cámaras de video, cámara de documentos, computadora o una máquina reproductora de video.

### *Herramientas y equipos auxiliares del sistema de videoconferencia.*

Internet: antes, durante o después de una sesión por videoconferencia, permite la comunicación permanente entre los participantes, ya que Internet es una red que contiene miles de redes de computadoras conectadas entre sí para intercambiar información.

Fax: es de gran utilidad para enviar y recibir material impreso antes, durante y después del encuentro por videoconferencia. Por otro lado, ayuda a establecer el contacto entre los sitios en caso de presentarse una falla en el equipo de videoconferencia.

Teléfono: a través de él se puede brindar asesoría técnica de un sitio a otro, así como informar de cambios o ajustes de último momento.

### *Funciones básicas que realiza el equipo de Videoconferencias*

- Establecer la comunicación a otro sitio: la comunicación se establece hacia la unidad multipunto.
- Control de audio: regular, en el sitio local, el nivel de volumen que se transmite a los demás sitios.



- Captura de imágenes: almacenar en memoria gráficos, dibujos, tomas de cámara, así como enviar y recibir de otros equipos de videoconferencias todo tipo de documentos guardados previamente.
- Selección y control de cámaras: cuando se trabaja con dos o más cámaras, mediante el equipo de videoconferencias se puede elegir la cámara cuya señal queremos transmitir. El equipo también puede controlar la cámara robótica para que ésta se mueva a posiciones preestablecidas por el usuario.

### *Funcionamiento de la Videoconferencia*

Las señales proporcionadas por las cámaras, el micrófono y equipos periféricos son enviadas al CODEC, dentro del cual se realiza un proceso complejo, que se realiza en tres etapas:

#### *Primera Etapa*

El CODEC convierte las señales de audio y vídeo a un código binario. A esto se le conoce como digitalizar. La información es entonces comprimida y multiplexada, eliminando los bits redundantes; de esta forma, los datos que se transmiten requieren menos ancho de banda en el canal de comunicación.

#### *Segunda Etapa*

Los datos son enviados a otro dispositivo de comunicación, el cual los transmite al sitio remoto por un canal de transmisión: cable coaxial, fibra óptica, microondas o satélite.

#### *Tercera Etapa*

A través del canal, el otro sitio recibe los datos por medio del dispositivo de comunicación, el cual lo entrega al CODEC que se encarga de desmultiplexar y decodificar a señales de audio y vídeo, las que envía a los monitores para que sean vistas y escuchadas por las personas que asisten al evento.

### 3.3 Teleconferencia

Mediante esta tecnología la cobertura territorial está en función de la "huella" del satélite que se esté utilizando, exponiendo el profesor el contenido de su curso en una sede emisora, transmitiéndose su clase a las sedes receptoras en las que los alumnos a través de monitores de televisión ven y escuchan al profesor y la interacción profesor-alumno se realizan a través de teléfono, fax o internet.

Una empresa que brinda este servicio en México es Edusat, Se describirá a continuación como se distribuye la señal que permite realizar una teleconferencia, además se dará a conocer cuales son los aditamentos con los que se debe contar para desarrollar una transmisión de este tipo.

#### Distribución de la señal Edusat

- I. La Dirección General de Televisión Educativa y el Instituto Latinoamericano de la Comunicación Educativa generan la señal de los canales de EDUSAT y la manda a través de microondas al centro de control operativo, ubicado en Contel Iztapalapa.
- II. La señal en Contel es comprimida y enviada al satélite Satmex V en forma digital.
- III. El satélite se encarga de distribuirla a todas las instituciones con receptor digital (figura 3.5).

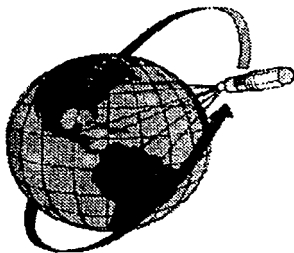


Figura 3.5 Distribución de la señal Edusat

El equipo necesario para implementar este sistema de comunicación consta de un receptor digital y una antena parabólica, además de ciertos accesorios como son el control remoto del receptor digital y el sistema de orientación de la antena.

Equipo de recepción DCI, Edusat

Receptor Digital

Marca General Instrument, modelo 310-D con capacidad para recibir hasta 8 canales de video y 16 canales de audio, utilizando la norma de compresión digital MPEG 1; cuenta además con un control remoto (figura 3.6).

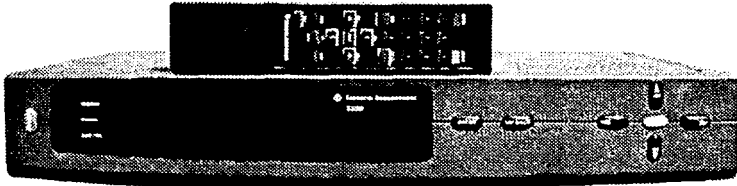


Figura 3.6 Receptor digital Edusat

Antena Parabólica. Está representada en la figura 3.7.

Diámetro mínimo: 1.8 metros

Montura: Polar

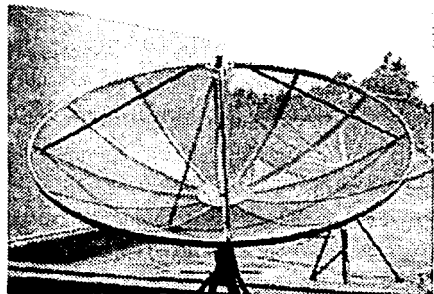


Figura 3.7 Antena parabólica

## Diagramas de conexión

En la figura 3.8 se muestra el diagrama de conexión de una televisión al receptor digital.

Conexión de una Televisión al Decodificador

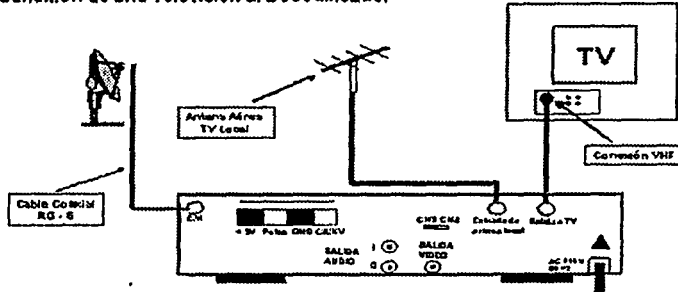


Figura 3.8 Diagrama de conexión de una televisión al decodificador

### 3.4 Técnicas de compresión/ descompresión por software y hardware vía Internet

En la actualidad existen tanto programas como tarjetas de comunicación que nos permiten llevar a cabo transmisiones de audio y video de una manera adecuada, para poder establecer pequeños sistemas de videoconferencia sin la necesidad de pagar costosos equipos.

El auge de estos pequeños sistemas, se debe en gran medida a las sofisticadas técnicas de compresión y descompresión que hay en audio y video, las cuáles permiten transmitir a gran velocidad, usando anchos de banda estrechos como los de los cables telefónicos.

Debido a los avances en la compresión y descompresión del sonido, además de la amplia gama de dispositivos de comunicaciones en la PC, el audio proveniente de los programas telefónicos en Internet puede rivalizar con el sonido de la telefonía normal, ya que el algoritmo de compresión utilizado no ocasiona pérdidas de información notables debido a los recortes de ancho de banda

Para que una conversación en Internet suene lo más parecido a una plática telefónica, el programa de compresión debe ser capaz de minimizar el tiempo que implica comprimir y

descomprimir el sonido. Existen varios algoritmos de compresión que difieren en la rapidez y la eficiencia con las que ejecutan sus procesos.

Un programa de compresión que produzca sonidos de alta calidad, requiere de una computadora con alta frecuencia de reloj; por el lado contrario, si se tiene una máquina con reloj de baja frecuencia, nos entregará un sonido de baja calidad.

En el esquema de la figura 3.9 se muestran los requerimientos mínimos de hardware para la instalación de un software de telefonía por Internet.

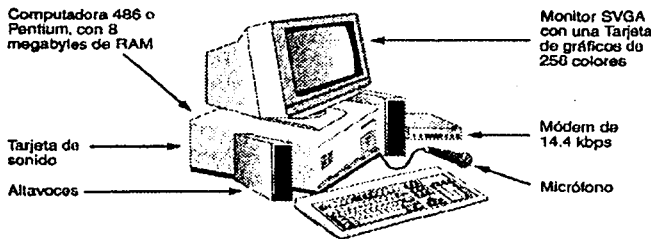


Figura 3.9 Requerimientos de hardware

Hoy en día existen alrededor de diez esquemas diferentes de compresión, conocidos comúnmente como codecs (abreviatura de compresión / descompresión), los cuales están incorporados al software existente para teléfono en Internet. Algunos codecs, que básicamente son ecuaciones matemáticas, funcionan mejor en las conexiones a velocidad baja; pero otros se dirigen a restituir las pérdidas de información o los vacíos sin sonido, mediante una característica llamada interpolación; sin embargo, existen otros que realmente se especializan en conexiones de velocidad alta, por ejemplo a 56k e ISDN.

En la tabla 3.1 se proporciona información sobre los codecs más utilizados y se hace referencia a lo que se puede esperar de ellos en cuanto a calidad de sonido, además de que se mencionan los requerimientos mínimos de hardware para cada caso.

Compresión	Bytes por segundo	CPU necesaria	Calidad de sonido
Simple	4000	486/33 mHz	Pobre
ADPCM	4000	486/33 mHz	Buena
Simple + ADPCM	2000	486/33 mHz	Mala
GSM	1650	486/66 mHz (aunque una Pentium a 75 mHz o superior funciona mejor)	Buena
Simple + GSM	825	Pentium a 75 mHz	No tan buena
ADPCM + GSM	800	486/33 mHz	Buena
LPC	650	Pentium a 90 mHz	Buena, pero depende de la calidad de la conexión
LPC-10	346	Pentium a 133 mHz	No mala

Tabla 3.1 Codecs

#### Información técnica de los codecs

- ADPCM reduce la velocidad de los datos por un factor de dos; de hecho reduce 8,000 bytes por segundo a 4,000 bytes por segundo. ADPCM requiere una computadora menos sofisticada que GSM para codificar y decodificar el sonido. Esto significa que las computadoras 486 de bajo rango y posiblemente algunas máquinas 386DX pueden usar el esquema de compresión ADPCM, siempre y cuando la conexión a Internet sea relativamente rápida.
- GSM reduce la velocidad de los datos por un factor de casi cinco; es decir, comprime 8,000 bytes por segundo a 1,650 bytes por segundo; sin embargo, la codificación es un proceso muy complicado que requiere una computadora bastante rápida, en el rango de 486 o Pentium.
- ADPCM + GSM reduce la velocidad de transferencia de datos a cerca de 800 bytes por segundo, pero requiere una computadora 486.
- LPC reduce la velocidad de transferencia de datos por un factor de 12, mediante una expresión matemática complicada; esto significa que se necesitará un coprocesador matemático de punto flotante. LPC puede ser muy sensible a los cortes que resultan de un elevado nivel de entrada de sonido (por colocar muy alto el volumen del micrófono o al hablar demasiado cerca de éste). Asimismo, LPC es muy sensible al

sonido de frecuencias altas, de manera que quienes tengan voz aguda pensarán que no les sirve.

Es así como el software y hardware aunado a las técnicas de compresión hacen posible la transmisión de sonido de una manera versátil para muchas personas que no pueden pagar grandes cantidades de dinero por equipos de videoconferencia o uso de satélite. La transmisión esta basada en el esquema de la figura 3.10.

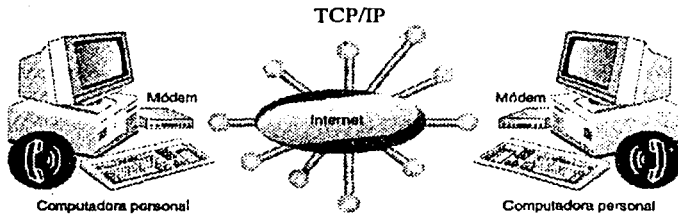


Figura 3.10 Diagrama de transmisión de la señal de voz vía Internet

Aquí podemos realizar un seguimiento en el que la voz como primera fase es grabada, después es convertida en señales digitales, las cuales son comprimidas por los codecs en pequeños paquetes de datos, que son finalmente enviados a través del Internet con la ayuda del protocolo de comunicación TCP/IP. Esta forma de comunicación se reduce a estar hablando de forma habitual por el teléfono normal (figura 3.11).



Figura 3.11 Por las técnicas de compresión y descompresión por software y hardware, podemos tener una comunicación semejante a la de la telefonía habitual.

**Análisis final**

Al realizar un análisis sobre las diferentes opciones que hay para transmisión de voz, se pueden manifestar las siguientes observaciones finales basadas en los objetivos que persigue el proyecto con la implantación de asesorías a distancia a bajo costo:

En el caso vía satélite, que se ha presentado, podemos decir que los costos de los equipos que se utilizarían, además de la renta por el uso del satélite en la transmisión son elevados, por lo cual se descarta esta opción, ya que la infraestructura que se necesita es demasiada y por lo tanto no cumple con nuestros objetivos.

El desarrollo del proyecto con la videoconferencia ya sea por conexión ISDN o línea dedicada, no es posible, por los motivos de que el país o los países en donde se requieran dar las asesorías deben contar con estos servicios, además de que los sistemas de videoconferencia deben cumplir con los protocolos estándares de comunicación para que se pueda efectuar el enlace; es decir, deben manejar el mismo protocolo de comunicación, como puede ser el H.320, el H.323 o el H.324. El aspecto económico aquí es importante, ya que se tienen que tomar en cuenta diversos factores como son los equipos y sus aditamentos que son muchos y variados; además, la sala en la que se lleve a cabo la transmisión debe poseer características específicas, por lo que no cumple con el objetivo de ser económico.

La teleconferencia es quizá un medio de transmisión económico ya que se equipara con los equipos de Sky o Direct TV; sin embargo, no es posible tener una interactividad en tiempo real de ambos lados de la transmisión, por lo cual no cumple con el objetivo de ser una comunicación bidireccional.

Finalmente, si se utilizan técnicas de compresión y descompresión por medio de software y hardware, se cumplen todos nuestros objetivos, ya que por medio de algoritmos podemos lograr una calidad de voz equiparable a la telefónica, con lo cual los recursos que se invierten para montar este sistema son relativamente bajos si los comparamos con los casos anteriores; así mismo, se consigue la optimización de los recursos que hay para lograr una transmisión a bajo costo y cumple con ser una comunicación bidireccional.



Es por eso que utilizar el Internet como medio de transmisión es la mejor opción junto con el sistema que se pretende desarrollar, para implantar asesorías a distancia para diplomados, maestrías y doctorados y en general, para cualquier tipo de cursos impartidos a distancia.

# CAPÍTULO 4

**TÉCNICAS DE  
COMPRESIÓN POR  
SOFTWARE O POR  
ALGORITMOS A TRAVÉS  
DE DSP's**

## CAPÍTULO 4

### Técnicas de compresión por algoritmos a través de DSP's

En el capítulo 3 punto 3.4, se hizo referencia a las técnicas de compresión por software y hardware utilizando como herramientas de trabajo: un programa o "paquete" cuya función es la de grabar, digitalizar, comprimir y enviar la voz por medio del Internet, con ayuda de un equipo de cómputo:

En el presente capítulo se observará que es posible también transmitir la voz utilizando un procesador digital de señales (DSP), en el cual se implementa un algoritmo de compresión y descompresión de voz, basado en el codec GSM.

#### 4.1. Compresión de señales de voz

Algoritmos más importantes en la codificación de voz

##### *El modelo LPC de voz*

Los códigos de Predicción Lineal (LPC) son una de las técnicas de análisis de voz más poderosas y uno de los métodos más útiles para obtener una buena calidad de voz a un "bit rate" bajo. Estos códigos proporcionan una buena estimación de los parámetros de la señal de voz, y son de fácil implementación en la computadora.

El método LPC parte del postulado de que la señal de voz es producida por un zumbador al final de un tubo. La glotis (el espacio entre las cuerdas vocales) produce el zumbido que es caracterizado por su intensidad y frecuencia (pitch). El tracto vocal (la garganta y boca) forma el "tubo" que queda caracterizado por sus resonancias que se llaman formantes.

LPC realiza el análisis de la señal de voz efectuando una estimación de los formantes, eliminando sus efectos de la señal de voz y estimando la intensidad y frecuencia del zumbido restante. El proceso de quitar el efecto de los formantes recibe el nombre de filtrado inverso, y la señal restante se llama residuo.

LPC sintetiza el segmento de voz invirtiendo el proceso: usa el residuo para crear una señal fuente, utiliza los formantes para crear el filtro (qué representa el tubo), y pasa la señal fuente a través del filtro, produciendo una señal de voz (figura 4.1). Puesto que la señal de voz varía con el tiempo, es decir, la señal de voz no es una señal estacionaria, sino cuasi-estacionaria, este proceso se hace en segmentos cortos de voz que reciben el nombre de tramas o "frames".

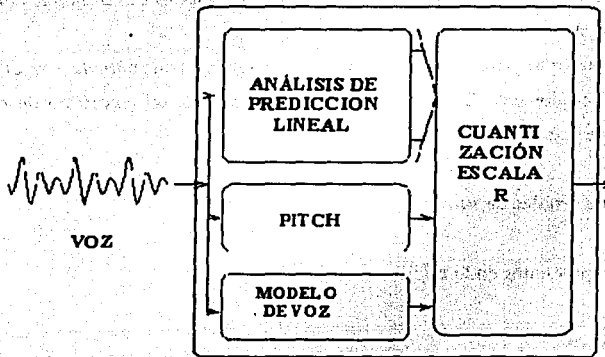


Figura 4.1 Diagrama del modelo LPC de voz

#### *El algoritmo multipulso (MPE, Multi-Pulse Excitation)*

El modelo del tracto vocal usado en predicción lineal (LPC) obtiene muy buenos resultados en la codificación de segmentos de voz. El problema principal con el predictor lineal es la excitación para el modelo del tracto vocal. El codificador divide el segmento de voz en sonidos graves y agudos. Esto tiene el efecto de producir un mismo rendimiento sonoro sintético. La excitación multipulso intenta rectificar este problema.

Cuando el segmento de voz se pasa a través del predictor lineal, el pitch o frecuencia ficticia de la señal, introduce un término de correlación larga en el segmento de voz, dando

como resultando una señal cuasi-periódica. Esta periodicidad no está reflejada por el predictor lineal y produce picos grandes en el residuo o señal de error. Este término de correlación larga puede ser eliminado pasando el residuo a través de un segundo predictor lineal. Este segundo predictor, es el predictor del pitch o predictor de tiempo largo, y está diseñado para no quitar la correlación de muestras adyacentes sino para quitar la correlación de los períodos adyacentes del residuo. Esto se logra insertando un retraso que corresponde al pitch o periodo en el predictor. En el codificador de multipulso la entrada excita la cascada de los dos predictores lineales con una serie de pulsos. Generalmente de cuatro a seis son los pulsos usados como excitación. La posición y amplitud de estos pulsos son determinadas usando un procedimiento de análisis-por-síntesis (es decir el segmento de voz es sintetizado para cada conjunto de pulsos y el conjunto que produce el error más bajo entre la señal reconstruida y la entrada es el conjunto escogido). Las situaciones de los pulsos son determinadas secuencialmente (es decir la búsqueda se lleva a cabo para encontrar la situación óptima para el primer pulso; entonces, la búsqueda se repite para encontrar el segundo pulso, y así sucesivamente). El codificador de multipulso posee una gran efectividad para "bit rates" del orden de 9.6 Kbps y menores, produciendo una alta calidad de voz.

#### *El estándar europeo GSM 06.10*

El método empleado por ISDN, y por los sistemas telefónicos actuales para el multiplexado de las líneas, del cual ya se ha hecho mención, se conoce como PCM. El bit rate es de 64 kbps; demasiado alto para ser usado en la red. La señal PCM contiene mucha redundancia. Para el proceso de compresión, la información de muestras anteriores, que no cambian muy rápidamente, se usa para predecir la muestra actual. Los coeficientes de la combinación lineal de las muestras anteriores, más una forma codificada del residuo, representan a la señal de voz, la cual se divide en segmentos de 20 ms., cada uno de los cuales se codifica con 260 bits, dando un bit rate total de 13 kbps. Éste método recibe el nombre de "Full-Rate speech coding", que se conoce como GSM 06.10. Por otra parte, un grupo de operadores de Norte América ha desarrollado el estándar GSM1900, que proporciona mejor calidad de voz utilizando el bit rate anterior de 13 kbps.

*El estándar americano VSELP*

El estándar VSELP está basado en la técnica CELP ("Coded Excited Linear Prediction"). Utiliza dos tipos de correlación: de tiempo corto y de tiempo largo. Debido a esta inclusión de la correlación de tiempo largo, el modelo CELP elimina correlaciones debidas explícitamente al pitch.

*El estándar GSM 06.10 para la implementación del algoritmo de compresión y descompresión de voz utilizando un DSP***Breve historia sobre GSM**

Las comunicaciones móviles por radio se usan desde hace 75 años. Los métodos utilizados en telefonía móvil se conocen desde hace 50 años. La telefonía móvil celular (analógica) apareció en el ámbito comercial en 1983.

En este tiempo, los sistemas de telefonía celular analógica estaban experimentando un rápido crecimiento en Europa; particularmente en Escandinavia y el Reino Unido, pero también en Francia y Alemania. Cada país desarrolló su propio sistema que era independiente de todos los demás en equipo y funcionamiento. Ésta era una situación indeseable, porque siendo Europa un continente relativamente pequeño, el enorme tráfico telefónico "internacional" no era posible por la incompatibilidad de los sistemas celulares. Debido a esto, la "Conference of European Posts and Telegraphs (CEPT)" formó un grupo de estudio llamado "Groupe Special Mobile (GSM)" para estudiar y desarrollar una tecnología propiamente europea y de carácter público para el sistema de telefonía móvil. El sistema propuesto debía seguir ciertos criterios:

- Alto nivel de calidad
- Bajo coste de mantenimiento
- Difusión internacional
- Posibilidad de implementación en equipos portátiles
- Máximo aprovechamiento espectral
- Compatibilidad con ISDN

En 1989, se transfirió responsabilidad de GSM al Instituto de Estándares de Telecomunicaciones Europeo (ETSI), y en 1990 se publicó la fase I de las especificaciones de GSM. El servicio comercial empezó a mediados de 1991, y a lo largo de 1993 había 36 redes GSM en 22 países.

### Servicios que proporcionan GSM

Usando las definiciones de ITU-T, los servicios de la telecomunicación pueden ser clasificados en tres rubros: servicios de portadora, teleservicios, y servicios suplementarios. El teleservicio básico apoyado por GSM es la telefonía. La señal de voz es codificada digitalmente y se transmite a través de la red GSM como una cadena digital de bits. Los usuarios GSM pueden enviar y recibir datos, a bit rates de 9600 bps.

### *El algoritmo de codificación GSM: GSM 06.10, RPE-LTP*

El estándar GSM 06.10 se compone de dos partes: un algoritmo de codificación y un algoritmo de decodificación. Para su análisis, estos procesos se dividen en módulos que a su vez se dividen en submódulos (véase figura 4.2).

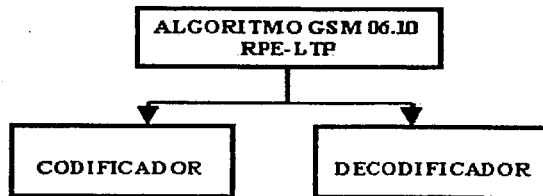


Figura 4.2 Algoritmo de codificación GSM 06.10

### Módulo codificador

El módulo codificador (figura 4.3), es el encargado de realizar la compresión de la señal analógica de voz. El algoritmo se implementa sobre un DSP, encargado de realizar la conversión A/D.

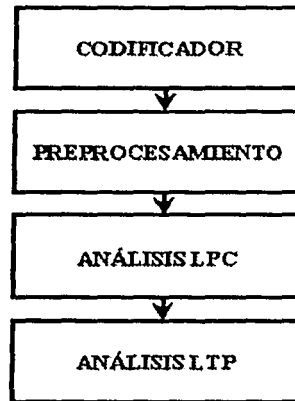


Figura 4.3 Módulo codificador

#### Submódulo de preprocesamiento

Tiene como función realizar un doble filtrado a la señal de entrada. Por un lado, debe eliminar el offset de la señal de entrada, mediante un filtro de offset y posteriormente, debe enfatizar la señal mediante un filtro de énfasis (figura 4.4).

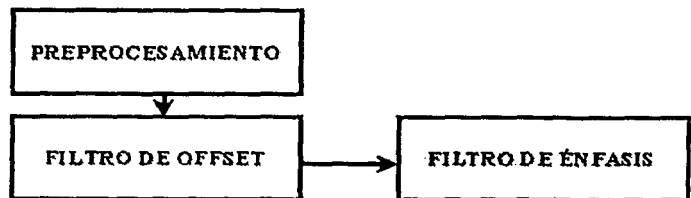


Figura 4.4 Submódulo de preprocesamiento



---

**Submódulo de análisis LPC**

Es un módulo de análisis de predicción lineal de tiempo corto, las partes que lo conforman son las siguientes (figura 4.5):

- **Bloque de segmentación:** La señal de entrada se divide en tramas de 20 ms. Como la frecuencia de muestreo es de 8 kHz, se recogen tramas de 160 muestras que no están solapadas. Sobre estas tramas se realiza un análisis LPC de orden  $p=8$ .
- **Autocorrelación:** Se calculan  $p+1$  valores de la función de autocorrelación para cada trama (en este caso, 9 valores).
- **Recursión de Schur:** Se usa este algoritmo para el cálculo de los coeficientes de reflexión. Estos coeficientes son los coeficientes del filtro FIR de tiempo corto. Reciben este nombre debido a la representación del tracto vocal como una serie de tubos cilíndricos, como ya se explicó en apartados anteriores.
- **Transformación de los coeficientes de reflexión a coeficientes "Log Area Ratio", LAR:** Los coeficientes de reflexión, se encuentran en el rango de  $-1 \leq r_p(n) \leq 1$ . Por tanto es conveniente para su procesamiento transformarlos a una escala logarítmica.
- **Cuantización y codificación de los coeficientes LAR:** Los coeficientes LAR, poseen rangos dinámicos distintos, debido a la función de densidad de distribución. Por esto, se hace necesario limitar los valores de estos coeficientes. El resultado de este bloque, son los parámetros LPC que se deben enviar al decodificador.
- **Decodificación de coeficientes LAR:** Para el filtrado de la señal es necesaria la utilización de los coeficientes de reflexión. Por tanto, hay que decodificar los coef. LAR que fueron codificados en el bloque anterior.
- **Interpolación de los coeficientes LAR:** Antes de pasar a la transformación a coef. de reflexión, hay que tener en cuenta que los coeficientes de filtro que han sido codificados y posteriormente decodificados, pueden presentar cambios muy abruptos debido a este proceso de codificación y decodificación. Por esta razón se hace necesaria una interpolación de estos

coeficientes para evitar transitorios espurios en la señal de voz reconstruida en el decodificador.

- Transformación a coeficientes de reflexión: Tras la interpolación de los coeficientes LAR, se procede al proceso inverso a la transformación a "Log Area Ratios", que es la transformación a coeficientes de reflexión.
- Filtrado de tiempo corto: Una vez obtenidos los coeficientes de reflexión se pasa la trama a través de un filtro FIR que es el análisis real del modelo LPC.

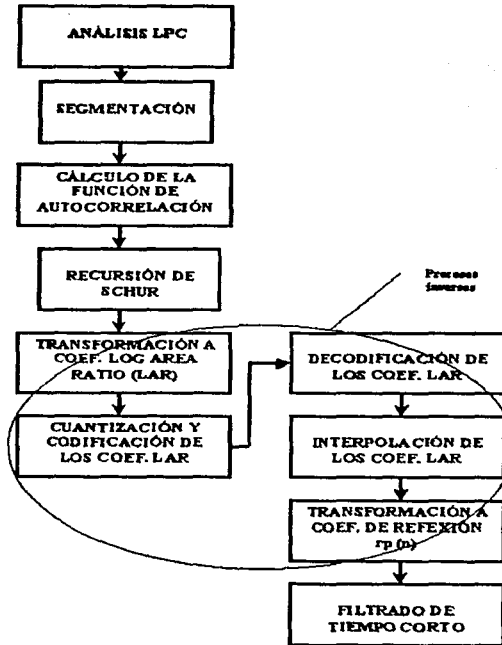


Figura 4.5 Submódulo análisis LPC

Una vez realizado el análisis de tiempo corto, se procede a realizar un análisis de tiempo largo, o análisis LTP, cuya principal misión es la de eliminar los picos producidos en la señal residuo a causa de la aparición del pitch o periodo ficticio de la señal de voz. El esquema a seguir es el que se presenta en la figura 4.6.

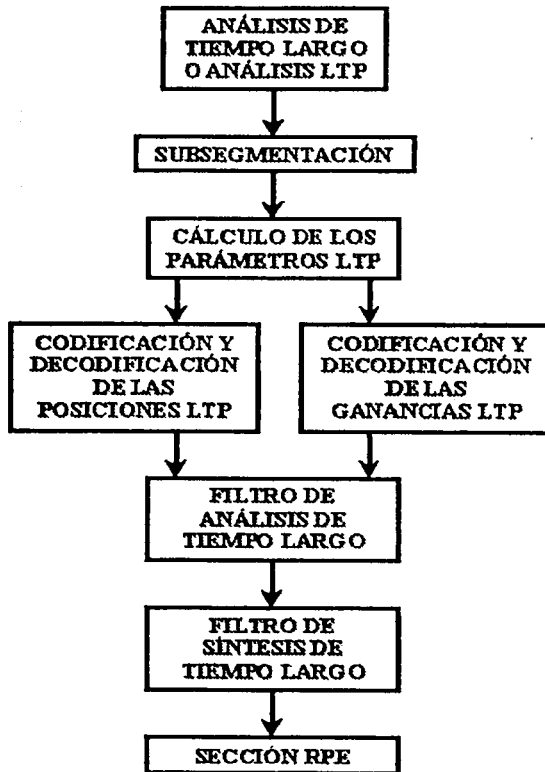


Figura 4.6 Análisis tiempo largo o análisis LTP

- Subsegmentación: Cada trama de 20 ms., (160 muestras) se divide en cuatro tramas de 5 ms., 40 muestras cada una.
- Cálculo de los parámetros LTP: Se compone de varios pasos de acuerdo con el siguiente orden:
  - Cálculo de la función de correlación cruzada de la señal residuo del subsegmento actual y muestras anteriores de la señal residuo reconstruida.
  - Búsqueda del pico de la función correlación cruzada para este intervalo.
  - Evaluación del factor de ganancia.
  - Codificación/decodificación de las posiciones LTP: Cada posición se codifica con 7 bits.
  - Codificación/decodificación de las ganancias LTP: Para cada subbloque, la ganancia calculada, se codifica con 2 bits.
  - Filtro de análisis de tiempo largo: A la señal de residuo procedente del filtrado LPC (señal residuo de tiempo corto), se le resta una estimación de esta señal, obteniéndose una señal residuo de tiempo largo.
  - Filtro de síntesis de tiempo largo: La señal residuo de tiempo largo, es reconstruida por bloques de 40 muestras. La estimación del residuo de tiempo largo se suma a esta señal reconstruida, dando como resultado la señal de residuo de tiempo corto reconstruida.
  - Sección RPE. Lleva a cabo el análisis multipulso.

#### Módulo decodificador

El módulo decodificador es el encargado de realizar la descompresión o decodificación de los datos recibidos del codificador. El decodificador recibe los datos, y realiza sobre ellos una síntesis LTP y a la salida de ésta, una síntesis LPC. Tras dicha síntesis, se realiza un postprocesamiento sobre los datos de salida. La mayoría de los bloques y funciones utilizados en el codificador, también son utilizados en el decodificador (véase figura 4.7).

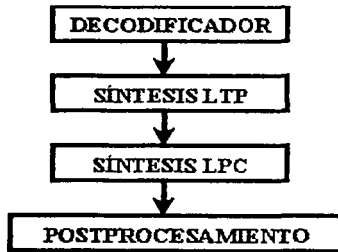


Figura 4.7 Módulo decodificador

La síntesis LTP, que es el proceso inverso al análisis LTP realizado en el codificador. Véase el esquema en la figura 4.8.

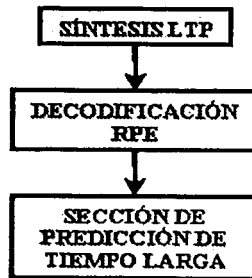


Figura 4.8 Síntesis LTP

- **Decodificación RPE:** Se deben realizar los siguientes pasos:
  1. Decodificación de las muestras RPE, (en el módulo codificador equivale a la cuantización APCM inversa).

2. Posicionamiento de la rejilla RPE. Es una interpolación de las muestras RPE en un factor 3. También fue realizado en el codificador.
- Sección de predicción de tiempo largo: La señal obtenida del submódulo anterior, es pasada a través del filtro de síntesis de tiempo largo, obteniéndose la señal residuo de tiempo corto, que puede ser pasada al módulo de síntesis LPC.

#### Submódulo de síntesis LPC

Este submódulo recibe los coeficientes LAR hallados en el decodificador, y aplica el mismo procedimiento para hallar los coeficientes de reflexión que en el caso del codificador, sigue el esquema que se presenta en la figura 4.9.

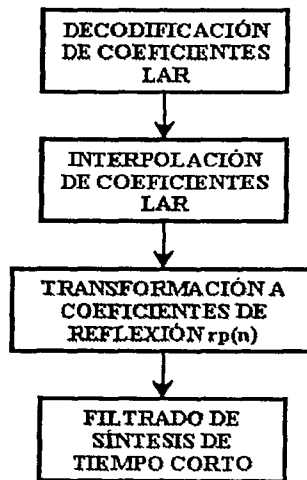


Figura 4.9 Submódulo de síntesis LPC

**Submódulo de posprocesamiento**

Tiene como misión eliminar el énfasis que fue introducido en la señal en el codificador por el filtro de énfasis en la sección de preprocesamiento. Es por tanto un filtrado inverso.

De esta manera puede llevarse a cabo la implementación de un algoritmo de compresión y descompresión a través de un DSP, utilizando el codec GSM para comunicaciones móviles.

# CAPÍTULO 5

## **ESTUDIO DEL HARDWARE**



## CAPÍTULO 5

### Estudio del hardware

#### 5.1 Requerimientos del equipo de cómputo para implantar el sistema de transmisión de voz

En la actualidad el avance que ha tenido la computación en cuanto a su arquitectura es grande y gracias a esto, se han podido mejorar los programas de telefonía en Internet. Sabemos que la eficacia de un programa de éstos se basa en la velocidad de transmisión, y en el tiempo de proceso en el que se comprime y descomprime la información. Afortunadamente las computadoras que hoy se encuentran en el mercado cumplen con todos los requerimientos para el procesamiento eficiente de la señal de voz.

Hace algunos años hablábamos de las computadoras 386 o 486 que en conjunción con los programas de telefonía no ayudaban mucho a tener transmisiones eficientes de voz, ya que los tiempos de compresión y descompresión de las señales eran grandes, debido a la velocidad que tenían las computadoras. Ahora, con los procesadores Pentium III o Pentium IV acercándose a los 2 GHz, es posible tener conversaciones con calidades sorprendentes. Hay que tomar en cuenta que además del procesador, hay otros dispositivos que contribuyen a mejorar la eficiencia de los programas de telefonía, como son los altavoces, la tarjeta de sonido, el modem y la estructura de la red.

Las características óptimas que debe tener una computadora personal para utilizar un sistema de transmisión de voz, son las siguientes (figura 5.1):

Hardware requerido:

PC mínimo estándar recomendable:

- \* Procesador Pentium
- \* 32 Mbytes en RAM
- \* 100 Mbytes de espacio libre en el disco duro para el Software
- \* Módem V.90/56k

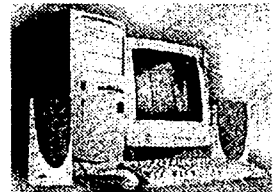


Figura 5.1 Requerimientos del equipo de cómputo

- \* Monitor VGA Color
- \* Conexión con un proveedor de servicio de Internet

PC recomendado como estándar óptimo:

- \* Procesador Celeron, Pentium III o Pentium IV de 600 Mhz (figura 5.2)
- \* 128 Mbytes en RAM
- \* 100 Mbytes de espacio libre en el disco duro
- \* MODEM V.90/56k
- \* Monitor SVGA
- \* Conexión con un proveedor de servicio de Internet



Figura 5.2 Procesadores de la familia Intel

#### Otros dispositivos

- \*Tarjeta de sonido bidireccional de 32 bits (pueden ser de la marca Sound Blaster)
- \*Un par de altavoces
- \*Micrófono unidireccional (capta sólo el sonido que viene de frente al micrófono)

#### 5.2 Tipos de módem

El módem es uno de los dispositivos fundamentales en la transmisión, ya que es el encargado de hacer la conversión de señal binaria a señal senoidal y viceversa. Dependiendo de la velocidad a la que transmiten los módems y sus aplicaciones, podemos encontrar los siguientes tipos básicos:

- \*Módem de marcación telefónica

\*Módem ISDN

\*Módem por cable

#### *Modem de marcación telefónica*

Los módem de marcación telefónica, como todos, convierten las señales digitales en señales senoidales, para que se puedan transmitir por una línea estándar. La razón del nombre de este dispositivo es que cuando el usuario marca un número telefónico, el MODEM lo identifica como tal y lo convierte en el código estándar (dos frecuencias para cada dígito).

Los módems de marcación telefónica vienen en varias velocidades (14.4, 28.8, 33.3 y 56 kbps) y con diferentes esquemas de compresión. En la actualidad, el módem que esta en el mercado es el de 56K PCI, que puede funcionar con cualquier procesador estándar.

#### *Módem ISDN*

ISDN son las iniciales de *Integrated Services Digital Network* (Red Digital de Servicios Integrados) y es la manera elegante de decir "servicios telefónicos digitales combinados". El ISDN permite la transmisión de voz, audio, video, fax y textos en una sola línea telefónica.

El ISDN puede manejar velocidades de transmisión más altas, debido a que la red dispone de canales de transmisión de alta capacidad (fibra óptica, cable coaxial o radio enlaces).

ISDN cuesta mucho más que los servicios de módem por cable y que una línea telefónica estándar, pero este costo se compensa con la mayor capacidad de transmisión del sistema.

Un módem ISDN se puede instalar en un puerto serie o dentro de un equipo, como si fuera una tarjeta adaptadora Ethernet. De este modo, si el usuario cuenta con el software adecuado, podrá obtener enlaces de gran calidad.

#### *Modem por cable.*

El modem por cable le permite al usuario realizar enlaces de más de 56 Kbits/seg. En este caso, la conexión a Internet por cable siempre está disponible y no es necesario comunicarse a la compañía para realizar la conexión; además, la línea telefónica del usuario

no se bloquea mientras utiliza la red y puede ver TV por cable con la misma conexión con la que recibe datos de Internet. La velocidad de transmisión de datos de un módem por cable puede llegar a 256 Kbps, que no se comparan con los 56 Kbps que ofrece el módem más accesible para los consumidores de marcación telefónica.

La conexión se realiza con un cable Ethernet, por lo que el usuario necesita un adaptador Ethernet en su computadora para conectarse al módem por cable. *Ethernet* es la manera elegante de nombrar un tipo de sistema de cableado, parecido a un cable telefónico grueso y que puede transferir datos a una velocidad de hasta 100 megabits/seg.

La mayoría de las computadoras personales; en particular las PC caseras, carecen de conexiones Ethernet; por ello es necesaria una tarjeta adaptadora cuando se planea usar un módem por cable. Las tarjetas adaptadoras Ethernet se conectan a la tarjeta madre de la computadora.

Para este tipo de módem por cable hay un inconveniente, el que de nada sirve tener este tipo de conexión si las personas con las que se lleva a cabo el enlace telefónico tienen un módem por marcación telefónica, ya que las velocidades de transmisión deben ser homogéneas para que el enlace tenga la calidad esperada.

### 5.3 Proveedores de Internet

Hablar de proveedores de Internet, es pensar que la elección de uno de éstos se basa en el costo y no en los servicios y características técnicas que nos puedan brindar. Podemos encontrar proveedores de servicio local, nacional o internacional. Por lo general, los primeros ofrecen precios competitivos pero sin acceso nacional, y junto con los proveedores de servicio nacional normalmente ofrecen software para que el usuario se conecte a Internet. Algunos proveedores de servicio comercial, a su vez, ofrecen características de lujo: secciones de charlas, servicios especializados de noticias e interfaces gráficas para navegar por la red fácilmente.

Con esta breve descripción se proponen algunos puntos que se deben tomar en consideración para la elección de un proveedor de Internet.

- **Servicios al cliente:** la capacidad que debe tener el proveedor de dar un eficaz soporte técnico a los usuarios en el momento que lo requieran.

- **Compatibilidad con la computadora de usuario:** verificar que el software preconfigurado para la conexión, si lo hay, sea compatible con la plataforma que la computadora maneja, con lo que este podrá correr sin dificultad. Además, mantendrá un servidor de archivos con las actualizaciones más recientes para navegadores.
- **Conexión eficiente:** de acuerdo a los requerimientos que se tengan, las conexiones que tienen anchos de banda grandes son las mejores, ya que transmiten a mayor velocidad la información y las aplicaciones son múltiples, como la videoconferencia en tiempo real.

# CAPÍTULO 6

## **SOFTWARE Y HARWARE PARA LA TRANSMISIÓN DE VOZ**

---

## CAPÍTULO 6

### Software y hardware para transmisión de voz

6.1 Evaluación del software y el hardware para transmisión de voz, disponible en el mercado.

Este capítulo tiene la finalidad de recabar la información técnica de los programas que en la actualidad funcionan como telefonía en Internet; además de ver cuales son los mejores accesorios que hay en el mercado para la transmisión de voz (como son los altavoces, y los micrófonos).

Para realizar una evaluación de los programas se han tomado como parámetros de medición la inteligibilidad, la facilidad de uso y el desempeño al realizar la conexión. El objetivo final de la evaluación es realizar mejoras en los programas de telefonía en Internet, a fin de que sean amigables para los usuarios.

Dentro del software para transmisión de voz podemos encontrar paquetes muy conocidos como son:

- Internet Phone de Intel
- Web.Max.Phone
- Internet Connection Phone de IBM
- Iris Phone
- WebPhone
- TeleVox
- SoftFone
- CoolTalk

A continuación se hará un breve análisis de cada uno de ellos.

#### Internet Phone de Intel

Fabricante: Intel Corporation

Plataforma: Windows 95

**Requerimientos mínimos:**

- ❑ Pentium a 90 MHz
- ❑ 5 MB de espacio libre en disco
- ❑ 16 MB en RAM
- ❑ Internet Explorer 2.0 de Microsoft o Navigator 2.0 de Netscape, si se desea utilizar la lista de directorio para buscar otros usuarios de Internet Phone de Intel. También se puede emplear de manera individual, para marcar directamente una dirección IP.

**Características:**

Internet Phone de Intel es uno de los primeros teléfonos en Internet que se adapta al estándar industrial del protocolo H.323 (ver Apéndice A); por ello, los usuarios de otros programas de teléfono en Internet que usan el mismo estándar pueden llamar a los usuarios de Internet Phone de Intel, sin importar el sistema operativo empleado.

El programa puede operar con o sin un navegador de la Web; sin embargo, si se utiliza el navegador de la Web (Netscape o Internet Explorer de Microsoft), se podrá encontrar con mayor facilidad a otros usuarios de Internet Phone que hayan descargado el programa y lo usen con frecuencia.

La conexión es rápida, a pesar de que en algunos momentos durante la transmisión se oyen cortes de voz, lo que ocasiona que se tenga baja calidad de audio. Podemos pensar que esto se debe a la versión de Internet Phone en la que se están realizando las pruebas, por lo que el programa no está actualizado para soportar las configuraciones actuales de las PC's .

**Web. Max. Phone**

**Fabricante:** Berkeley Systems

**Plataforma:** Windows 95, Windows 98

**Requerimientos mínimos:**

- ❑ Una PC 486 a 66 MHz
- ❑ 8 MB en RAM
- ❑ 5 MB de espacio libre en el disco sólo para el programa del teléfono; 18 MB para todo el sistema Web.Max.



**Características:**

Web. Max. Phone es fácil de configurar y puede funcionar con o sin los productos acompañantes: Web. Max. Retriever, Web. Max. TV y Web. Max. Security. A decir verdad la única información necesaria para configurar el Web. Max. Phone es el nombre del usuario y la dirección de correo electrónico correspondiente. De todo lo demás se encarga el servidor negociador al que uno se conecta al abrir el programa; sin embargo, se puede incluir una imagen que se transmitirá a la pantalla del otro usuario de Web. Max. Phone, cuando se establezca la conexión.

Literalmente no hay forma de controlar la compresión de la señal de voz; sólo hay algunos botones para agregar y quitar nombres de personas y direcciones de correo electrónico del directorio; además de un gran botón TALK (hablar) que uno debe oprimir para comunicarse. Simplemente se agrega una lista de nombres y la dirección de correo electrónico de una persona; luego, hay que hacer clic en el botón de marcación, para que el programa haga el resto. Un indicador ubicado en la parte inferior de la pantalla informa cuando llama alguien, el estado de la llamada y el estado de cualquier transferencia de archivos que se efectuó en esos momentos.

Por desgracia, no hay un directorio de interlocutores en línea; por lo cual, uno necesita conocer el nombre de alguien que ya tenga el programa. Una vez que se sepa este dato, el procedimiento se reduce a introducir la dirección del correo electrónico y oprimir el botón de marcación. La computadora del destinatario contesta de manera automática cuando recibe la llamada, pero el usuario debe aceptar o rechazar la llamada. La calidad del audio es aceptable, aun cuando en ciertos momentos se escuchan pequeñas interrupciones en la transmisión, pero bajo estas condiciones se pueden mantener conversaciones inteligibles.

**Internet Connection Phone de IBM**

**Fabricante:** IBM Corporation

**Plataforma:** Windows 95

**Requerimientos mínimos:**

- PC 486 a 66 MHz (aunque funciona bien en una máquina a 33 MHz)

- 8 MB en RAM
- 5 MB de espacio libre en el disco

**Características:**

Internet Connection Phone de IBM está basado enteramente en servidor y solo muestra a los usuarios que estén ejecutando el programa en ese preciso momento.

Después de seleccionar a la persona a quien se desea llamar, uno puede ajustar la compresión del sonido sobre la marcha, mientras avanza la conversación; asimismo, puede efectuar cambios en los controles del micrófono y el volumen mediante botones deslizables, lo que hace de este programa uno de los más completos con una elevada calidad de sonido.

Además se puede ver el nombre y la dirección de correo electrónico de los usuarios que estén conectados, con solo hacer clic sobre ellos en la lista mostrada por el servidor.

La interfaz es muy sencilla; lo único que se debe hacer es proporcionar nombre y dirección de correo electrónico del destinatario.

**IRIS Phone**

**Fabricante:** IRIS Systems

**Plataforma:** Windows 3x, Windows NT 3.5x, Windows 95

**Requerimientos mínimos:**

- PC 486/40 MHz
- 4 MB en RAM
- 2 MB de espacio libre en el disco

**Características:**

IRIS Phone es un programa con características incorporadas, incluye una agenda electrónica, correo de voz con la opción de grabar saludos para diferentes personas, capacidad de grabar conversaciones, un registro de llamadas bastante detallado, características especiales de procesamiento de llamadas (diferentes tonos de sonido para

determinados interlocutores), capacidad de reconocer el sistema operativo que tiene la máquina e instalar la versión apropiada del programa, de manera que se aproveche al máximo la capacidad del mismo.

La interfaz es simple y con un solo botón se tiene acceso a la mayoría de las características. El programa se puede ejecutar en segundo plano, para que el usuario trabaje en la computadora al mismo tiempo que recibe llamadas y graba mensajes de correo de voz.

El programa está diseñado para módems lentos y conexiones a personas que vivan muy lejos. Cuando la conexión es mala, se puede ajustar la velocidad de compresión del módem sobre la marcha. La calidad de sonido es muy buena, una vez ajustadas las especificaciones iniciales que correspondan a la configuración.

Cada vez que iris se inicia y conecta al servidor, se reciben automáticamente los mensajes de correo de voz que hayan llegado. Durante la charla, se pueden transferir textos.

### WebPhone

Fabricante: NetSpeak Corporation

Plataforma: Windows 3.1/WFW3.11, Windows NT, WIN-OS/2 Warp, Windows 95

### Requerimientos mínimos:

- PC 486DX a 33 MHz
- 4 MB en RAM
- 5 MB de espacio libre en el disco

### Características:

La interfaz de WebPhone parece un teléfono celular, pero actúa como un sistema telefónico totalmente equipado con la calidad de sonido equiparable al convencional (inteligible). Ofrece 4 líneas para entradas de llamadas, llamada en espera con música de fondo y características Do Not Disturb (no molestar) y de bloqueo.

También ofrece un tablero para texto de conversación. Tiene la capacidad para establecer conferencias con interlocutores múltiples al mismo tiempo y transferencia de llamadas. Se puede llamar usando la dirección de correo electrónico o la dirección IP de la otra persona. Si se desconoce la dirección de correo electrónico del interlocutor o la dirección IP es dinámica y cambia todo el tiempo, uno puede buscar el nombre de la persona en el servidor WebPhone. La primera vez que se realiza una conexión con alguien, el nombre se puede almacenar en la agenda, de modo que posteriormente, basta con marcar desde la página Web.

#### **TeleVox**

**Fabricante:** Voxware, Inc.

**Plataforma:** Windows 3.11, Windows NT versión 3.51, Windows 95

#### **Requerimientos mínimos:**

- PC 486DX a 66 MHz**
- 8 MB en RAM**
- 2 MB de espacio libre en el disco**

#### **Características:**

La interfaz es fácil de usar; ofrece operación a manos libres para los que tengan tarjetas de sonido bidireccionales. Cuando TeleVox se inicia por primera vez, el software se conecta a un servidor y este proporciona una lista de nombres de los usuarios de TeleVox que figuren en el directorio. Uno puede bloquear llamadas, identificar a su interlocutor antes de hablar con él y crear grupos privados.

Los diferentes controles de TeleVox le permiten cambiar el tono de voz de la misma forma que un ecualizador, para poder tener una conversación lo suficientemente inteligible y de buena calidad; además se puede realizar conversación por texto y transferir y recibir archivos durante la conversación.

## WebTalk

Fabricante: Quarterdeck Corporation

Plataforma: Windows 3.1, Windows para grupos de trabajo, Windows 95

### Requerimientos mínimos:

- PC 486 a 33 MHz
- 8 MB en RAM
- 13MB de espacio libre en el disco

### Características:

WebTalk ofrece una calidad de sonido muy buena, especificaciones de compresión ajustables y compatibilidad de plataforma, de manera que los usuarios de Windows y Mac pueden conversar entre sí. Este programa trabaja con los navegadores de Mosaic, Web de Quarterdeck, Netscape o Internet Explorer y ofrece conexión automática a otros usuarios de WebTalk con sólo hacer clic en la liga de página Web. Es posible emplear el programa en sí, para llamar a otros usuarios ya sea de manera directa con una dirección IP o mediante el servidor Quarterdeck.

## Hardware

### *Altavoces*

Los altavoces, mejor conocidos como bocinas, son los dispositivos más fáciles de comprar y encontrar dentro de un almacén que se dedique a la venta de equipo de cómputo; los podemos encontrar de diferentes tamaños, formas y potencias como pueden ser de 120, 180, 300, 480, 600 watts etc. y la única manera de verificar la calidad del sonido es probándolos.

*Tarjetas de sonido*

La tarjeta de sonido entre sus funciones tiene la de realizar las conversiones A/D y D/A. Como el convertidor A/D no puede digitalizar todos los valores de la señal analógica, es necesario seleccionar "unos pocos" y entregárselos para que los convierta a binario. Este proceso se llama "Muestreo". El proceso de digitalización en sí, produce un número constante de bits por cada muestra que se genera. Adicionalmente a estas funciones, la tarjeta; o mejor dicho, su hardware, realizan los procesos de compresión y expansión de la voz.

Algunas de las características que se deben tomar en cuenta para elegir una tarjeta de sonido son las siguientes:

Además de ser una tarjeta PCI de 32 bits (figura 6.1), la tarjeta de sonido debe contar de preferencia con la especificación DSP, que suelen aparecer impresas en los costados de la caja o en el manual del usuario.

Las letras DSP son las iniciales de *Digital Signal Processor* (procesador de señal digital); esto indica que la tarjeta puede desempeñar diversas tareas al mismo tiempo; por ejemplo, comprimir y descomprimir sonidos. El uso de una tarjeta de sonido con capacidad de multitareas liberará de cierto trabajo al procesador de la computadora.

Otros de los aspectos a tomar en cuenta es que la tarjeta sea bidireccional; es decir, que tenga la capacidad de transmitir el audio en ambos sentidos simultáneamente.



Figura 6.1 Tarjeta de sonido ideal para la implantación del sistema de transmisión de voz.

*Micrófonos*

Un micrófono de buena calidad también hace la diferencia entre una conversación inteligible y otra ininteligible. En el mercado existen comúnmente dos tipos de micrófonos: los omnidireccionales y los unidireccionales. Los primeros captan el sonido proveniente de todo el entorno; los segundos sólo captan los sonidos que llegan de frente. Por lo anterior, se recomienda que para entablar conversaciones por telefonía en Internet se ocupen los unidireccionales.

# CAPÍTULO 7

## **DESARROLLO**



---

## Capítulo 7

### Desarrollo

#### 7.1 Desarrollo de un sistema de comunicaciones para impartir asesorías a distancia a bajo costo

##### Introducción

##### *Definición del sistema*

El proyecto del cual nos ocupamos ahora, trata del desarrollo de un sistema de comunicaciones que sea capaz de transmitir audio vía Internet, para la implementación de asesoría a distancia para los diplomados que se imparten en el extranjero, a través de la programación de software (sockets) que integra la compresión y descompresión de datos, ayudados por la tecnología de hardware de la cual hoy podemos hacer uso.

##### *Objetivos del sistema*

El objetivo primordial del proyecto radica en lograr un enlace bidireccional o unidireccional de voz a través del Internet para llevar a cabo asesoría en los diplomados que se imparten a nivel internacional. Para este desarrollo habrá que apoyarnos de las técnicas de compresión y descompresión de datos, además de elementos de programación que comunican procesos por el Internet llamados Sockets y de un equipo de hardware que cumpla con las características mínimas requeridas para poder llevar a cabo el proceso de comunicación (figura 7.1).

La comunicación buscará cumplir en lo posible los siguientes requisitos:

- Llamada por Internet con calidad telefónica
- Comunicación bidireccional o unidireccional

- Aprovechamiento de la infraestructura con la que se cuenta a través de la optimización de los recursos para conseguir una transmisión a bajo costo.

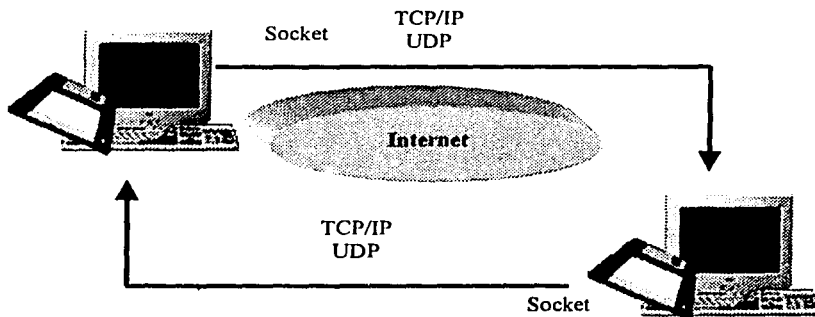


Figura 7.1 Esquema generalizado de la idea fundamental del proyecto

### *Partes del sistema*

#### Parte Hardware

Los elementos fundamentales para la transmisión son una computadora personal a la que se le conectan dispositivos de entrada y salida de datos como es el caso de los altavoces y micrófonos.

#### Información técnica óptima recomendable en la computadora personal

- Procesador Celeron/ Pentium III/ Pentium IV 500 MHz -1GHz; 500 MHz – 1.5 GHz o superior; 1.2 GHz – 2.0 GHz o superior, respectivamente.
- Memoria Ram de 128 Mb o en su defecto el 90% de la máxima que soporta la tarjeta madre.
- Tarjeta de sonido full-duplex (bidireccional) de 32 bits

- Altavoces de buena calidad de audio de los watts que se deseen.
- Micrófono

**Información técnica mínima recomendable en la computadora personal**

- Procesador Celeron/ Pentium III 300 MHz - 450 MHz; 450 MHz, respectivamente.
- Memoria Ram de 64 Mb - 128 Mb.
- Tarjeta de sonido full-duplex (bidireccional) de 32 bits
- Altavoces de buena calidad de audio de los watts que se deseen.
- Micrófono

**Software**

La mayor parte el proyecto esta basado en la programación que interactúa con los distintos dispositivos que hacen posible la transmisión del audio. Es por eso que la parte de software en su intervención en el proceso de transmisión se puede observar en la figura 7.2.

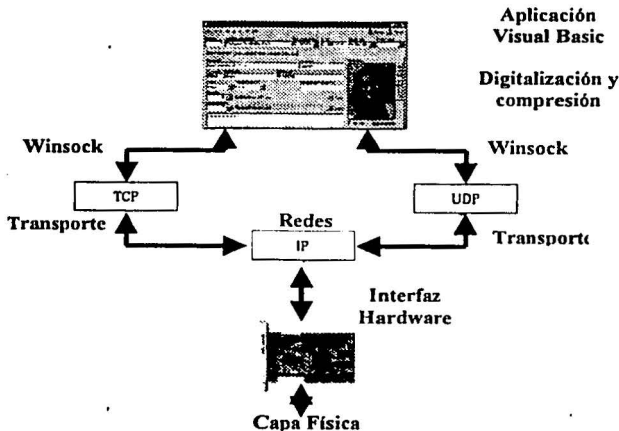


Figura 7.2 Diagrama general de la transmisión

Como se puede apreciar, el desarrollo de la aplicación tiene siempre una comunicación bidireccional con las capas de los protocolos; es decir, la comunicación esta basada en la arquitectura de los protocolos que se dividen en cinco capas, las cuales se describen a continuación:

- Capa de aplicación: contiene la lógica necesaria para llevar a cabo las aplicaciones del usuario, es decir, aquí se encuentra la interfaz de usuario que es la que recaba la información a transmitir.
- Capa de transporte o capa origen-destino (TCP y UDP): se encarga de controlar que los datos emanados de las aplicaciones lleguen correctamente y en orden a su destino en el caso de TCP, para UDP no verifica que los datos hallan llegado satisfactoriamente.
- Capa de red (IP): se encarga del encadenamiento a través de varias redes a través de la IP de la máquina.
- Capa de enlace a la red: es la responsable del intercambio de datos entre el sistema final y la red a la cual se está conectando.
- Capa física: es la encargada de utilizar el medio de transmisión de datos.

Con lo antes mencionado, damos paso a un análisis de nuestro sistema, en el que se expone los conceptos e ideas en las cuales se ha basado la elección de los métodos y forma de programación que son puntos importantes en el desarrollo del proyecto. Cabe aclarar que en el presente capítulo sólo se analiza el sistema de comunicaciones por medio de software y hardware, como la solución óptima que tenemos con respecto a las diferentes opciones de transmisión de voz propuestas en el capítulo tres.

### *Análisis del sistema*

#### **Exposición del problema**

Hoy en día, en la "era de las comunicaciones", la transmisión de datos entre distintos puntos geográficamente se hace imprescindible; además, nos encontramos ante la problemática de que los volúmenes de información cada vez se hacen mayores. Esto

implica, que debido a las limitaciones de los anchos de banda de los canales de transmisión, la información debe ser comprimida para que pueda ser transmitida.

En la actualidad, las técnicas de compresión de datos son muy utilizadas para reducir el tamaño de los archivos; esto reduce el espacio necesario en la unidad de almacenamiento y reduce también el tiempo o el ancho de banda de transmisión.

Podemos observar que se ha notado un cambio drástico en los tiempos de compresión y descompresión para escuchar y ver audio y video respectivamente; esto se debe al uso de técnicas cada vez mas perfeccionadas de codificación, las cuales parten de usar algoritmos que identifican la información que debe ser procesada y almacenada, eliminando bits no indispensables. Posteriormente esta información es reconstruida durante el proceso de decodificación, la cual, se realiza por medio de aplicaciones que se instalan de manera local, en las que intervienen de igual forma algoritmos, que hacen posible la presentación de esa información ante el usuario.

Las técnicas de compresión son diversas; cualquiera de ellas puede lograr la compresión de la voz, pero difieren en la velocidad con la que ejecutan el proceso; sin embargo, con las máquinas modernas de alta velocidad, esto no es significativo. La diferencia entre los algoritmos de compresión se expondrá a continuación.

**Algoritmos sin pérdidas:** Estos tipos de compresión utilizan el análisis estadístico para identificar y eliminar de los bits no necesarios, de manera que al descomprimir, se recupera absolutamente toda la información original. Estos algoritmos no logran una reducción notable en la cantidad de bytes producidos por el convertidor analógico digital, pero pueden proporcionar a los usuarios una calidad de voz muy buena.

**Algoritmos con pérdidas:** Analizando cuidadosamente la información recibida, estos algoritmos pueden decidir qué parte de la información es menos necesaria, a fin de que su pérdida no deteriore demasiado la calidad de la señal obtenida en el extremo receptor. El nivel de compresión logrado depende de dos cosas: de qué tanta información repetitiva contenga la señal y de qué tanta información indispensable podamos eliminar sin causar un daño importante a la señal. La compresión que se obtiene con estos procedimientos puede llegar a ser hasta de 100 a 1, con pérdidas no significativas. Este tipo de compresión se aplica con mayor frecuencia al vídeo, ya que en las imágenes se presenta una gran cantidad de información repetida.

En concreto, es desarrollar un sistema de compresión-descompresión y transmisión de voz por Internet.

### *Exposición de soluciones*

Dentro de las soluciones que se han formulado para satisfacer las necesidades que se presentan en las asesorías a distancia a bajo costo, tenemos las siguientes en materia de programación:

- Se hace imprescindible el uso de un lenguaje de programación sobre el cual recaerá la tarea de digitalización, compresión y descompresión del audio y voz. Los lenguajes de programación propuestos son: Visual C++, Visual Basic, Visual J++, C++, J++.
- El uso estructurado de una programación desarrollada específicamente para la comunicación vía Internet.
  - Implementación de componentes Active X (ocx).
  - Programación de Sockets llamados comúnmente Winsocks.

### **Soluciones en materia de equipo de computo de Internet:**

- Se busca la optimización de recursos, por lo que se proponen las características antes descritas como óptimas y mínimas recomendables para una transmisión eficiente.

### **Proveedores de servicios de Internet**

- Se busca el que tenga mayor desempeño de transmisión, menos interrupciones y cortes que puedan afectar a los objetivos del proyecto, dando como resultado una transmisión deficiente. Entre los proveedores de este servicio habrá que analizar a Prodigy de TELMEX, Axtel, Terra, Infosel y Red UNAM, entre los más conocidos.

*Solución escogida*

De acuerdo con los puntos descritos en los apartados anteriores, se ha adoptado como lenguaje de programación el Visual Basic, en una programación estructurada basada en módulos, que desempeñan cada uno, una función en particular en la que interviene la programación de los sockets que son parte fundamental en el enlace por la Internet entre los usuarios. Para el desarrollo y pruebas de transmisión, se eligieron equipos de cómputo con las características mínimas recomendables. El proveedor de servicios de Internet que por su eficiencia nos puede ayudar a que nuestra transmisión sea un éxito es Axtel. Los argumentos que soportan la solución escogida son los siguientes:

- Se ha adoptado como lenguaje de programación Visual Basic, por ser considerado uno de los más sencillos, debido en gran parte a su flexibilidad, por el entorno visual que facilita el desarrollo de programas de 32 bits para Windows; además de que contiene controles para el desarrollo de comunicaciones cliente-servidor como son el Winsock y los Active X.
- Por la forma en que se plantea el desarrollo del programa, es conveniente el uso del control Winsock del lenguaje de programación Visual Basic, ya que cumple con todas las funciones que requerimos para poder realizar un enlace entre usuarios por Internet; por el contrario el uso de componentes Active X originaría la compra de librerías y licencias para poder mantener en uso y funcionamiento el programa desarrollado.
- Todo el sistema de comunicaciones se probará en equipos de especificaciones recomendable mínimas, ya que debemos probar el sistema en las peores condiciones para poder realizar ajustes dentro del programa. Además, no se sabe con que equipos cuentan los alumnos de un diplomado que requieran asesorías.
- El proveedor de servicios de Internet Axtel (figura 7.3), es el que hemos elegido por su capacidad y rapidez de transmisión, ya que por la eficiencia de sus servidores, es posible obtener una tasa alta de trasmisión.



Figura 7.3 Axtel.net

Ya teniendo las herramientas sobre las cuales se basa el desarrollo de nuestro proyecto, a continuación se describirá de manera detallada las partes que conforman el programa de transmisión. Esta metodología nos ha de servir para entender el funcionamiento de dicho programa.

### *Entorno de desarrollo*

#### **Análisis de la parte software**

Como hemos podido observar en la primera parte del capítulo, se ha dado a conocer el diagrama general de la transmisión de voz y audio. Ahora es pertinente analizar ese mismo diagrama pero en bloques; es decir, desde que la voz es digitalizada, codificada y comprimida, hasta el momento de la transmisión en la que intervienen los protocolos de comunicación, el streaming audio y los Winsocks como enlace de aplicaciones sobre red.

#### **Introducción a la compresión de audio**

El sonido es una onda que se propaga a través del aire u otros medios físicos; está formado por variaciones alternativas de presión y puede detectarse por la medida del nivel de presión en un punto determinado. Si las ondas sonoras contienen información que deseamos transmitir, lo más conveniente es someterlas a un proceso de digitalización. Actualmente, las matemáticas nos proporcionan algoritmos para realizar digitalmente todos los procesos que antes se hacían en forma analógica; por ello, la mayoría de las operaciones realizadas sobre señales de audio son digitales, ya que tanto el almacenamiento como el procesado y la transmisión de la señal en forma digital ofrecen ventajas muy significativas frente los métodos analógicos. La información expresada en forma binaria, es menos influenciada por el ruido introducido al sistema. Además, se tiene la capacidad de incluir códigos de protección contra errores, así como encriptación. Con los mecanismos de decodificación adecuados, se pueden tratar simultáneamente señales de diferentes tipos por un mismo canal. La desventaja principal de la señal digital es que requiere un ancho de



banda mucho mayor que el de la señal analógica; de ahí que se hayan realizado (y se estén realizando) grandes esfuerzos para optimizar los procesos de compresión.

Para comprimir las señales analógicas tendrán que realizar los procesos que se describen a continuación, englobados en el concepto general de "Digitalización".

- Muestreo.
- Cuantificación.
- Codificación.

### Digitalización

El proceso de digitalización se compone de dos fases: muestreo y codificación analógica digital. En el muestreo, se almacena en un capacitor el voltaje de la señal en instantes discretos. El teorema de Nyquist dice que si la señal analógica tiene una frecuencia máxima de  $B$  Hz, entonces, la frecuencia de muestreo es como mínimo  $2B$  muestras/seg. Por tanto, siendo el margen superior de la audición humana en torno a los 20KHz, la frecuencia que garantiza un muestreo adecuado para cualquier sonido será de 40000 muestras/seg. Las frecuencias de la voz suelen ser muy inferiores a los mencionados 20 KHz; de tal manera que codificación de la voz, generalmente se realiza a no más de 10000 muestras/seg.

Después del proceso de muestreo se realiza la codificación. Aquí se presenta el problema de que el convertidor no puede digitalizar exactamente los valores muestreados, por lo que tiene que realizar un proceso de redondeo o de truncamiento; llamado generalmente cuantización o cuantificación. De modo que el número infinito de posibles valores de la señal analógica, tiene que ser redondeado o truncado a  $2^n$ , siendo  $n$  el número de bits con los que se realiza la conversión. Cuantos más bits se utilicen para la conversión, mayor precisión obtendremos y por tanto menor el error al atribuir una amplitud concreta al sonido en cada instante.

### Codificación

La codificación es la conversión de los valores cuantificados de la señal analógica a números binarios, para lo cual hay una enorme cantidad de algoritmos y circuitos. La señal binaria obtenida del convertidor se le llama PCM. La codificación subbanda o SBC (*Sub-*

*Band Coding*) es un método potente y flexible que permite codificar señales de audio eficientemente. El principio básico del **SBC** es la limitación del ancho de banda por eliminación de información en frecuencias no indispensables para la inteligibilidad. El resultado no es el mismo que la señal original, pero si el proceso se realiza correctamente, el oído humano no percibe gran diferencia.

La mayoría de los codificadores **SBC** utilizan el mismo esquema: primero tenemos un filtro o un banco de ellos, o algún otro mecanismo que descompone la señal de entrada en subbandas. A continuación se aplica un modelo psicoacústico que analiza las bandas y determina los niveles de enmascaramiento utilizando los datos psicoacústicos de que dispone. Considerando estos niveles de enmascaramiento, se cuantifican y codifican las muestras de cada banda. Si dentro de una banda hay una componente por debajo de un determinado nivel, se desecha. Si la componente supera dicho nivel, se calculan los bits necesarios para codificarla. Por último se empaquetan los datos según el estándar correspondiente que se esté utilizando, de manera que se puede recomponer la señal en el extremo receptor. La decodificación es muy sencilla, ya que no hay que aplicar ningún modelo psicoacústico. Se analizan los datos y se recomponen las bandas de frecuencia; las cuáles al ser superpuestas dan lugar a una señal similar a la original.

Los diagramas de las figuras 7.4 y 7.5 muestran la codificación y decodificación en forma general de la señal de voz respectivamente.

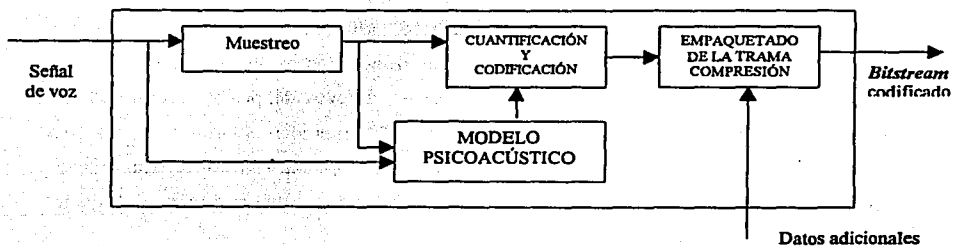


Figura 7.4 Diagrama Codificador

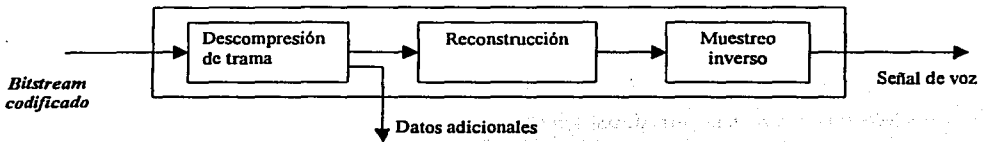


Figura 7.5 Diagrama Decodificador

Al observar los diagramas en los bloques de empaquetado de la trama en el codificador y la descompresión de la trama en el decodificador, veremos que aquí están los procesos de compresión y descompresión de los datos, en los cuales se utiliza el algoritmo de compresión / descompresión de Huffman.

Analicemos cómo se realiza la compresión por el algoritmo de Huffman y qué es el streaming audio.

### Algoritmo de compresión de Huffman

#### Generalidades

Se trata de un algoritmo de compresión que presenta características de encriptación. Este algoritmo se basa en asignar códigos de distinta longitud de bits a cada uno de los caracteres de un archivo. Si se asignan códigos más cortos a los caracteres que aparecen más a menudo, se consigue una compresión del archivo. Hay que aclarar que al hablar de caracteres nos estamos refiriendo tanto a símbolos alfanuméricos como a valores binarios correspondientes a la señal analógica.

Esta compresión es mayor cuando la variedad de caracteres diferentes que aparecen es menor. Por ejemplo: si el texto se compone únicamente de números o mayúsculas, se conseguirá una compresión mayor. Para recuperar el archivo original es necesario conocer el código asignado a cada carácter; que generalmente se envía antes de empezar a transmitir los datos.

Pasos del algoritmo

- \* Calcular las veces que aparece cada carácter (o elemento) del archivo.
- \* Ordenarlos de menor a mayor.
- \* Cada elemento se convierte en un árbol (de un sólo elemento)
- \* Mientras la lista tenga más de un elemento:
  - Sacar los dos primeros árboles formando un nuevo árbol con ellos.
  - Sumar sus frecuencias.
  - Insertar ordenadamente el árbol en la lista.
- \* Los códigos en binario de cada carácter se asignan en función del camino a recorrer en el árbol. Si en el recorrido se toma una rama izquierda se pone un 1 y si se coge una rama derecha se pone un 0, así hasta llegar al elemento.

Tomemos un texto corto, por ejemplo:

"ata la jaca a la estaca"

1) Contamos las veces que aparece cada carácter y hacemos una lista enlazada:

espacio(5), a(9), c(2), e(1), j(1), l(2), s(1), t(2)

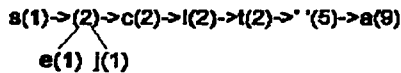
2) Ordenamos por frecuencia de menor a mayor

e(1), j(1), s(1), c(2), l(2), t(2), espacio(5), a(9)

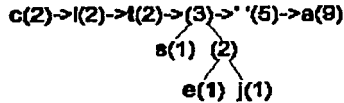
3) Consideremos ahora que cada elemento es el nodo raíz de un árbol.

e(1)->j(1)->s(1)->c(2)->l(2)->t(2)->' '(5)->a(9)

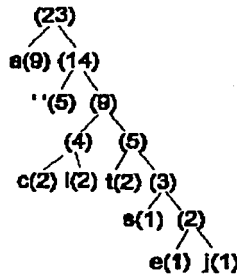
4) Fundimos los dos primeros nodos (árboles) en un nuevo árbol, sumamos sus frecuencias y lo colocamos en el lugar correspondiente:



Y sucesivamente:



El resultado final es:



5) Asignamos los códigos, las ramas a la izquierda son ceros, y a la derecha unos (por ejemplo), es una regla arbitraria.

a	'	c	l	t	s	e	j
0	10	1100	1101	1110	11110	111110	111111

6) Y traducimos el texto:

a	t	a	'	l	a	'	j	a	c	a	'	a	'	l	a	'	e	s	t	a	c	a
0	1110	0	10	1101	0	10	111111	0	1100	0	10	0	10	1101	0	10	111110	11110	1110	0	1100	0

7) Y sólo queda empaquetar los bits en grupos de ocho, es decir en bytes:

01110010	11010101	11111011	00010010	11010101	11110111	10111001	10000000
----------	----------	----------	----------	----------	----------	----------	----------

En total 8 bytes y el texto tenía originalmente 23. Es fácil comprobar que el mensaje preparado de esta manera tiene una sola forma de decodificarse, o sea, que es "unívoco".

Este proceso es mucho más recomendable para textos grandes que para cortos ya que hay ocasiones en que la codificación sale de mayor tamaño que el mensaje que había que codificar. Así es como el algoritmo de Huffman aplicado en la compresión y descompresión de información es implantado en el proceso de decodificación de audio / voz con una longitud variable sin pérdidas de información. Cabe señalar que los métodos de longitud variable se caracterizan, en general, por asignar palabras cortas a los eventos más frecuentes, dejando las largas para los eventos más infrecuentes.

### Streaming Audio

Uno de los factores que contribuyen al desarrollo óptimo de una buena transmisión de audio dentro del bloque de comunicaciones (transmisión), es la intervención de la tecnología streaming dentro del proceso recepción de la señal, ya que el streaming se utiliza para aligerar la descarga y ejecución de audio y vídeo en la web, permitiendo escuchar y visualizar los archivos mientras se están descargando.

Si no utilizamos streaming, para mostrar un contenido multimedia en la Red, tenemos que descargar primero el archivo entero en nuestro ordenador y más tarde ejecutarlo, para finalmente ver y oír lo que el archivo contenía. El streaming permite que esta tarea se realice de una manera más rápida y que podamos ver y escuchar su contenido durante la descarga.

El streaming funciona de la siguiente manera. Primero nuestra computadora (el cliente) se conecta con el servidor y éste le empieza a mandar el archivo. El cliente comienza a recibir el archivo y construye un buffer en el que se guarda la información. Cuando se ha llenado el buffer con una pequeña parte del archivo, la computadora lo empieza a mostrar y a la vez continúa con la descarga.

El sistema está sincronizado para que el archivo se pueda ver mientras que se descarga, de modo que cuando el archivo acaba de descargarse, también ha acabado de visualizarse. Si en algún momento la conexión sufre descensos de velocidad, se utiliza la información que hay en el buffer, de modo que se puede aguantar un poco ese descenso. Si la comunicación se corta demasiado tiempo, el buffer se vacía y la ejecución del archivo se suspende hasta que se reanude la transmisión de la señal. En la figura 7.6 se puede observar el diagrama

del seguimiento de la transmisión a nivel lógico con la intervención del streaming cuando forma un buffer de almacenamiento.

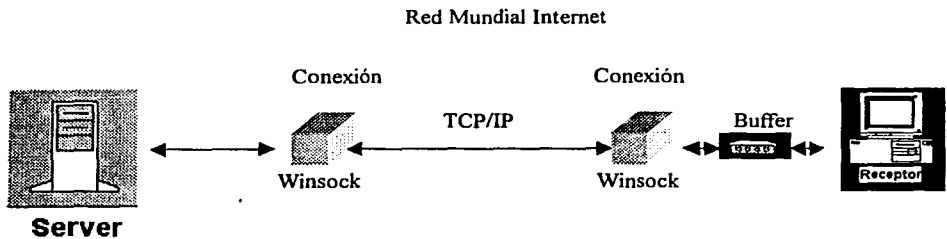


Figura 7.6 Diagrama Streaming

En las figuras 7.2 y 7.6 en las que se muestra el diagrama general de la transmisión y el diagrama streaming respectivamente, hay diversos elementos que guardan una estrecha relación por las funciones que desempeñan, ya que no sería posible una transmisión si suprimiéramos alguno de estos elementos; como por ejemplo, no podría llevarse a cabo el streaming audio si no existiera la capa de transporte, en dónde los protocolos de comunicación intervienen; tal es el caso de los protocolos TCP y UDP.

Para la transmisión de los datos, se elige como protocolo de transporte al TCP por la confiabilidad que ha mostrado a lo largo de su existencia. A continuación se detallan las semejanzas y diferencias que hay entre TCP y UDP con las que se ratifica y soporta la decisión tomada.

### TCP y UDP

La capa de transporte tienen como objetivo garantizar la comunicación fiable y a la vez eficiente, proporcionando una comunicación de extremo a extremo. Entre los protocolos de transporte encontramos al TCP y al UDP. Con los protocolos de transporte viene relacionada la dirección IP, encargada de realizar dos funciones: la segmentación y reensamblado de los mensajes y la supervisión del orden de llegada.

La estructura de los mensajes de control de los protocolos de este nivel se compone de:

- Un identificador de host por medio de su dirección IP.
- Un identificador de proceso por medio de la asignación de un puerto.
- Un identificador del protocolo que se está utilizando.

#### UDP (User Datagram Protocol)

Es un protocolo de transporte no orientado a conexión que proporciona un servicio de "datagramas de usuario", que se caracteriza por no incluir mecanismos que eviten la pérdida de información; es decir, no ofrece fiabilidad; lo cual implica que las aplicaciones que lo utilicen deben responsabilizarse de este tipo de problemas. Las direcciones de transporte UDP, añaden un número de puerto a la dirección IP. Los números de puerto identifican a los procesos emisor y receptor.

#### TCP (Protocolo de Control de Transporte)

Ciertas aplicaciones de usuario requieren asegurarse de que todos los datagramas han llegado correctamente a su destino. Los mecanismos para realizar estos servicios residen en TCP, el cuál debe ser capaz de establecer y gestionar sesiones (conexiones lógicas) entre los usuarios locales y los remotos. TCP, por tanto debe tener conocimiento de las actividades de los usuarios para dar soporte a la transferencia de sus datos por la red. TCP reside en el nivel de transporte del modelo de niveles convencional; está situado entre IP y los niveles superiores; ha sido diseñado para residir en los ordenadores o en las máquinas que se ocupan de conservar la integridad de la transferencia de datos entre extremos.

TCP es quién se tiene que encargar de las tareas de fiabilidad, control de flujo, secuenciamiento, aperturas y cierres. Puede soportar otros protocolos, como los de aplicación FTP (Protocolo de transferencia de archivos) y SMTP (Protocolo de transferencia de correo simple).



## Características TCP

TCP es un protocolo orientado a conexión. Esto quiere decir que TCP mantiene información del estado de cada cadena de datos de usuario que circula por él. Es responsable de la transferencia de datos entre extremos por la red o redes hasta la aplicación de usuario receptora. Es responsable de la transferencia fiable de cada uno de los caracteres que recibe del nivel superior correspondiente.

Cada byte transmitido lleva asignado un número de secuencia. El módulo TCP receptor utiliza una rutina de checksum para comprobar la posible existencia de daños en los datos producidos en el proceso de transmisión. Si los bytes son aceptables, se envía una aceptación positiva (ACK) al módulo TCP remitente. En cambio, si los datos han resultado dañados, el receptor los descarta y utiliza un número de secuencia para informar al remitente del problema. También, TCP emplea temporizadores para garantizar que no transcurra un lapso demasiado grande antes de la transmisión de aceptaciones desde el nodo receptor y/o de la transmisión de datos desde el nodo transmisor.

TCP recibe datos de un protocolo de nivel superior de forma orientada a cadenas; es decir, se envían caracteres separados y no bloques, tramas, datagramas, etc. Los datos son enviados byte a byte, y cuando llegan al nivel TCP, los bytes son agrupados para formar segmentos. Dichos segmentos se transfieren a IP para su transmisión.

TCP proporciona transmisión en modo full duplex entre las entidades que se comunican; así, la transmisión se puede efectuar en ambos sentidos sin necesidad de esperar la señal de indicación de cambio de sentido. Aparte de todo esto, también permite especificar niveles de seguridad y prioridades de las conexiones.

## La dirección IP

La IP está formado por 4 bytes y a cada aplicación que se encuentra registrada dentro de un servidor se le asigna un número de 16 bits llamado Puerto, que es fundamental especificarlo cuando se está programando. De tal manera que para poder establecer una

conexión con un servidor es necesario especificar su dirección que esta formada por su IP y por el número del puerto, por ejemplo:

254.90.128.56 : 8889

El rango de número de puerto es del 0 al 65535.

Los primeros números del 0 al 1023 se encuentran reservados para servicios del sistema como: http, FTP, Telnet, etc.

En este largo proceso sobre el cual ha viajado la señal de voz, desde su digitalización hasta el punto donde esta lista para transmitirse, viene aquí la última fase por la que pasa, en donde el protocolo TCP utiliza el Winsocket para comunicar el cliente con el puerto del servidor.

#### Winsock

Mejor conocido en la tecnología local como socket; es un componente que permite enlazar dos programas que se ejecutan a través de la red. Los procesos que se comunican pueden estar ejecutándose en el mismo o en distintos sistemas mientras se ejecuten por red. La comunicación entre procesos a través de sockets se basa en la filosofía CLIENTE-SERVIDOR; por tal razón, tanto el cliente como el servidor deberán usar el mismo protocolo de comunicación para que se realice un enlace eficiente.

El funcionamiento del socket es el que a continuación se describe:

Todo comienza cuando una aplicación servidor normalmente escucha a un puerto específico esperando una petición de conexión de un cliente. Cuando llega una petición de conexión, el cliente y el servidor establecen una conexión dedicada sobre la cual pueden comunicarse. Durante el proceso de conexión, el cliente es asignado a un número de puerto, y toma un socket en ese puerto. El cliente habla al servidor escribiendo sobre el socket y obtiene información del servidor cuando lee de él. Similarmente, el servidor obtiene un número nuevo de puerto local (necesita un nuevo puerto para poder continuar escuchando peticiones de conexión del puerto original). El servidor también toma un

socket en un determinado puerto local y comunica con él mediante la lectura y escritura sobre él (figura 7.7 y figura 7.8).

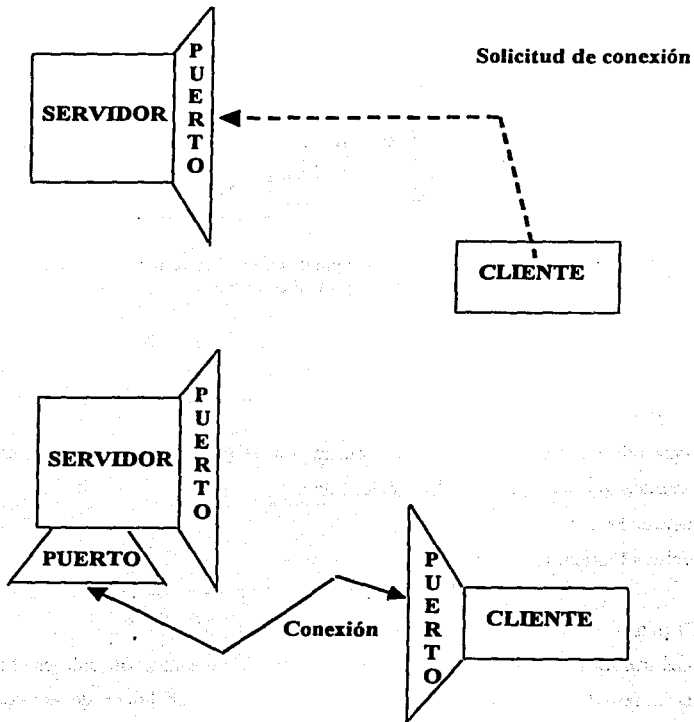


Figura 7.7 Diagrama de funcionamiento de un Socket

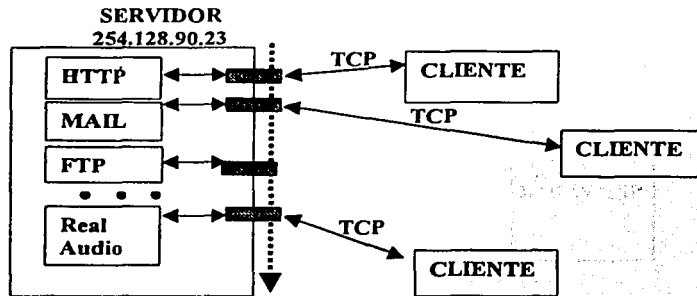


Figura 7.8 Diagrama del funcionamiento de los puertos en un servidor al recibir una petición de conexión por parte del cliente

### Tipos de Sockets

Hay diferentes tipos de sockets para la comunicación de procesos por Internet, pero son los que ocupan nuestro interés, los cuales son:

- Sockets Stream
- Sockets Datagram

### Sockets Stream

Son los más utilizados, hacen uso del protocolo TCP porque este protocolo es el que nos asegura el orden de los objetos durante la transmisión; están libres de errores; si se envían por el socket tres objetos "A, B, C", llegarán a su destino en el mismo orden "A, B, C". Al hacer uso de este socket, primero hay que establecer una conexión entre un par de sockets, ya que uno de los sockets hace la petición y el otro la recibe. Una vez que ambos sockets estén conectados, se pueden utilizar para transmitir información en ambas direcciones.

**Sockets Datagrama**

Hacen uso del protocolo UDP, y no necesitan de una conexión accesible como los sockets stream; son más eficientes pero no se garantiza la fiabilidad en la transmisión porque los datos se envían y se reciben en paquetes, cuya entrega no está garantizada porque los paquetes pueden ser duplicados, perdidos o llegar en un orden diferente del que se enviaron.

Como se ha mencionado, el protocolo de comunicación con datagrama es un protocolo sin conexión; es decir, que cada vez que se envíen datagramas es necesario enviar el descriptor del socket local y la dirección del socket que debe recibir el datagrama. Como se puede ver hay que enviar datos adicionales cada vez que se realice una comunicación.

A través de esta explicación de ambos sockets, podemos darnos cuenta por qué se ha tomado la decisión de elegir al socket stream como elemento de comunicación entre procesos, pero para apoyar esta decisión se ha de mostrar las diferencias entre un socket y otro en la tabla 7.1.

SOCKET DATAGRAMA	SOCKET STREAM
En cada envío de los datagramas hay que enviar con ellos el descriptor con el socket local y la dirección del socket que va a recibir los datagramas, por lo cual éstos son más grandes que los stream.	Antes de la transmisión hay que establecer la conexión entre dos sockets, lo que implica un tiempo empleado en la conexión, situación que no se presenta con los sockets datagrama.
Hay un límite de tamaño de los datagramas, establecidos en 64 kilobytes.	No hay límite de tamaño en los paquetes de información.
El socket datagrama conjuntamente con su protocolo UDP no garantizan que los datagramas enviados sean recibidos en el mismo orden por el socket de recepción.	Todos los datos que se envían se leen inmediatamente en el orden en que se enviaron.

Tabla 7.1

En la figura 7.9 se presenta el diagrama en el que se observa como se encuentran distribuidos los sockets en una comunicación vía Internet.

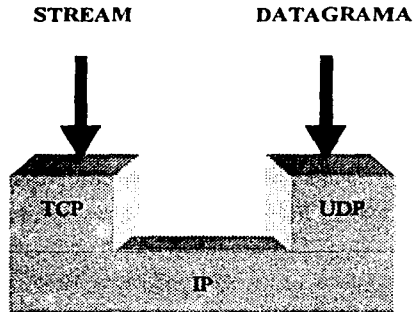


Figura 7.9 Arquitectura de la comunicación con sockets.

### *Implementación*

Con el marco teórico visto con anterioridad, estamos ya en la posibilidad de llevar el aspecto teórico a la implementación y desarrollo de un sistema de comunicaciones capaz de transmitir audio vía Internet.

Para el desarrollo del sistema, habrá que ayudarnos de la programación modular orientada a objetos, para simplificar y optimizar el trabajo; por lo cual, el programa estará dividido en módulos que desempeñarán una función específica dentro de la transmisión (véase figura 7.10).

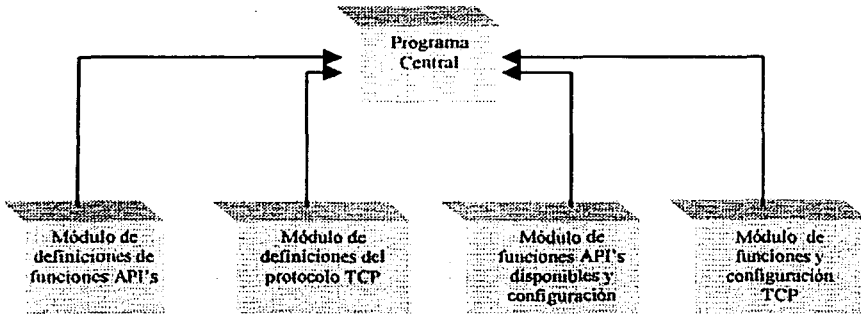


Figura 7.10 Diagrama a bloques que representa la estructura modular del programa de transmisión de audio

En la figura 7.11 podemos observar la Interfaz de usuario del programa de transmisión. La explicación de la estructura del programa se presentará a continuación, posteriormente se darán a conocer las referencias de instalación y funcionamiento del programa.

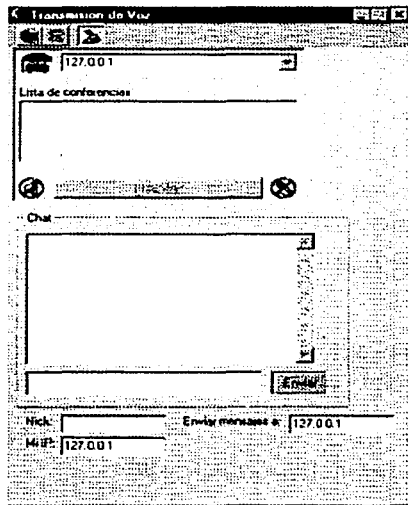


Figura 7.11 Interfaz de usuario para transmitir Voz via Internet

## Estructura del programa "Transmisión de Voz"

Recordemos que para transmitir voz por Internet, debemos llevar a cabo un proceso de digitalización y compresión de ésta, antes de ser transmitida. Para ello es necesario conocer los procedimientos que se pueden utilizar del sistema operativo Windows. Tales son los casos de las funciones de compresión y digitalización que requerimos para llevar a cabo la transmisión. Estos procedimientos reciben el nombre de API's (Interfaz de Programación de Aplicaciones), las cuáles permiten realizar cualquier actividad bajo Windows; como modificar el menú del sistema, ejecutar ficheros de sonido, controlar la finalización de Windows, trabajar con gráficos, etc. La utilización de las API's es esencial para el desarrollo del programa, ya que se tomarán del sistema operativo Windows las funciones de digitalización y compresión que están preestablecidas en él; así mismo, se introducirá una librería (dll) que no está registrada en el sistema, llamada "WaveStream", la cual integra funciones complementarias como el streaming audio que son fundamentales para la descompresión instantánea de la señal de voz.

En la estructura del programa es fundamental definir como primer paso los tipos de procedimientos (API's), ya que estos serán llamados en diversas ocasiones durante la ejecución del programa. Los procedimientos de audio tomados de la API se definirán en un módulo, que además de establecer las funciones o estructuras que se utilizarán en el tratamiento de la señal de voz, incluirán el dimensionamiento de variables que son utilizadas por algunas funciones tales como el sonido que se debe escuchar al llamar a recibir una llamada o el estado de la barra de herramientas al efectuarse una comunicación. En este mismo módulo se ha de cargar la librería en donde se encuentran localizadas estas funciones dentro del sistema operativo, la cual se llama Winmm.dll.

El módulo dónde se definen estas funciones es el siguiente:

### Option Explicit

```
Public Const WAVE_FORMAT_PCM = &H1
Public Const WAVE_FORMAT_ADPCM = &H11
Public Const WAVE_FORMAT_IMA_ADPCM = &H11
Public Const WAVE_FORMAT_DVI_ADPCM = &H11
Public Const WAVE_FORMAT_DSPGROUP_TRUESPEECH = &H22
```



```
Public Const WAVE_FORMAT_PCM610 = &H31
Public Const WAVE_FORMAT_MSN_AUDIO = &H3
Public Const TIMESLICE = 0.2
Public Const NoOfRings = 1
Public Const phoneHungUp = 3
Public Const phoneRingIng = 2
Public Const phoneAnswered = 1
Public Const mikeNO = 6
Public Const mikeOFF = 7
Public Const mikeON = 8
Public Const speakNO = 9
Public Const speakOFF = 10
Public Const speakON = 11
Public Const RingInId = 101
Public Const RingOutId = 102
Public Const tbCALL = 2
Public Const tbHANGUP = 3
Public Const tbAUTOANSWER = 5
Public Const SND_SYNC = &H0
Public Const SND_ASYNC = &H1
Public Const SND_NODEFAULT = &H2
Public Const SND_MEMORY = &H4
Public Const SND_LOOP = &H8
Public Const SND_NOSTOP = &H10
Declare Function sndPlaySound Lib "winmm.dll" Alias "sndPlaySoundA" (ByVal SoundData As Any, ByVal
uFlags As Long) As Long
```

De la misma forma se define un módulo que contendrá la declaración de algunas constantes que se manejarán en la configuración del socket TCP. El módulo es el que se escribe a continuación:

```
Option Explicit
Public Const MINTCP = 1
Public Const MAXTCP = 32
Public Const VOICEPORT = 701
Public Const NULLPORTID = 0
```

El siguiente módulo tiene la función de desplegar en un cuadro de texto las direcciones IP, el número de puerto en que se realiza la transmisión y los Hosts de los usuarios que se encuentran en comunicación. Así mismo controla la conexión y desconexión de los usuarios que entablan una comunicación. Dicho evento se ve reflejado en la aparición y desaparición de la información que despliega el cuadro de texto (dirección IP, número de puerto y host) al entablar o finalizar una comunicación entre ambos usuarios. Por último se comienzan a llamar a las API's del sistema para dar inicio a la digitalización y compresión de la señal de voz. El módulo es el siguiente:

Option Explicit

Public Sub InitServerList(ServerList As ComboBox)

    ServerList.AddItem "127.0.0.1"

    ServerList.AddItem "VBLABWK3"

    ServerList.AddItem "VBLABWK9"

End Sub

Public Sub DebugSocket(TCPSocket As Winsock)

    Debug.Print "TCPSocket.RemoteHost", TCPSocket.RemoteHost

    Debug.Print "TCPSocket.RemoteHostIP", TCPSocket.RemoteHostIP

    Debug.Print "TCPSocket.RemotePort", TCPSocket.RemotePort

    Debug.Print "TCPSocket.LocalHostName", TCPSocket.LocalHostName

    Debug.Print "TCPSocket.LocalIP", TCPSocket.LocalIP

    Debug.Print "TCPSocket.LocalPort", TCPSocket.LocalPort

    Debug.Print "TCPSocket.State", TCPSocket.State

    Debug.Print "====="

End Sub

Public Sub ResPlaySound(ResourceId As Long)

    Dim sndBuff As String

    sndBuff = StrConv(LoadResData(ResourceId, "WAVE"), vbUnicode)

    Call sndPlaySound(sndBuff, SND\_SYNC Or SND\_MEMORY)

End Sub

Public Sub AddConnectionToList(Socket As Winsock, ConnList As ListBox)

    Dim MemberID As String

    MemberID = Socket.RemoteHostIP & " [" & \_

        Format(Socket.RemotePort, "0") & "]" - [" & \_

        Format(Socket.LocalPort, "0") & "]"

    ConnList.AddItem MemberID

```

ConnList.ItemData(ConnList.NewIndex) = Socket.Index
End Sub

Public Sub RemoveConnectionFromList(Socket As Winsock, ConnList As ListBox)
    Dim Conn As Long
    Dim MemberID As String
    MemberID = Socket.RemoteHostIP & " [" & _
        Format(Socket.RemotePort, "0") & "]" - [" & _
        Format(Socket.LocalPort, "0") & "]"
    For Conn = 0 To ConnList.ListCount - 1
        If (ConnList.List(Conn) = MemberID) Then
            ConnList.RemoveItem Conn
        End If
    Next
End Sub

Public Sub GetIdxFromMemberID(Sockets As Variant, MemberID As String, Index As Long)
    Dim Idx As Long
    Dim LocPortID As Long
    Dim RemPortID As Long
    Dim RemoteIP As String
    Dim sStart As Long
    Dim sEnd As Long
    Dim Socket As Winsock

    sStart = 1
    sEnd = InStr(1, MemberID, " ") - 1
    If (sEnd > 1) Then
        RemoteIP = Mid(MemberID, sStart, sEnd)
        sStart = InStr(sEnd, MemberID, "[") + 1
        If (sStart > 1) Then
            sEnd = InStr(sStart, MemberID, "]") - 1
            If (sEnd > 2) Then
                RemPortID = Val(Mid(MemberID, sStart, sEnd))
                sStart = InStr(sEnd, MemberID, "[") + 1
                If (sStart > 1) Then
                    sEnd = InStr(sStart, MemberID, "]") - 1
                    If (sEnd > 2) Then
                        LocPortID = Val(Mid(MemberID, sStart, sEnd))
                        For Each Socket In Sockets

```

```

    If ((Socket.RemoteHostIP = RemoteIP) And _
        (Socket.RemotePort = RemPortID) And _
        (Socket.LocalPort = LocPortID) And _
        (Socket.Index > 0)) Then
        Index = Socket.Index
        Exit Sub
    End If
Next
End If
End If
End If
End If
End Sub

```

El último módulo que se programa tienen la característica de contener secuencias de captura de errores como medida de protección al programa, es decir, se piensa en lo que puede realizar el usuario para quebrantar el sistema y originar la interrupción de la transmisión o la generación de errores que son notificados en pantalla. Para evitar estas situaciones se programan estas secuencias de captura de errores, que aunque el usuario genere un error, el programa no interrumpirá su ejecución.

También en este módulo se establecen los parámetros de condiciones iniciales y finales después de efectuar la transmisión del socket como son: el número de puerto que se deja a la escucha de una petición de conexión, las configuraciones necesarias para realizar la conexión a una máquina remota que funge como servidor, las secuencias que permiten a otros equipos realizar conexión con nuestra PC y las instrucciones necesarias para concluir la comunicación (desconexión), dejando el puerto libre para alguna otra petición de conexión.

El módulo es el siguiente:

```

Option Explicit
Public Function InstanceTCP(TCPArray As Variant) As Long
    Dim Ind As Long
    InstanceTCP = -1
    On Error GoTo InitControl

```

```
For Ind = MINTCP To MAXTCP
  If (TCPArray(Ind).Name = "") Then
    End If
  Next
```

InitControl:

```
  On Error GoTo ErrorHandler

  If ((Ind >= MINTCP) And (Ind <= MAXTCP)) Then
    Load TCPArray(Ind)
    InstanceTCP = Ind
  End If
Exit Function
```

ErrorHandler:

```
  Debug.Print Err.Number, Err.Description
  Resume Next
End Function
```

Public Function Connect(Socket As Winsock, RemHost As String, RemPort As Long) As Boolean

```
  Connect = False
  Call CloseListen(Socket)
  Socket.LocalPort = 0
  Call Socket.Connect(RemHost, RemPort)

  Do While ((Socket.State = sckConnecting) Or _
    (Socket.State = sckConnectionPending) Or _
    (Socket.State = sckResolvingHost) Or _
    (Socket.State = sckHostResolved) Or _
    (Socket.State = sckOpen))
    DoEvents
  Loop

  Connect = (Socket.State = sckConnected)
End Function
```

---

```
Public Function Listen(Socket As Winsock) As Long
```

```
    If (Socket.State <> sckListening) Then
```

```
        If (Socket.LocalPort = 0) Then
```

```
            Socket.LocalPort = VOICEPORT
```

```
        End If
```

```
        Call Socket.Listen
```

```
    End If
```

```
End Function
```

```
Public Function CloseListen(Socket As Winsock) As Long
```

```
    If (Socket.State = sckListening) Then
```

```
        Socket.Close
```

```
    End If
```

```
End Function
```

```
Public Sub Disconnect(Socket As Winsock)
```

```
    If (Socket.State <> sckClosed) Then
```

```
        Socket.Close
```

```
    End If
```

```
End Sub
```

Ya definidos los módulos que conforman el programa, damos paso a la descripción general del programa central de comunicaciones. Es preciso recordar que los módulos que con anterioridad se han descrito tienen funciones (API's) y constantes que serán llamadas por el programa central cuando éste las utilice.

Para describir el funcionamiento del programa central, se darán las características de operación de cada evento que conforma el programa con la finalidad de evitar confusiones a lo largo del programa.

---

**Evento click del botón Hablar y el evento Mouseup del botón Hablar**

Al pulsar el botón hablar, el programa activará de forma automática los dispositivos de audio del sistema, llevando a cabo una verificación del funcionamiento de los mismos, para la grabación, digitalización, compresión y transmisión de la voz por Internet. La interfaz esta diseñada para semejar un funcionamiento idéntico al que tiene Windows, por lo que en este evento también se manipula la apariencia del cursor del mouse, al cambiar de su apariencia de flecha a la de un reloj de arena cuando procesa información. Este suceso ocurre mientras el puerto del socket detecte información perteneciente al procesamiento de la señal de voz; además, en este mismo evento, el icono de la bocinita que se encuentran en el costado izquierdo del botón hablar (véase la figura 7.11), se manipula, de tal forma que cuando se comienza la grabación de la voz, el icono pasa del estado inhabilitado al habilitado, lo cual puede distinguirse por la desaparición de la raya roja que atraviesa al icono. De la misma forma, cuando se acaba de transmitir la señal de voz y se deja de pulsar el botón "hablar" (evento Mouseup), el icono de la bocinita emite unas ondas, haciendo referencia a que se está transmitiendo y reproduciendo la señal de voz que había sido previamente digitalizada y comprimida del lado del usuario con el que se tiene comunicación.

Una secuencia que se ha incluido, es la que indica que hay una persona que se encuentra hablando, lo cual es una medida de seguridad para que no se genere algún error durante la ejecución del programa al intentar hablar más de una persona a la vez. Esto se puede observar en el cuadro de texto que es llamado servidor de lista de usuarios, en el que se despliega la dirección IP de la persona con la que se esta comunicando, porque ésta se selecciona con color azul siempre que se encuentra transmitiendo y se deselecciona cuando deja de transmitir. A continuación se muestra el código de este evento.

```
Private Sub cmdTalk_Click()
```

```
    Dim rc As Long
```

```
    Dim iPort As Integer
```

```
    Dim itm As Integer
```

```
    If (Not wStream.Playing And wStream.PlayDeviceFree And _
```

```

Not wStream.Recording And wStream.RecDeviceFree) Then
wStream.Playing = True
cmdTalk.Caption = "&Reproduciendo"
Screen.MousePointer = vbHourglass
iPort = wStream.StreamInQueue
Iport1 = iPort
Do While (iPort <> NULLPORTID)
    inLight.Picture = ImgIcons.ListImages(speakON).Picture
    inLight.Refresh
    For itm = 0 To ConnectionList.ListCount - 1
        If (ConnectionList.ItemData(itm) = iPort) Then
            ConnectionList.TopIndex = itm
            ConnectionList.Selected(itm) = True
            Exit For
        End If
    Next
    If cnt = 1 Then cnt = 3
    If cnt = 0 Then cnt = 1
    rc = wStream.PlayWave(Me.hWnd, iPort)
    Call wStream.RemoveStreamFromQueue(iPort)
    iPort = wStream.StreamInQueue
    inLight.Picture = ImgIcons.ListImages(speakOFF).Picture
    inLight.Refresh
Loop
ConnectionList.TopIndex = 0
If (ConnectionList.ListCount > 0) Then
    ConnectionList.Selected(0) = True
    ConnectionList.Selected(0) = False
End If
Screen.MousePointer = vbDefault
cmdTalk.Caption = "&Hablar"
wStream.Playing = False
End If
End Sub

Private Sub cmdTalk_MouseUp(Button As Integer, Shift As Integer, x As Single, Y As Single)
wStream.Recording = False
End Sub

```



Evento Click sostenido del botón Hablar

En este evento al quedarse uno pulsando el botón Hablar con el puntero de mouse, se inicializa la grabación, digitalización y compresión de la señal de voz, quedando lista para ser transmitida y reproducida del lado del usuario con el que se mantiene la conversación. Cabe aclarar que en este punto la información a transmitir es enviada al puerto de salida del socket, ya que éste es el que mantiene el enlace de aplicaciones entre usuarios por Internet y el que transmite la información.

```
Private Sub cmdTalk_MouseDown(Button As Integer, Shift As Integer, x As Single, Y As Single)
    Dim rc As Long
```

```
    If (Not wStream.Playing And _
        Not wStream.Recording And _
        wStream.RecDeviceFree And _
        wStream.PlayDeviceFree) Then
        wStream.Recording = True
        cmdTalk.Caption = "&Hablando"
        Screen.MousePointer = vbHourglass
        outLight.Picture = ImgIcons.ListImages(mikeON).Picture
        outLight.Refresh
```

```
    rc = wStream.RecordWave(Me.hWnd, TCPSocket)
```

```
    outLight.Picture = ImgIcons.ListImages(mikeOFF).Picture
    outLight.Refresh
    Screen.MousePointer = vbDefault
    cmdTalk.Caption = "&Hablar"
```

```
    If Not wStream.Playing And _
        wStream.PlayDeviceFree And _
        wStream.RecDeviceFree Then
        Call cmdTalk_Click
```

```
    End If
```

```
End If
```

```
End Sub
```

Evento doble Click en la lista de servidor de usuarios

Este evento tiene la finalidad de terminar una conversación con otro usuario, con tan sólo dar doble Click en la dirección IP que se muestra en la lista de servidor de usuarios. En el momento en el que se ejecuta este evento, la dirección IP del usuario desaparece de la lista al mismo tiempo que el socket realiza la desconexión de las aplicaciones y los dispositivos de audio se desactivan quedando listos para una siguiente conexión. Además al realizarse la desconexión, los iconos de la bocina y micrófono los observará el usuario desactivados, ya que nuevamente volverán a aparecer sobre ellos las líneas rojas que los atraviesan en señal de que han quedado inhabilitados.

El código es el siguiente:

```
Private Sub ConnectionList_DblClick()
```

```
    Dim MemberID As String
```

```
    Dim Idx As Long
```

```
    If (ConnectionList.Text = "") Then Exit Sub
```

```
    MemberID = ConnectionList.List(ConnectionList.ListIndex)
```

```
    Call GetIdxFromMemberID(TCPsocket, MemberID, Idx)
```

```
    Call RemoveConnectionFromList(TCPsocket(IdIdx), ConnectionList)
```

```
    Call Disconnect(TCPsocket(IdIdx))
```

```
    Unload TCPsocket(IdIdx)
```

```
    cmdTalk.Enabled = (ConnectionList.ListCount > 0)
```

```
    Tools.Buttons(tbHANGUP).Enabled = (ConnectionList.Text <> "")
```

```
    If Not cmdTalk.Enabled Then
```

```
        inLight.Picture = ImgIcons.ListImages(speakNO).Picture
```

```
        outLight.Picture = ImgIcons.ListImages(mikeNO).Picture
```

```
    End If
```

```
End Sub
```

Dentro del programa también hay procedimientos que no hacen referencia a ningún evento como es el caso del procedimiento "Load" y el procedimiento "Unload". En el primero de ellos siempre se cargan las condiciones iniciales del programa, como son todas aquellas funciones de las API's que contienen las funciones de grabación, digitalización y compresión de la señal de voz. En este caso en la función de digitalización obtenemos una señal PCM, la cuál se comprimirá a través del codec GSM610 (formato de compresión que se basan en los algoritmos de Huffman para comprimir datos). Dentro de este procedimiento también se cargan las configuraciones iniciales de la barra de herramientas, controles (botones, cuadros de textos, iconos etc.) que se despliegan en la interfaz de usuario y configuraciones muy específicas del socket.

Por su parte el procedimiento "Unload" se encarga de realizar una verificación de la desconexión total de las aplicaciones en red, además de cerrar el programa en su totalidad.

Los códigos de ambos procedimientos son los siguientes:

```
Private Sub Form_Load()
```

```
Dim rc As Long
```

```
Dim idx As Long
```

```
Dim TCPidx As Long
```

```
CLOSINGAPPLICATION = False
```

```
Call InitServerList(txtServer)
```

```
txtServer.Text = txtServer.List(0)
```

```
imgStatus = ImgIcons.ListImages(phoneHungUp).Picture
```

```
Set wStream = CreateObject("WaveStreaming.WaveStream")
```

```
Call wStream.InitACMCodec(WAVE_FORMAT_PCM610, TIMESLICE)
```

```
cmdTalk.Enabled = False
```

```
Tools.Buttons(tbHANGUP).Enabled = (ConnectionList.Text <> "")
```

```
inLight.Picture = ImgIcons.ListImages(speakNO).Picture
```

```
outLight.Picture = ImgIcons.ListImages(mikeNO).Picture
```

```
Call Listen(TCPSocket(0))
```

```
Text2.SelStart = Len(Text2.Text)
```

```
udp.RemoteHost = "127.0.0.1"
```

```
udp.RemotePort = 1001
```

```
udp.Bind
```

```

Text4.Text = udp.LocalIP
End Sub

Private Sub Form_Unload(Cancel As Integer)


    Dim Idx As Long
    Dim Socket As Winsock

    CLOSINGAPPLICATION = True
    For Each Socket In TCPSocket
        Call Disconnect(Socket)
    Next
    Set wStream = Nothing
End
udp.Close
End Sub

```

Dentro del programa ha llegado el momento de explicar las estructuras de programación que tiene el socket; no hay que olvidar que sin él no habría comunicación de aplicaciones por Internet.

Como se había mencionado en apartados anteriores dentro de la tesis, “dos programas se pueden comunicar entre sí por medio de un control que en la mayoría de los lenguajes de programación se denomina Winsock Control”.

El socket en Visual Basic es un control que contiene una serie de propiedades, procedimientos y eventos esenciales para lograr una transmisión de datos. Visualmente el socket es de esta forma. 

A continuación se hablarán de las propiedades, procedimientos y eventos que posee el socket, con la finalidad de poder entender posteriormente el código de algunos de sus procedimientos y métodos.

Lista de la propiedades más importantes:

**LocalIp:** devuelve la dirección IP de la máquina local, por ejemplo (132.248.96.101).

**LocalHostName:** devuelve el nombre de la máquina local.

**RemoteHost:** establece el equipo remoto al que se quiere solicitar la conexión.

**LocalPort:** establece el puerto que se quiere dejar a la escucha de una petición de conexión.

**RemotePort:** establece el número del puerto remoto al que se quiere uno conectar.

**State:** verifica si el control Winsock está siendo utilizado o no.

#### Lista de procedimientos más importantes

**Accept:** utilizado sólo en aplicaciones de servidor TCP. Este procedimiento se utiliza para aceptar una conexión entrante cuando se ejecuta el evento **ConnectionRequest**.

**GetData:** recupera el bloque actual de datos y lo almacena en una variable tipo **variant**.

**Listen:** toma el número de puerto que se indicó para la transmisión y lo deja a la escucha de una petición de conexión.

**SendData:** envía datos a un equipo remoto.

#### Lista de eventos más importantes

**ConnectionRequest:** se produce cuando el equipo remoto solicita una conexión. Sin este evento no se puede llevar a cabo la conexión.

**Connect:** se produce cuando el equipo local se conecta al equipo remoto y se establece una conexión.

**Close:** se produce cuando el equipo remoto cierra la conexión. Las aplicaciones deben usar el evento **close** para cerrar correctamente una conexión TCP.

**DataArrival:** se produce cuando llegan nuevos datos.

A partir de esta descripción de las propiedades, procedimientos y eventos del control Winsock, se presentarán los códigos de eventos que se han descrito con anterioridad.

#### Evento Close

Cierra la conexión, desconecta el puerto remoto del puerto local. Por otra parte remueve por seguridad todos las direcciones IP que se encuentran en la lista de servidor de usuarios; además, según el estado habilitado o deshabilitado del botón "Hablar", es que se permite mostrar los iconos habilitados de la bocina y el micrófono. El código es el siguiente:

```
Private Sub TCPSocket_Close(Index As Integer)

    Call RemoveConnectionFromList(TCPSocket(Index), ConnectionList)
    Call Disconnect(TCPSocket(Index))

    cmdTalk.Enabled = (ConnectionList.ListCount > 0)
    If Not cmdTalk.Enabled Then
        inLight.Picture = ImgIcons.ListImages(speakNO).Picture
        outLight.Picture = ImgIcons.ListImages(mikeNO).Picture
    End If

    Tools.Buttons(tbHANGUP).Enabled = (ConnectionList.Text <> "")
    If cmdTalk.Enabled Then
        imgStatus = ImgIcons.ListImages(phoneHungUp).Picture
    End If

    Unload TCPSocket(Index)

End Sub
```

### Evento ConnectionRequest

En este evento, como se indicó, se solicita la petición a conexión. Dentro de esta secuencia de programación, en la barra de herramientas se puede observar un icono que despliega un mensaje de auto contestación, el cual si es habilitado despliega un cuadro de mensaje en el que se pregunta ¿Quieres contestar la llamada de la dirección IP xxx.xxx.xxx.xxx? cuando hay una petición de conexión. Si es aceptada la conexión, la dirección IP del usuario es agregada a la lista de servidor de usuarios y se habilitan los controles necesarios para poder entablar la conversación; por el contrario si la petición no es aceptada, el programa no ejecuta ningún otro evento. El código se muestra a continuación:

```
Private Sub TCPSocket_ConnectionRequest(Index As Integer, ByVal requestID As Long)

    Dim rc As Long
    Dim idx As Long
    Dim RemHost As String
```

```

If (TCPSocket(Index).RemoteHost <> "") Then
    RemHost = UCase(TCPSocket(Index).RemoteHost)
Else
    RemHost = UCase(TCPSocket(Index).RemoteHostIP)
End If

If (Tools.Buttons(tbAUTOANSWER).Value = tbrUnpressed) Then
    rc = MsgBox("Incomming call from [" & RemHost & "]..." & vbCrLf & _
        "Do you wish to answer?", vbYesNo)
Else
    rc = vbYes
End If

If (rc = vbYes) Then
    Idx = InstanceTCP(TCPSocket)
    If (Idx > 0) Then
        TCPSocket(Idx).LocalPort = 0
        Call TCPSocket(Idx).Accept(requestID)
        Call AddConnectionToList(TCPSocket(Idx), ConnectionList)

        imgStatus = ImgIcons.ListImages(phoneRingIng).Picture
        Call ResPlaySound(RingInId)
        imgStatus = ImgIcons.ListImages(phoneAnswered).Picture
        cmdTalk.Enabled = True
        Tools.Buttons(tbHANGUP).Enabled = (ConnectionList.Text <> "")
    End If
End If
End Sub

```

### Evento Connect

Este evento realiza la conexión entre el equipo local y el equipo remoto. En este instante se agrega a la lista de servidor de usuarios la dirección IP de la máquina remota en la lista de usuarios local y la dirección IP de la máquina local en la lista de usuarios de servidor remoto. Cuando se realiza la conexión, se despliega el icono de un teléfono sonando en

señal de que alguien quiere establecer una conversación. Posteriormente cuando se acepta la conversación se despliega un icono de un teléfono descolgado. Veamos el código:

```
Private Sub TCPSocket_Connect(Index As Integer)

    Call AddConnectionToList(TCPSocket(Index), ConnectionList)
    imgStatus = ImgIcons.ListImages(phoneRingIng).Picture
    Call ResPlaySound(RingOutId)

    imgStatus = ImgIcons.ListImages(phoneAnswered).Picture
    cmdTalk.Enabled = True
    Tools.Buttons(tbHANGUP).Enabled = (ConnectionList.Text <> "")
End Sub
```

#### Evento DataArrival

Este evento es el encargado de procesar la información que ha sido transmitida por Internet, por lo que aquí se encuentran las funciones del streaming audio. El evento se ejecuta mientras en el puerto del socket se detecte la presencia de bytes, los cuales son redireccionados al buffer del socket. Cuando la información se encuentra en el buffer se comienza el proceso de recuperación de la señal; es decir, se inician las etapas de descompresión y decodificación por medio de las funciones streaming audio, las cuáles permiten seguir recibiendo datos en el buffer para la recuperación de la señal que es reproducida de manera instantánea después de haber sido transmitida. El código es el que se muestra a continuación:

```
Private Sub TCPSocket_DataArrival(Index As Integer, ByVal bytesTotal As Long)
    Dim rc As Long
    Dim WaveData() As Byte
    Dim WaveData1() As Byte
    Static ExBytes(MAXTCP) As Long
    Static ExData(MAXTCP) As Variant
    Dim ExData1(MAXTCP) As String
    Dim i As Long
    Dim x As Long
```



With wStream

If (TCPSocket(Index).BytesReceived > 0) Then

Do While (TCPSocket(Index).BytesReceived > 0)

If (ExBytes(Index) = 0) Then

If (.waveChunkSize <= TCPSocket(Index).BytesReceived) Then

Call TCPSocket(Index).GetData(WaveData, vbByte + vbArray, .waveChunkSize)

i = UBound(WaveData)

x = InstrBt(WaveData, 255)

Label3.Caption = "One " + Str(x)

Label1.Caption = "One " + Str(WaveData(i))

Call .SaveStreamBuffer(Index, WaveData)

Call .AddStreamToQueue(Index)

Else

ExBytes(Index) = TCPSocket(Index).BytesReceived

Call TCPSocket(Index).GetData(ExData(Index), vbByte + vbArray, ExBytes(Index))

End If

Else

Call TCPSocket(Index).GetData(WaveData, vbByte + vbArray, .waveChunkSize - ExBytes(Index))

ExData(Index) = MidB(ExData(Index), 1) & MidB(WaveData, 1)

WaveData1 = ExData(Index)

i = UBound(WaveData1)

x = InstrBt(WaveData1, 255)

Label3.Caption = "Two " + Str(x)

Label1.Caption = "Two " + Str(WaveData1(i))

Call .SaveStreamBuffer(Index, WaveData1)

Call .AddStreamToQueue(Index)

ExBytes(Index) = 0

ExData(Index) = ""

End If

Loop

If (Not .Playing And .PlayDeviceFree And \_

Not .Recording And .RecDeviceFree) Then

Call cmdTalk\_Click

End If

End If

End With

End Sub

Finalmente las últimas secuencias de programación son las que se encargan de habilitar ciertos controles de la interfaz, dependiendo de las selecciones que se realicen en la barra de herramientas. En el caso de seleccionar el icono de la barra de herramientas en el que se despliega la leyenda "Llamar", se habilitará si es que se realizó la conexión el botón "Hablar". Cuando se selecciona el icono "Colgar", desaparece de la lista de servidor de usuarios la dirección IP de la persona con la que entablamos la comunicación. Por último si es seleccionado el icono de "Auto contestación", sólo se observará que el icono cambia de un teléfono colgado a uno descolgado en espera de que alguien quiera establecer una conexión con nosotros.

De esta manera es como se ha programado esta sencilla interfaz que es útil para comunicarse con otros usuarios, ya sea en red local o por Internet. A continuación se dará a conocer el manual del usuario que contiene las instrucciones de instalación y uso; posteriormente se exponen los resultados experimentales de las pruebas que se efectuaron a este programa de transmisión de voz por Internet.

## Manual de Usuario

Este apartado tiene como finalidad, dar una descripción del uso del programa que se ha desarrollado para la transmisión de voz.

### Descompresión del programa

Como primer paso inserte en la unidad de 3<sup>1/2</sup> el disco que acompaña esta tesis. Cuando tenga a la vista la información que posee el disco flexible, visualizará dos carpetas una llamada "Voz" y otra llamada "Manual", como lo muestra la figura 7.12.

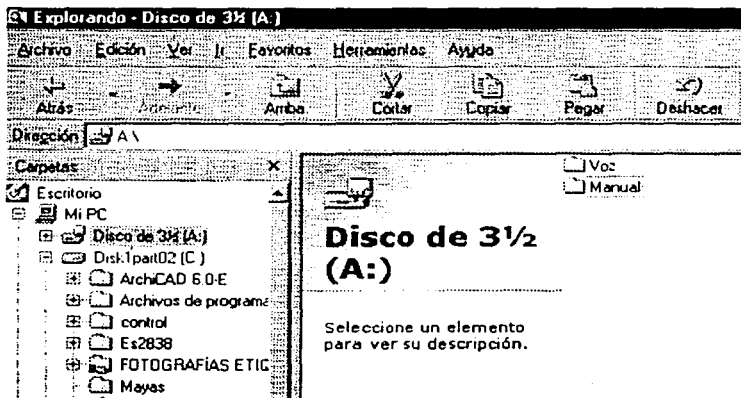


Figura 7.12

Abra la carpeta llamada Voz y encontrará de igual forma un archivo llamado "Voz" en formato comprimido (zip) como lo muestra la figura 7.13, el cual debe ser descomprimido por el programa Winzip.

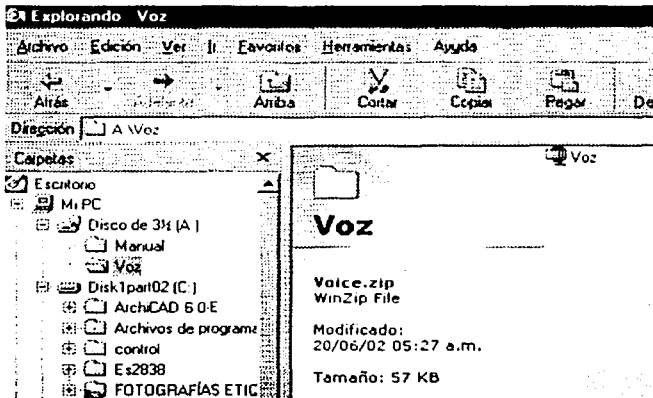


Figura 7.13

Cuando ejecute el programa Winzip para la descompresión del archivo "Voz", encontrará archivos que en su conjunto componen el código fuente del programa, además se encuentra el archivo ejecutable del programa con el nombre de "Voz.exe" (véase la figura 7.14).

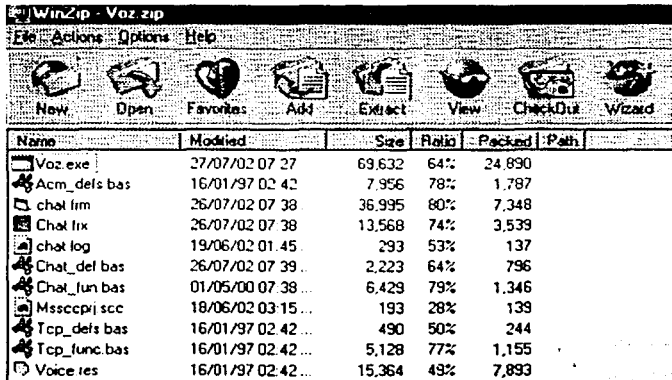


Figura 7.14

Para poder utilizar sin restricciones los archivos, es necesario extraerlos del archivo comprimido, por lo que se necesitará crear una carpeta con el nombre que desee y ubicarla de preferencia en la raíz de la unidad C:\, para nuestro caso la carpeta creada la llamaremos "Proyecto"; después de haber creado la carpeta, vaya a la barra de herramientas del programa Winzip y pulse el icono Extract, a continuación hay que especificar la ubicación de la carpeta que se acaba de crear, ya que en ésta quedarán grabados los archivos que se observan en la figura 3, finalmente pulsamos aceptar para terminar la descompresión.

Ahora vaya a su carpeta donde extrajo los archivos que se acaban de descomprimir y examínela, aquí se encuentran una serie de archivos que son fundamentales para la ejecución del programa de transmisión. En la figura 7.15 se muestran los archivos que deben estar contenidos en su carpeta.

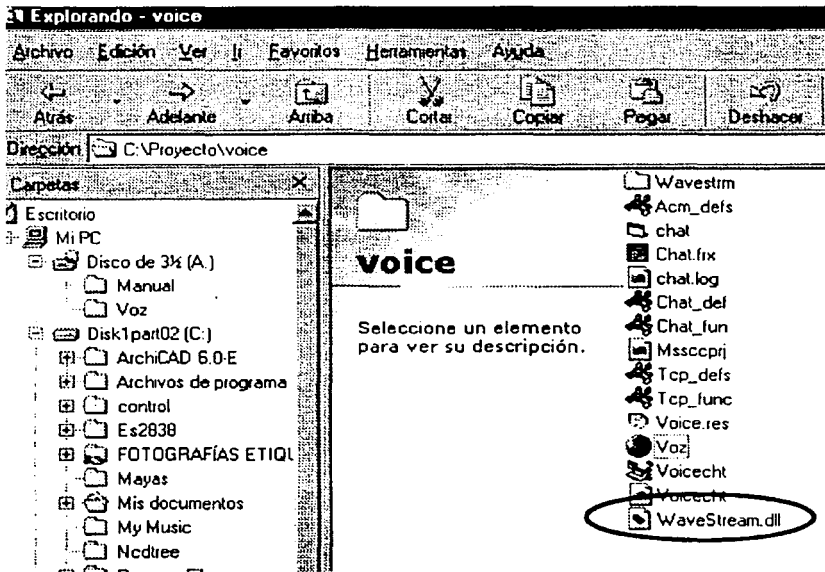


Figura 7.15

En la misma figura 7.15 esta marcado el archivo WaveStream.dll, el cuál es una librería que se debe dar de alta en el sistema, ya que de omitir este paso no se podrá ejecutar el programa.

#### Alta de la librería WaveStream.dll

Para dar de alta la librería WaveStream debemos escribir en el cuadro de mensaje que se despliega en la opción ejecutar (pulsar el botón de Inicio – Ejecutar), la siguiente sintaxis “regsvr32 C:\Proyecto\voice\Wavestream.dll”, finalmente pulsar aceptar. Si la librería fue cargada correctamente aparecerá un mensaje como el que se muestra en la figura 7.16.

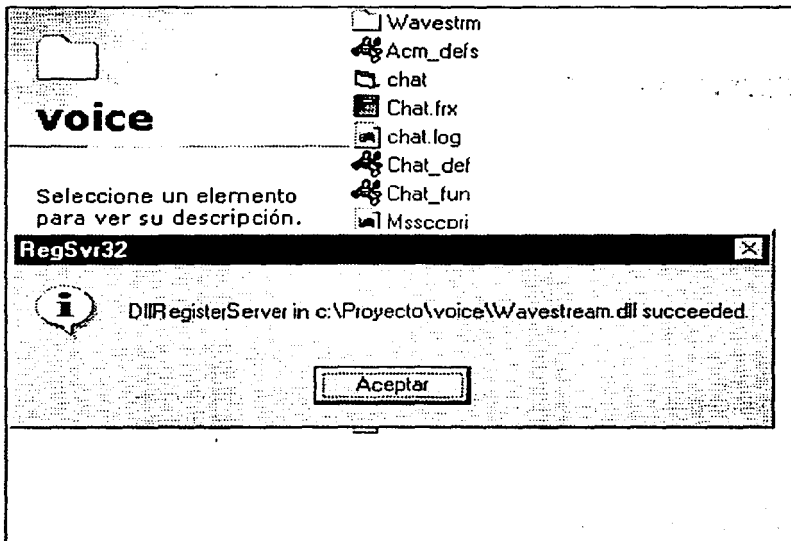


Figura 7.16

Ahora ya que esta dada de alta la librería podemos ejecutar el programa.

## Ejecución del programa

Volvamos a la carpeta en la que se descomprimió el archivo Voz.zip, para nosotros es la carpeta "Proyectos", aquí podremos ver un archivo que tiene un icono en forma de mundo llamado "Voz", este archivo es el ejecutable del programa, sólo hay que dar doble click sobre él para comenzar a operar el programa (figura 7.17).

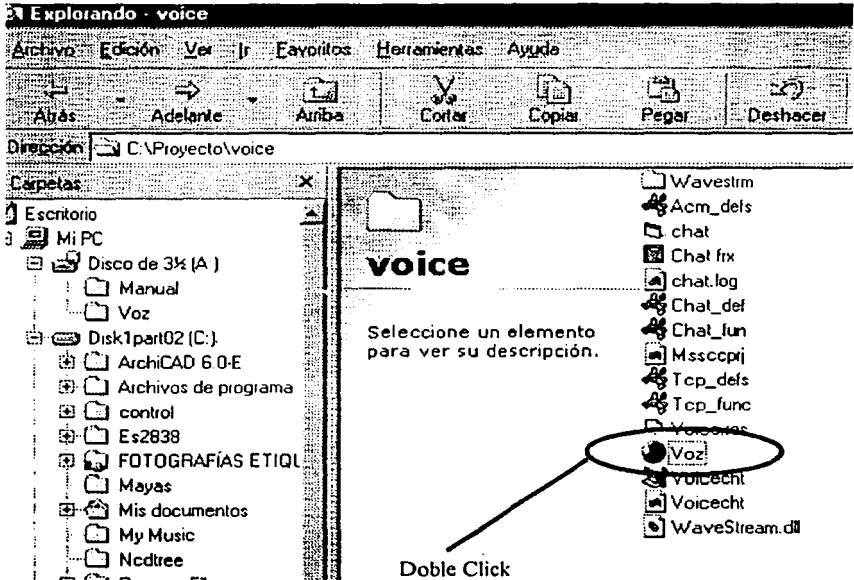


Figura 7.17

La imagen que se observa después de dar doble click sobre el icono "Voz.exe", es la interfaz del programa de transmisión de voz. Para familiarizarnos con dicha interfaz, a continuación en la figura 7.18 se describen la función de los controles que aparecen en ella.

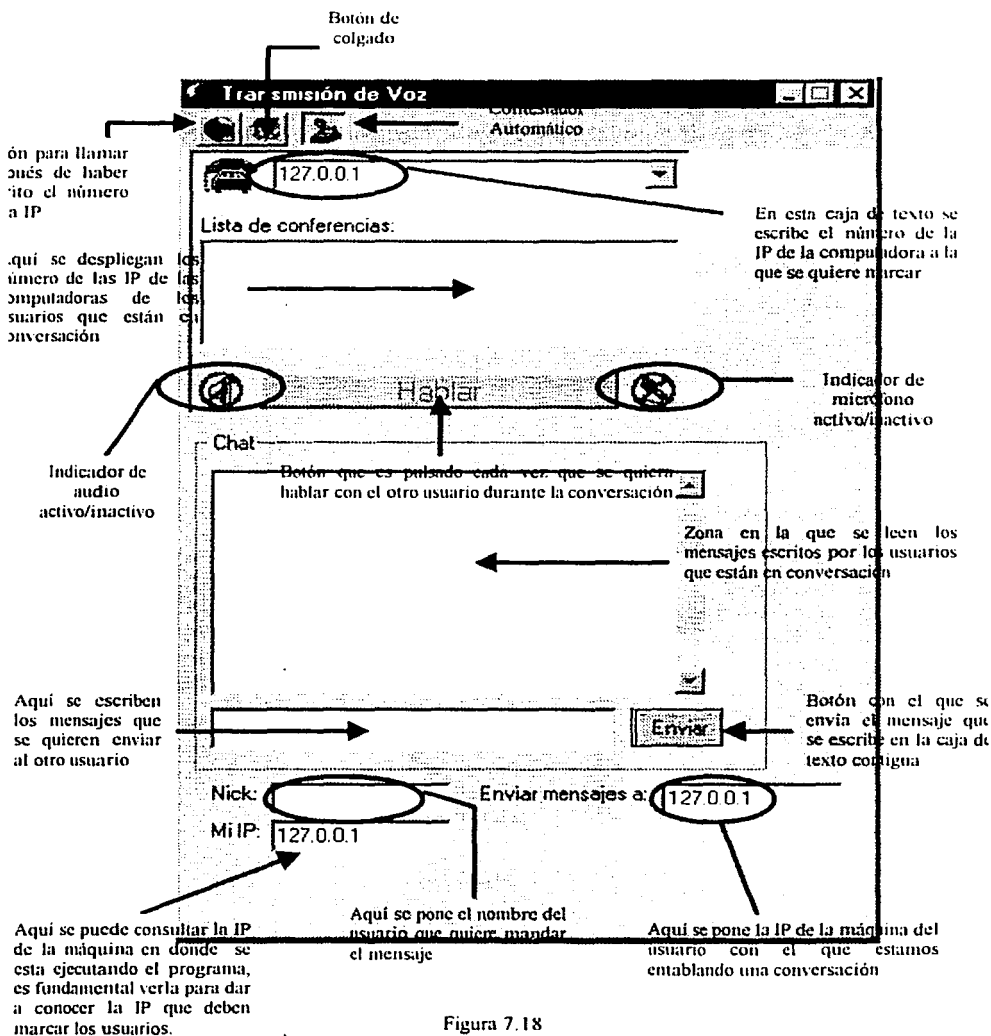


Figura 7.18



---

## Uso del Programa

Después de ver la interfaz y reconocer los controles que conforman el programa de transmisión, dará los pasos a seguir para poder entablar una comunicación.

1. Primero debe establecer una conexión con su proveedor de Internet para poder realizar la transmisión.
2. Abra el programa "Voz" como ya se indico con anterioridad y verifique la IP que tiene en la leyenda de la interfaz "Mi IP", este paso es fundamental porque usted dará a conocer ese número a las personas con las que quiere entablar una comunicación.
3. Para dar a conocer la IP de su máquina, será necesario que estén enlazados todos los participantes de la asesoría a través del servicio Messenger de yahoo o hotmail, porque este servicio de conferencia por texto, nos permite dar a conocer al mismo tiempo todos los números de las IP de los usuarios que estarán participando en la asesoría.
4. Ya teniendo el número de la IP del usuario con el que desee entablar comunicación, solo resta escribirla en la caja de texto que esta contigua a la imagen de un teléfono, finalmente en la barra de menú se pulsa el icono "Llamar".
5. En este momento estará marcando al usuario con el que desea hablar (aquí el programa del otro usuario comenzará a emitir un sonido de llamado igual al del teléfono convencional), si el usuario al que esta llamando le responde (en este caso el usuario si desea tomar la llamada tendrá que pulsar el botón "Hablar" para iniciar la conversación) comenzará usted a escucharlo.
6. Para que usted pueda responder al mensaje, deberá de igual forma pulsar el botón "Hablar", sólo cuando vaya a responder a otro usuario.
7. En caso de que deseen entablar comunicación por texto (chat), deben volver a escribir el mismo número de la IP que marcaron al inicio de la transmisión además deberán escribir su nombre en el recuadro cuya leyenda dice "Nick". Posteriormente deben escribir los mensajes en la caja de texto que está al lado izquierdo del botón enviar. Finalmente sólo tienen que pulsar el botón "Enviar".

8. Para finalizar la conversación se tiene que pulsar el icono de la barra de herramientas "Colgar" o dar doble click en la dirección IP que se muestra en el cuadro de lista de usuarios, de esta manera usted podrá entablar otra conversación con otro usuario, siguiendo el mismo procedimiento antes descrito.

**Notas importantes:**

Es fundamental conocer la IP de la máquina en la que trabaja, ya que si el número que da a conocer es erróneo, no podrán establecer comunicación con usted.

El modo de operar del programa se parece a los Walkitokies, por lo que un usuario puede hablar a la vez pulsando el botón "Hablar".

Cuando se finaliza la conversación, deben forzosamente pulsar ambos usuarios el botón de la barra de herramientas "Colgar" o dar doble click sobre las direcciones IP que se despliegan en el cuadro de lista de usuarios, de lo contrario no podrán establecer comunicación posteriormente con otros usuarios.

## 7.2 Resultados Experimentales

Después de haber desarrollado un modelo de comunicación a través de software y hardware, y llegar a su implementación real, se han obtenido resultados interesantes que es conveniente describir a detalle para dar una conclusión final a este proyecto.

Primeramente las pruebas se han dividido en dos fases:

- Las pruebas sobre red Lan
- Las pruebas sobre red mundial (Internet)

La decisión de realizar estas dos pruebas en ambas redes, tiene la finalidad de observar el desempeño del programa en una red local que se basa en tarjetas de red para su comunicación entre PC's, y por otro lado ver el desempeño con marcación a la red mundial a través de un módem. Hay que recordar que el programa realiza un marcado a un número IP que tiene una computadora, por lo que el programa se puede implementar para cualquier tipo de red ya sea LAN, WAN o MAN.

Así mismo, hay que tener presente que los diplomados se imparten en el extranjero, por lo que el punto de conexión entre ambas entidades se debe de realizar por Internet, sabiendo previamente la dirección IP de la máquina con la que se efectuará el enlace.

### Resultados a pruebas sobre red LAN

Las condiciones bajo las cuales se realizaron las pruebas sobre esta red fueron:

#### 1ª. Prueba

Equipos conectados en una topología tipo bus a través de cable coaxial (figura 7.19); tarjetas de red 10 Mbps PCI Adapter, full duplex; tarjetas de sonido ISA de 16 bits y PCI de 32 bits ambas full duplex, además tarjeta de sonido que soporta la tecnología XP con controladores digitales; procesadores Pentium I a 70 Mhz., AMD K6-2 a 500Mhz., Pentium III a 450 Mhz., Celeron a 700 Mhz. Y Pentium IV a 1.6 Ghz; sistemas operativos Windows 98 segunda edición, Windows Milenium y Windows XP.

Se establecieron estas condiciones de trabajo para ver el funcionamiento de la transmisión para diferentes equipos que cubren y no cubren nuestros estándares preestablecidos como mínimos y óptimos. Así mismo como el programa lo exige, son de

conocimiento en la red LAN todas las direcciones IP de las máquinas que participan en la prueba, por lo que el marcado de una máquina a otra para efectuar un enlace es sencillo.

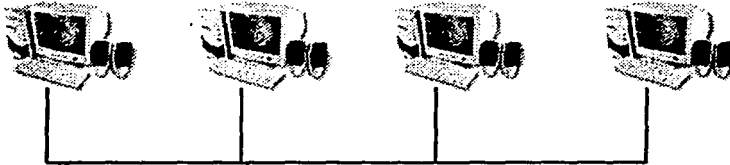


Figura 7.19 Diagrama de conexión de topología bus en una red LAN utilizando cable coaxial

### Resultados

Se ha visto que el mejor desempeño para la realización de la transmisión lo tienen los equipos que están a la vanguardia tecnológica, como es el caso de los procesadores Pentium IV, ya que por su alta capacidad en el desarrollo de procesos nos proporcionan compresión y descompresión de la información transmitida en tiempos mínimos, lo que origina una calidad de audio aceptable, con lo cual no se escuchan cortes o interrupciones en el audio.

A su vez, el equipo que no cumple con los requisitos mínimos recomendables tuvo un gran desempeño en la transmisión, ya que se le introdujo al sistema una tarjeta de audio de 32 bits de alta tecnología para transmisiones de videoconferencia y un micrófono unidireccional que capta los sonidos que llegan directamente de frente, con lo que se compensó su deficiencia de hardware.

Del resto de los equipos, sólo cabe decir que han dado una buena transmisión que es lo que se esperaba de ellos, porque cumplen con nuestros requerimientos al 100%.

La transmisión sobre cable coaxial es buena, no ha presentado fallas y se considera finalmente un buen canal de comunicación que transmite a una velocidad de 10 Mbps y es recomendable para este proyecto, pero en la realidad este cable coaxial se ha

sustituido por el UTP, por ser más económico y fácil de instalar en redes de gran tamaño, como es el caso de edificios en que el cableado de la red esta basado en su mayor parte en UTP.

### 2ª Prueba

Esta prueba se ha realizado bajo las mismas condiciones que la anterior; la única diferencia es que se ha usado una topología de red en estrella (figura 7.20); es decir, las computadoras han estado conectadas a un punto central en donde se encuentra un concentrador. En esta red se ha utilizado cable UTP.

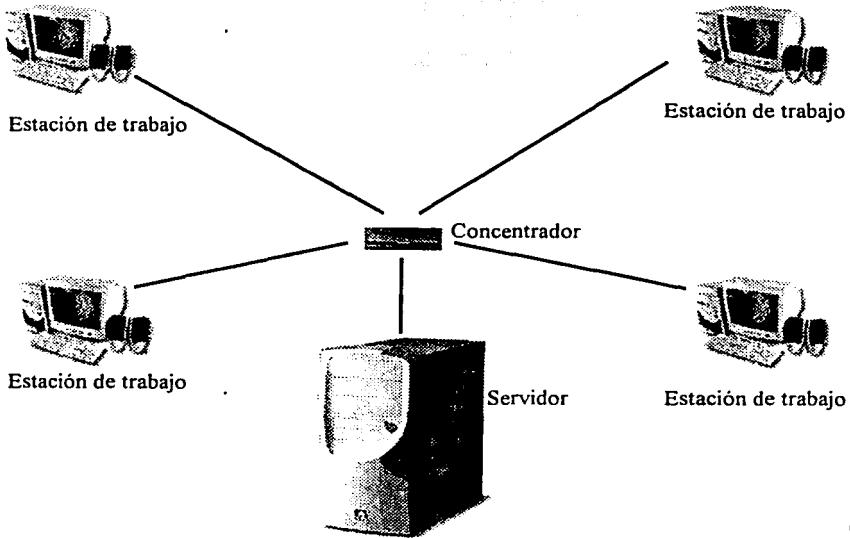


Figura 7.20 Diagrama de conexión topología estrella utilizando cable UTP en una red LAN

### Resultados

Se han tenido satisfactorios resultados en esta transmisión. La calidad del audio ha mejorado por la razón de que el cable UTP tienen una capacidad de transmisión de 1-100 Mbps, lo que ocasiona que se tenga un mayor flujo de información por el canal de comunicación. Los comentarios hechos a los equipos de cómputo en el apartado anterior de la 1ª prueba son los mismos para esta prueba. Como se ha mencionado, el único factor que nos ha ayudado a una mayor eficiencia es el cable UTP.

### Pruebas realizadas sobre la red mundial (Internet)

Esta prueba tienen características totalmente diferentes a las pruebas realizadas sobre la red LAN, por los siguientes motivos:

- La conexión a la red mundial se puede hacer por módem.
- La conexión a la red mundial se puede realizar por la red local utilizando las configuraciones respectivas de puertas de enlace (Gateways).
- Se necesitan en ambos casos saber con exactitud el punto IP de la máquina a marcar.

Para el desarrollo de esta prueba nos enfrentamos a un problema, saber el punto IP de la máquina a marcar cuando ésta no se encuentra dentro de una red LAN.

Actualmente los proveedores que brindan el servicio de Internet ya no dan a los usuarios una dirección IP fija, por lo que el segmento de direcciones IP que tiene el servidor se asignan aleatoriamente cuando el usuario realiza una petición de conexión; es decir, nunca tienen una dirección IP fija. Por el contrario si una máquina realiza una conexión a través de la red local, siempre se puede saber la IP que realiza la conexión a Internet y en tal caso no hay problema alguno.

Como el programa que se ha desarrollado debe adaptarse a los entornos de trabajo a los que se enfrenta, hemos de dar solución a este pequeño problema que es determinante para realizar un enlace.

Las acciones realizadas para dar solución a este problema fueron:

1. Dentro del programa que realiza la conexión se ha implementado una función que al momento de ejecutarse nos despliega en un encabezado la dirección IP que tiene la máquina después de realizar la conexión con Internet. Este encabezado muestra un conjunto de cuatro números separados por puntos, el cual hace referencia a la dirección IP que le asignó el servidor en esa conexión.
2. Al saber la dirección IP de las máquinas cuando éstas no se encuentran en una red local, lo que hace falta en el caso del instructor es, darla a conocer a los alumnos del diplomado; esto se puede realizar a través de dos modos:
  - a. De un correo electrónico en el que se indique la dirección IP del instructor, la cuál los alumnos deberán de marcar para realizar un enlace.
  - b. Conectándose ya estando en Internet a un programa como el Messenger de Yahoo, ya que todos los usuarios al momento de contactar a una persona dentro de la red marcan la misma dirección IP del servidor al entrar a ese programa.

Analizando las dos posibles opciones tenemos los comentarios siguientes:

- Un correo electrónico como medio para informar la IP del instructor es una buena opción, pero se corre el riesgo de que los alumnos no lo vean a tiempo o que no puedan acceder en ese instante a su cuenta de correo; además, el instructor deberá estar conectado tiempo antes de que se lleve a cabo la asesoría para mandar a tiempo ese correo electrónico, por lo que esta opción se toma como una medida de seguridad y alternativa para dar a conocer la dirección IP de interés.
- La conexión a un servidor intermedio es la opción más viable y segura para dar a conocer la IP que tiene el instructor (figura 7.21 y figura 7.22), ya que estando los alumnos dentro del programa Messenger, se informa a todos su dirección IP y desde ese momento incluso se pueden determinar los turnos en los que irán pasando los alumnos y la duración que tiene la asesoría. Todo esto se lleva a cabo en tiempo real, ya que el Messenger tiene un chat incluido que permite la comunicación instantánea.

Así hemos resuelto este problema que surge dentro de la realización de la prueba sobre la red mundial (Internet).

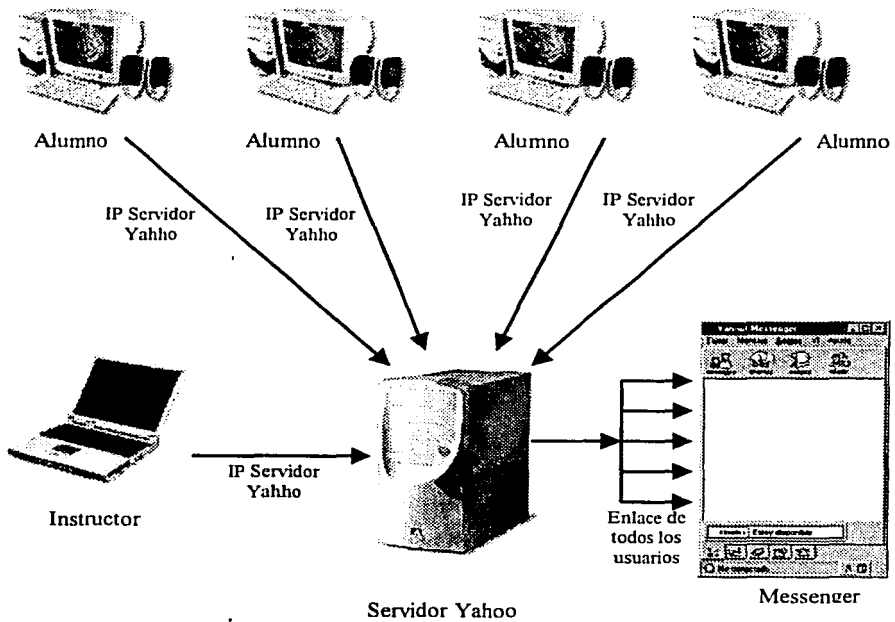


Figura 7.21 Diagrama del comienzo del enlace al servidor Yahoo que contiene el servicio del programa Messenger. Todos lo usuarios marcan la misma IP al ingresar al programa Messenger.



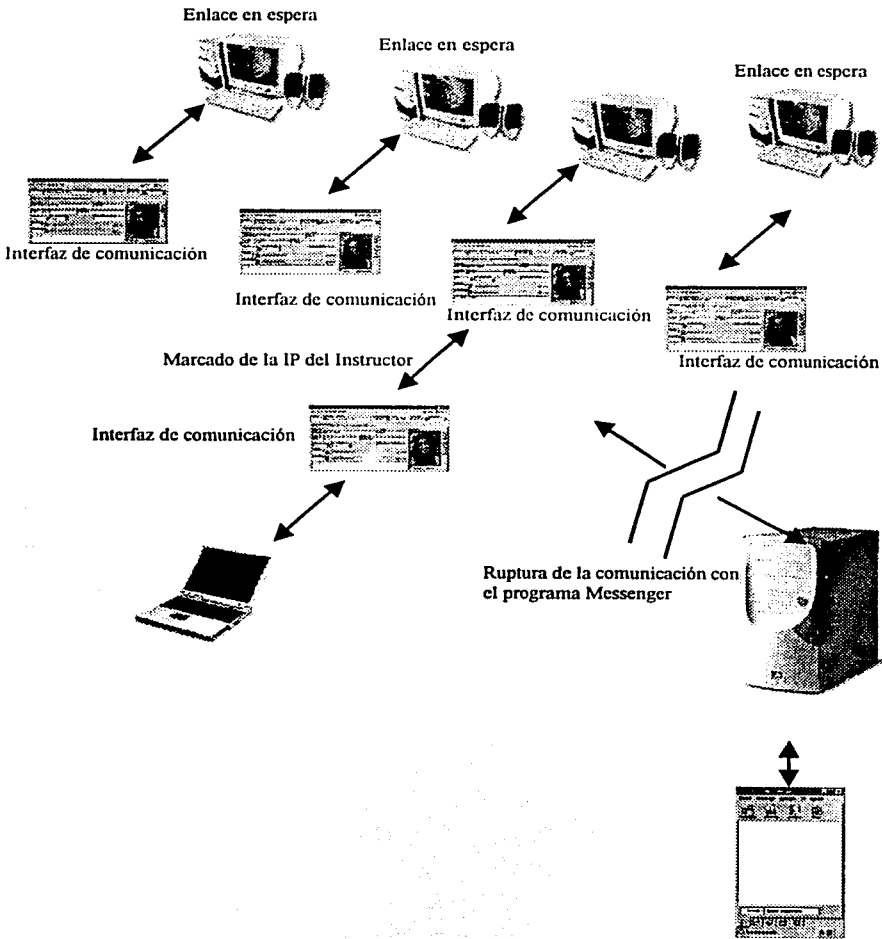


Figura 7.22 Diagrama en la que se comunica la dirección IP del instructor y comienza el marcado de esa dirección por los alumnos para llevar a cabo la asesoría. En este momento se hace ruptura con el enlace del programa Messenger.

## Resultados

Los equipos utilizados fueron de las mismas características que los usados en las pruebas sobre red LAN, por lo que podemos decir que fue una transmisión aceptable, en la que se logró enlazar las computadoras a través de la red mundial por la vía de los módems. Cabe mencionar que la triangulación en la que había un servidor de por medio al iniciar el enlace fue una solución realmente acertada, ya que nos evita mandar un correo electrónico que no sabemos si llegará.

De manera general se considera que los resultados experimentales bajo las condiciones con las cuales se desarrollaron han sido satisfactorios y han cubierto con las expectativas que se han planteado desde el inicio del desarrollo de este proyecto. Así es como se ha logrado tener un enlace a nivel mundial para dar asesorías a bajo costo sin requerir complejos sistemas de videoconferencias, que no pueden ser solventados en la mayoría de los casos por la falta de recursos o infraestructura destinados al desarrollo de un proyecto.

# CAPÍTULO 8

## CONCLUSIONES

---

## CAPÍTULO 8

### Conclusiones

Después del desarrollo de este proyecto, me he dado cuenta que siempre hay varias soluciones para la solución de un problema, de tal forma que el análisis y la investigación son factores importantes para determinar de todas las soluciones posibles, cuál es la más conveniente y eficaz para resolver a plenitud el problema planteado.

Es así desde el primer momento en que se presenta la necesidad de dar asesoría a los diplomados a través de algún medio que fuese eficiente, que se piensa en los sistemas más complejos de transmisión de audio y video como lo son la videoconferencia con los múltiples aditamentos que ella requiere y el Internet satelital que de igual forma requiere de equipos especiales de gran desempeño para las transmisiones, por lo que ambas propuestas no se pueden costear para un proyecto que esta pensado a largo plazo, ya que el uso del satélite y las horas de transmisión en el caso del Internet satelital son en su conjunto de precios elevados, que no pueden ser costeados con los recursos e infraestructura que se tiene destinados para este proyecto. Otro punto importante radicaba en el que la solución escogida cubriera los objetivos y expectativas que se habían planteado desde el inicio como son:

- Inversión a largo plazo.
- Crecimiento del sistema para adaptarse a las necesidades del medio.
- Comunicación unidireccional o bidireccional con optimización de recursos.

Teniendo estos objetivos como generales, se da una nueva solución al problema y radica en la teleconferencia, sistema de comunicación que funciona con los mismos principios de Direc TV o Sky, el cual resulta económico comparado a los dos antes mencionados, pero no cubre los objetivos originales: tener la capacidad de realizar una comunicación unidireccional o bidireccional aunado a que era de vital importancia en la renta de un canal de transmisión, tener una programación continua como si fuese un canal de televisión comercial. La respuesta de la compañía que brinda el servicio de teleconferencias "EDUSAT", resolvía el problema de la comunicación bidireccional con el Internet, solución válida pero no efectiva, por lo que fue descartada esta solución.

Es aquí cuando la investigación y el análisis se hacen presentes para dar una solución certera que resuelva la problemática que se tiene de comunicación.

Por tal motivo dentro de esta ardua tarea de investigación, surgen soluciones que pueden ser para muchos las más sencillas con la creencia de no poder resolver o satisfacer los objetivos de un proyecto que dé solución al problema planteado. Aun con esto, creo yo que estas soluciones, si se saben llevar con una metodología y un plan de trabajo estratégico son las más eficaces y flexibles, para poder ser adaptadas en cualquier medio de trabajo, por la sencilla razón de que las soluciones que se han comercializado requieren un estándar de trabajo rígido, es decir, requieren de la compra de todo una serie de dispositivos costosos y complejos para que el proyecto se desarrolle de manera eficaz, dando como resultado a largo plazo un sistema que tarde o temprano quedará obsoleto por la constante revolución tecnológica que se vive en la actualidad.

Pero, si tenemos una solución que tenga elementos que se puedan modificar constantemente conforme avanza la tecnología, es una inversión a largo plazo que nunca tendrá pérdidas, pues la propia tecnología ayudará a que se realice cada vez mejor el desarrollo del proyecto.

Con todo lo antes mencionado, es como tomé la decisión final previo una investigación y análisis, de emprender un proyecto que fuese flexible y que pudiese ser mejorado conforme la tecnología y los recursos lo permitieran, dando como resultado el desarrollo de un sistema de comunicaciones que se basa en un programa con interfaz amigable al usuario que es capaz de comunicar dos entidades geográficas diferentes a través del Internet.

Como principio fundamental y recordando el inicio de la era de las comunicaciones en las que sólo se transmitía voz y audio, los objetivos en particular en el desarrollo de este proyecto son:

- La transmisión de sólo voz y audio a través del Internet.
- Comunicación unidireccional o bidireccional.
- Optimización de recursos para conseguir una transmisión a bajo costo.

A partir de estos objetivos se generó el desarrollo del proyecto, el cual condujo a una nueva investigación de bastantes elementos que fueron conformando poco a poco lo que hasta el final fue un programa de comunicación que realiza la conexión entre dos computadoras

que pueden estar en el mismo lugar geográfico o no, las cuales son capaces de tener una comunicación semejante a la del teléfono convencional.

Algunos elementos en los que se ha fundamentado el desarrollo del proyecto son la programación de Winsocks que permiten la comunicación de procesos dentro del Internet; el Streaming que permite la descarga y ejecución al instante de la señal de voz y audio; la compresión y descompresión de la señal de voz y audio a través del algoritmo de Huffman, que al conjuntar a todos éstos e implementarlos a un lenguaje de programación, da como resultado, un programa en la que intervienen funciones ya preestablecidas que desempeñan trabajos muy específicos como son la digitalización, codificación, compresión y transmisión de la señal de voz y audio.

Así, en el desarrollo de este proyecto, se han tenido que ir resolviendo algunas incógnitas y dudas que han sido solamente esclarecidas cuando se han hecho las pruebas. Un ejemplo de ello es la interrogante que se abrió en las pruebas realizadas sobre una de las redes, en la que se enfrentó una problemática que no se había contemplado, pues era de vital importancia el conocimiento de la dirección IP del instructor en cada sesión. Si éste no realizaba la conexión de Internet por medio de una red local. Para dar solución a dicha interrogante se tuvo que realizar un análisis que nos permitiera conocer y dar a conocer esa dirección IP que se necesitaba. El resultado final, la incrustación de una función dentro del programa que nos permite saber después de conectarse al Internet la dirección IP que nos ha asignado el servidor que nos provee del servicio y la intervención de un servidor intermedio (punto de enlace en la primera fase de conexión) que permite enlazar a todos los alumnos en un mismo sitio (hablamos del programa Messenger en el que se contactan a personas sobre la red) por el cual se les comunica la dirección IP que de antemano ya es conocida, para finalmente mediante la marcación de ésta realizar un enlace de audio de forma privada por medio del programa que se ha desarrollado.

Las pruebas de las que se habla en el apartado anterior se han ejecutado en redes de área local y la red mundial (Internet), para que de este modo el programa se pueda implementar sin ninguna dificultad en cualquier medio de trabajo y así tener un parámetro con el que se pueda medir su desempeño.

La prueba en red local resultó satisfactoria en ese entorno de trabajo. En esta ocasión, a propósito se hizo intervenir equipos con características que no cumplieran con los requisitos

mínimos, ya que cuando se implemente el sistema no sabemos desde donde lo pueda ejecutar el usuario. En pocas palabras, quizá no lea los requerimientos mínimos y realice todo lo contrario. Pero aun teniendo estas condiciones de trabajo, las pruebas fueron exitosas, ya que un equipo con tarjeta de audio de 32 bits y un buen micrófono unidireccional producen una magnífica calidad de voz y audio.

Por su parte las pruebas desarrolladas sobre la red mundial (Internet) han sido de igual forma satisfactorias, teniendo igualmente la intervención de equipos que no cumplen con los requisitos marcados como mínimos. Pero ya no tiene caso decir, que las computadoras que en la actualidad están con la tecnología a la vanguardia resultan formidables para realizar una transmisión de esta índole, ya que sus componentes están hechos en la mayoría de los casos para la completa navegación por Internet.

El resultado final que se tuvo fue una transmisión con calidad telefónica semidireccional, ya que habla sólo una persona a la vez, y sobre todo con recursos que se han optimizado al máximo, porque sólo un programa, una tarjeta de audio, un micrófono y un buen proveedor de servicios de Internet hacen posible el buen funcionamiento de una transmisión inteligible para los usuarios, por lo cual los objetivos propuestos en este proyecto han sido cumplidos.

# APÉNDICE A



## Apéndice A

Este apartado tiene como finalidad analizar los diferentes protocolos de comunicación que utilizan los equipos que se emplean en las videoconferencias.

Debido a la necesidad que tuvieron las empresas de comunicarse de una manera más eficiente con aplicaciones multimedia, fue necesaria la creación de protocolos de comunicación estándar que controlaran y permitieran el aprovechamiento de las herramientas multimedia (voz, video y datos). Como ejemplos de estos protocolos de videoconferencia tenemos: H.320, T.120, H.324 y H.323, los cuáles fueron propuestos por la Unión Internacional de Telecomunicaciones (UIT-T); permitiendo el establecimiento de videoconferencia de alta calidad, sobre ISDN ó ATM (estándar H.320) ó videoconferencia en las que no es necesario cumplir con altas exigencias de calidad como es el caso de los estándares H.323 y H.324. Así mismo también es posible establecer una videoconferencia de muy alta calidad (H.310) para aplicaciones muy especiales. Por supuesto, que la elección de una u otra depende de la calidad de servicio deseado por el grupo de usuarios.

Estos protocolos de comunicación permiten interconectar todos los dispositivos integrantes de una red para poder llevar la información del transmisor al receptor. En el caso de la videoconferencia, los protocolos de comunicación son los encargados del funcionamiento en conjunto de manera correcta y ordenada de los elementos que conforman el sistema total. Dentro de estos elementos se pueden encontrar el codec, los monitores, las cámaras, las bocinas, los micrófonos, etc.

### *Análisis de H.320*

El H.320 describe normas para la videoconferencia punto a punto y multipunto en las Redes Digitales de Servicios Integrados ISDN. La tecnología H.320 requiere típicamente redes separadas para la transmisión del video y los datos. Esto supone doble cableado y un incremento en el costo de implantación del sistema. Las características de dicho protocolo son las que a continuación se mencionan:

- Tiene dos resoluciones de imagen: QCIF (Quarter Common Intermediate Format) proporcionando una resolución de imagen de 352 píxeles por 288 líneas y CIF

(Common Intermediate Format) que maneja una resolución de 176 píxeles por 144 líneas.

- Soporta velocidades de 7.5, 10, 15 y 30 cuadros por segundo.
- Especifica tres compresores de audio. G.711 banda estrecha con velocidades de transmisión de 48 a 64 kbps; G.722 banda amplia con velocidades de transmisión de 48 a 64 kbps; G.728 banda estrecha con velocidades de transmisión de 16 kbps.
- La velocidad de transmisión se ajusta a la empleada como estándar que es de 384 kbps.

En la figura A.1 se muestra el diagrama de conexiones de una videoconferencia ISDN con el protocolo H.320.

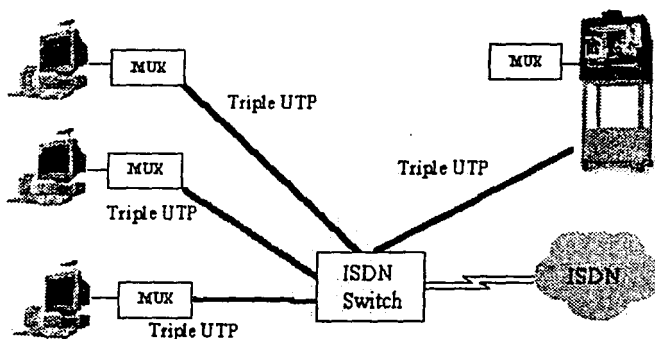


Figura A.1 Esquema de conexión de una Videoconferencia utilizando el protocolo H.320

Se han realizado diferentes pruebas que soportan este protocolo en equipos de videoconferencia para sistemas de grupo y de escritorio, los cuales se mencionan en la tabla A.1:

<i>Sistemas de Grupo</i>		<i>Sistemas de Escritorio</i>	
Equipo	Fabricante	Equipo	Fabricante
Venue 2000-50/30	Picture Tel	LivePro	Picture Tel
Concorde 4500ZX	Picture Tel	Smax	Vtel
PC 5100	Sony	WG-5000	Vtel
PC 3000	Sony	C-Phone WS	CPhone
PC 2000	Sony	Intel ProShare	Intel
LC 3000	Vtel		
TC 2000	Vtel		
Media Max	Vtel		
FRED	Vtel		
Radiance	CLI		
Rembrandt	CLI		
	Phillips		
	Tanderm		
	GTP		

Tabla A.1 Equipos de videoconferencia que soportan el protocolo H.320

Las principales diferencias que existen en los equipos son: el costo, configuración, operación y soporte.

### *Análisis de H.323*

Este protocolo de comunicación fue definido por la ITU-T para la comunicación multimedia a través de redes de área local. Fue diseñado específicamente para tecnologías que no proporcionan una calidad de servicio garantizado (como pueden ser las tecnologías TCP/IP e IPX corriendo sobre redes Ethernet, Fast Ethernet y Token Ring).

El H.323 es un estándar bastante flexible, ya que es independiente del hardware y del sistema operativo; además su diseño permite la compatibilidad de la red

independientemente de su topología y su interacción en cualquier ambiente basado en el protocolo de Internet IP.

Tiene la capacidad de codificar y decodificar audio de acuerdo al algoritmo G.711, aun cuando carece de la capacidad de codificar y decodificar video; además, si están presentes las aplicaciones que comparten datos deben cumplir con el estándar T.120.

El estándar T.120 define un servicio de comunicación de datos para uso en ambientes de conferencia multimedia (audio, video y datos). T.120 proporciona la facilidad de establecer y manejar comunicaciones interactivas (conferencias en las que se comparten datos).

Su característica fundamental es: permitir el compartimento de datos entre los participantes de una videoconferencia a través de las redes ISDN, PSDN, CSN, TCP/IP y IPX LAN.

#### *Análisis de H.324*

La recomendación H.324 proporciona los protocolos para el funcionamiento en líneas telefónicas analógicas de módems de alta velocidad tipo V.34, cubriendo de esta forma la transmisión en tiempo real de video, audio y datos o cualquier combinación de éstos, dando como resultado un sistema económico basado en computadora personal.

H.324 proporciona estándares como el G.723 de codificación de audio de 5.3 a 6.4 kbps y el H.263 de codificación de video.

Los equipos que resultaron compatibles con el protocolo H.324 después de realizar varias pruebas en ellos son: los equipos Modular Videophone UVC-50 y Videoteléfono C-Phone.

Gracias a estos protocolos de comunicación de videoconferencia, ha sido posible el desarrollo de múltiples proyectos educativos y laborales, que han permitido enlazar países y continentes con la única finalidad de reunir en un espacio a millones de gentes con un sólo objetivo, "el de comunicarse".

## **BIBLIOGRAFÍA**

---

## Bibliografía

Faúndez Zanuy Marcos, *Tratamiento digital de voz e imagen*, México: Alfaomega, 2001, págs. 271

Cornella Alfons, *Información didital para empresa*, México: Alfaomega, 1997, págs. 180

Peasley Richard, *Descubre Visual Basic 6*, España: Prentice Hall, 1999, págs. 673

Charte Francisco, *Programación con Visual Basic 6*, España: Anaya, 1998, págs. 771

Appleman Dan, *Desarrollo de componentes COM/Actives con Visual Basic 6*, España: Prentice Hall, 2000, págs. 801

### Internet

*Visual Basic Internet Programming – MS Winsock Control – Getting Started, USA:*  
[www.monografias.com](http://www.monografias.com), 2000.

*Client Server Programming with Winsock – Winsock Properties & Events, USA:*  
[www.lawebdelprogramador.com](http://www.lawebdelprogramador.com), 2001.